Shinsuke Ikeda · Hideaki Kiyoshi Kato
Fumio Ohtake · Yoshiro Tsutsui *Editors*

Behavioral Interactions, Markets, and Economic Dynamics

Topics in Behavioral Economics

Springer

Behavioral Interactions, Markets, and Economic Dynamics

Shinsuke Ikeda • Hideaki Kiyoshi Kato
Fumio Ohtake • Yoshiro Tsutsui
Editors

Behavioral Interactions, Markets, and Economic Dynamics

Topics in Behavioral Economics

Springer

Editors
Shinsuke Ikeda
Institute of Social and Economic Research
Osaka University
Ibaraki, Osaka, Japan

Hideaki Kiyoshi Kato
Graduate School of Economics
Nagoya University
Nagoya, Aichi, Japan

Fumio Ohtake
Institute of Social and Economic Research
Osaka University
Ibaraki, Osaka, Japan

Yoshiro Tsutsui
Faculty of Economics
Konan University
Kobe, Hyogo, Japan

ISBN 978-4-431-55500-1 ISBN 978-4-431-55501-8 (eBook)
DOI 10.1007/978-4-431-55501-8

Library of Congress Control Number: 2015950212

Springer Tokyo Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer Japan KK is part of Springer Science+Business Media (www.springer.com)

Preface

For the purpose of providing new and broader directions for the future development of behavioral economics and finance, this book collects important contributions in behavioral economics/finance and related topics among journal publications of Japanese researchers to date. By applying new insights from behavioral economics/finance, we are interested in extending the reach of the standard theories in our own fields. A project to edit readings and/or handbooks on behavioral economics/finance for the promotion of economic research came about naturally as a result of our frequent interactions when running academic meetings on behavioral economics, especially those of the Association of Behavioral Economics and Finance (ABEF), the Japanese Economic Association (JEA), and the Nippon Finance Association (NFA). In addition, these meetings gave us access to important works that were motivated by behavioral economics. We therefore have compiled and edited a couple of independent volumes in an attempt to capture the many worthy articles that lie within this topic. The first, titled *Behavioral Economics of Preferences, Choices, and Happiness*, focuses on works on behavioral economics; and the second, *Behavioral Interactions, Markets, and Economic Dynamics: Topics in Behavioral Economics*, on economics-oriented studies on topics in behavioral economics. This book is the latter.

Three features characterize the present book. First, it focuses on economic studies examining the interactions of multiple agents or market phenomena using behavioral economics models. As current behavioral economics models are not necessarily good at analyzing phenomena from the viewpoints of market equilibrium and agent interactions, this feature of the book will help readers consider new possibilities for behavioral economics models as well as for general economic models. In contrast, the other book focuses on more behavioral, single-agent issues, such as decision making, preference formation, and subjective well-being. The two books thus are complementary.

Second, the chapter authors have added newly written addenda to the original articles, in which they discuss their own subsequent works, and provide supplementary analyses, detailed information on the underlying data, and/or recent literature surveys. The addendum of each chapter is based on discussion at the Development of Behavioral Economics and Finance Conference held in February 2014. During this conference, participants, including the authors of the book chapters, discussed the original studies to be included in these volumes in light of contributions, limitations, and implications for future research developments. We accordingly believe that this work creates a bridge between the original studies and future research development.

Third, reflecting the diverse fields of the editors, this book as well as the companion volume, captures broad influences of behavioral economics on various topics in economics. The topics of this book cover parental altruism, economic growth and development, the relative and permanent income hypotheses, wealth distribution, asset price bubbles, auctions, search, contracts, personnel management, and market efficiency and anomalies in financial markets. The remainder of this preface provides a brief introduction to the parts of the book.

Part I is composed of two chapters that address intergenerational interactions under parents' altruism. In Chap. 1, Professor Hideo Akabayashi develops a unique dynamic principal-agent model to endogenously describe a child's development, his time preference formation, and the parents' interventions under asymmetric information. Akabayashi successfully explains child maltreatment by parents as an equilibrium outcome under their divergent misbeliefs about the child's ability. He also characterizes families that are at risk of child maltreatment. In Chap. 2, Professors Vipul Bhatt and Masao Ogaki propose another model of parents' strict intervention behavior toward their children. Unlike Akabayashi, they assume perfect information and thereby focus on a positive aspect of parental intervention in the form of "tough love," where the parent in their model allows the child to suffer in the short run via lower childhood transfers (e.g., allowances) so that she grows up to be more patient in the long run. The authors also extend the model to account for the child's leisure choice to emphasize the distinction between exogenous and endogenous changes in income when examining the redistributive neutrality property of altruism models.

Part II begins with important research by Professors Hiroaki Hayakawa and Yiannis Venieris in Chap. 3, which was originally published in the *Journal of Political Economy*. In 1977, when the field of behavioral economics had not yet appeared, they made contributions that are behavioral-economics oriented. First, they address heuristic cognition-saving decision making under bounded rationality. Second, they focus on the critical role of social interdependence in endogenous preference formation. The authors describe the consumer behavior that identifies with and emulates a chosen reference group for heuristic decision making. In doing so, they derive indifference curves under social interdependence based on two axioms and four basic assumptions. The implications for consumer theory too are discussed. In Chap. 4, Professor Hayakawa further extends the ideas in the previous chapter by presenting an axiomatic theory for the analysis of boundedly rational consumer choice. To describe heuristic decision making, the author focuses on the important roles of social norms and reference groups as sources of low-cost heuristics and proposes a model of a sequential two-step choice making procedure to satisfy physical and social wants. Classical theories of consumption externalities developed by Leibenstein, Veblen, and Duesenberry are re-interpreted using the proposed framework. In Chap. 5, Professors Koichi Futagami and Akihisa Shibata address the effect of consumers' status/wealth preferences on endogenously determined steady-growth rate. When consumer preferences are personally interdependent due to status preferences, effective time preferences are shown to depend on relative wealth holdings producing rich, and sometimes paradoxical, implications for growth and wealth distribution. In Chap. 6, Professor Katsunori Yamada provides further macroeconomic implications of status preferences. He develops a capital-accumulation model with two consumption goods for normal and conspicuous purposes in order to characterize the properties of equilibrium dynamics in the bandwagon-type and snob-type economies. The Sombartian oscillating dynamics are duplicated as an equilibrium outcome of the growth-impeding effect of conspicuous consumption. This characteristic is seen particularly in the bandwagon-type economy. Chapter 7, written by Professors Yoshiyasu Ono and Junishiro Ishida, develops a new dynamic behavioral model to describe unemployment due to demand shortage. In this process, two behavioral assumptions are incorporated: workers' concern for fairness, which provides a microfoundation for a behavioral version of the Phillips curve, and the insatiable desire for money, which plays a critical role in producing persistent demand shortage. Monetary and fiscal policies are then evaluated in light of their effectiveness in reducing unemployment in the short and long run.

The four studies in Part III contribute to the literature of time preference in macroeconomics. Chapter 8 is based on the *Review of Economics and Statistics* article written by Professors Masao Ogaki and Andrew Atkeson. The authors examine the empirical validity of the models of wealth-dependent intertemporal elasticity of substitution (IES) and the wealth-dependent rate of time preference (RTP) using panel data from India in which there were large fluctuations in consumption data. By incorporating the subsistence consumption level, the estimation result shows that IES depends positively on wealth, whereas RTP is wealth-independent. In contrast, in Chap. 9 Professor Kazuo Ogawa uses aggregate time-series data of Japan, Taiwan, and Korea to show that the RTP of each country's representative consumer depends on the income level. In particular, he compares the empirical validity of the three alternative RTP schedules—flat, upward, and U-shaped—to show that the RTPs of Japan and Taiwan are characterized by a U-shaped schedule. The estimated turning points in the two countries are found to be consistent with their historical loci of economic growth. Chapters 10 and 11 comprise theoretical contributions to the RTP issue. In Chap. 10, Professor Shinsuke Ikeda extends an endogenous RTP model to characterize luxury and necessity good consumption in terms of good specific RTP and IES. Preferences for luxury are shown to affect capital accumulation and wealth distribution. In Chap. 11, Professors Ken-ichi Hirose and Ikeda examine the implications of decreasing marginal impatience. As is often empirically observed, RTP is decreasing in wealth. The authors show its dynamic implications for stability property, multiple equilibria, and the possibilities of consumption-satiated equilibria.

Part IV analyses bubbles and the ensuing crashes. Chapter 12, authored by Professors Robert J. Shiller, Fumiko Kon-Ya and Yoshiro Tsutsui and published in the *Review of Economics and Statistics*, investigates why the Japanese stock market crashed between 1989 and 1992. To answer this question, they collect parallel time series data on expectations, attitudes, and theories from market participants in both Japan and the United States for the period 1989–1994. Such a survey is unique, especially in the early 1990s. They find a relationship between the crash and changes in both Japanese price expectations and speculative strategies. In Chap. 13, Professors Shinichi Hirota and Shyam Sunder conduct an economic experiment to explore how investor decision horizons influence the formation of stock price bubbles. The experiment consists of long- and short-horizon sessions. These sessions differ by receiving either the determined dividend (the long-session) or the expected future price when the subjects exit (the short-session). They find that price bubbles emerge more frequently in the short-horizon session, suggesting that the difficulty of performing backward induction from future dividends is important to the emergence of price bubbles.

Part V contains three chapters concerning experimental markets. It begins with Chap. 14, which is authored by Professors Soo Hong Chew and Naoko Nishimura. It is well-known that the English and second-price auctions generate the same revenue when bidders have independent private valuations of an auctioned object. That is, both auctions exhibit the *revenue equivalence theorem*. However, if the auctioned object involves risk, the theorem breaks down when bidders are non-expected utility maximizers, since submitting one's valuation is no longer a dominant strategy for them under second-price sealed-bid auctions. In this chapter, the authors experimentally examine whether their subjects have expected utility preferences and, if not, whether they exhibit choices consistent with the Allais paradox. The authors show that the two experimental auction markets do not support the revenue equivalence theorem when they introduce a risky auctioned object. Additionally, the English auction yields higher seller revenue than the second-price auction for the subject pool where the Allais type is predominant, as predicted by the theoretical examination under non-expected utility preferences. In Chap. 15, Professors Yoichi Hizen, Keisuke Kawata, and Masaru Sasaki examine the properties of a committee search, in which a decision is made by a group of multiple agents rather than by a single agent. Recently, Albrecht, Anderson, and Vroman (AAV) theoretically analyzed the properties of decision-making in the case of committee search. However, there exist no empirical studies on committee search, mainly because of the difficulty in collecting suitable data. A unique feature of this chapter is the use of laboratory experiments to collect original data in order to test the AAV's propositions. Specifically, the authors examine the propositions that the average search duration is increasing in the number of votes required to stop committee search and that it is also increasing in the number of group members. Overall, the experimental outcomes are consistent with the implications suggested by the AAV model. Chapter 16 is authored by Professors Toshiji Kawagoe and Hirokazu Takizawa. The authors investigate cheap-talk games with private information using an experiment. They find that when the interests of the sender and receiver are aligned, informative communication frequently arises. While babbling equilibrium play is observed more frequently in conflicting interest cases, a substantial number of players tend to choose truth-telling. In other words, they found over-communication, truth bias, and truth-detection bias, which are not predicted by equilibrium refinement theories. They explain these results using a level-k model, which is a non-equilibrium theory of players' initial responses to games that reflect the strategic thinking of players.

Part VI contains three attempts to extend contract theory by applying the insights of behavioral economics. Chapter 17 is Professor Hideshi Itoh's initial attempt to develop a behavioral contract theory. By incorporating players' other-regarding preferences, such as inequity aversion and status preferences, into the standard moral hazard models of principal-agent relationships, he shows that other-regarding preferences interact with moral hazard in some important ways. For example, a principal is worse off when his agent cares about the principal's income. In the presence of symmetric self-regarding agents, the principal is shown to be able to optimally exploit his agents' other-regarding behaviors by designing contracts appropriately. Further development of behavioral contract theory is surveyed in the addendum of the chapter and found in the two subsequent Chaps. 18 and 19, both of which are written by Professor Junichiro Ishida. In Chap. 18, Ishida incorporates self-esteem concerns as a behavioral motive into a simple principal-agent framework. By specifying the agent as benefiting from having a positive self-image (expected self-attributes), he provides a unique model that describes "self-handicapping" behaviors to withhold effort with the intention of obscuring his own attributes. An important implication is that uncertainty reduces agency costs and thereby increases the effort incentive because uncertainty reduces the need for self-handicapping. In Chap. 19, Ishida again considers a principal-agent model in which the agent does not have perfect knowledge about his innate ability (attributes). When the principal has superior knowledge about the agent's ability and decides whether to promote the agent based on the private information, promotion decisions act as credible signals of the principal's evaluation and have the "looking-glass" effect on the agent's self-confidence. The principal's strategic promotion policy that incorporates the "looking-glass" effect potentially explains why demotions are rare in practice, even when employees' incompetence level increases, a phenomenon otherwise known as the Peter Principle

Part VII contains four chapters on anomalous stock return behavior against market efficiency. In Chap. 20, Professor Takahiro Azuma, Katsuhiko Okada, and Yukinobu Hamuro examine the media's influence on stock returns, focusing on investor behavior surrounding revisions of sell-side analysts' ratings. Azuma et al. find that media-covered stocks show significantly lower post-announcement returns than non-media-covered stocks. A more careful examination of media-covered stocks finds that while downgraded stocks show little difference in post-event returns regardless of the degree of sentiment, upgraded stocks do show a difference. These results are consistent with the view that heavy-media-coverage stocks are overpriced due to individual investors' noise trading. In Chap. 21, Professors Yoshio Iihara, Hideaki Kiyoshi Kato, and Toshifumi Tokunaga document the winner–loser effect in the Japanese stock market. Surprisingly, the well-known stock return regularity that is a characteristic of American and other nations' stock markets, momentum, is not observed in Japan. Instead, a significant short-term return reversal exists for the portfolio of the formation period of 1 month. Iihara et al. argue that investor overreaction may be a possible cause for the 1-month return reversal. Although a number of studies have examined Japanese stock markets since this paper was first written, no momentum effect has been reported except the conditional momentum effect in our addendum. Either the Japanese market is more efficient or our theoretical model is still immature or both. In Chap. 22, Professors Katsuhiko Okada, Nobuyuki Isagawa, and Kenya Fujiwara examine the Japanese stock market response to additions to the composition of the Nikkei Stock Average. This study is an extension of several U.S. studies that focus on stock price effects associated with a change in the composition of the S&P 500 index. All these studies find stock price increases for the added firms. Since the price increase is temporary, a large demand shock such as the excess demand of index arbitragers for shares of the newly added firms moves the price. This finding implies that the demand curve is downward sloping, which is inconsistent with the market efficiency assumption of a horizontal demand curve. In Chap. 23, one of the long-lived anomalies, the Sell-in-May effect, is carefully re-examined using Japanese stock return data. Although Professors Shigeki Sakakibara, Takashi Yamasaki, and Katsuhiko Okada document a similar stock return seasonality, the pattern is not exactly the same. Sakakibara et al. find stock returns are higher for the first 6 months of the year even though the Sell-in-May effect implies that stock returns are higher from November to April. For some reason, Japanese markets do not respond to this global market trend in a timely fashion. The authors call this anomaly the "Dekanshobushi effect." Interestingly, this anomaly still exists.

Ibaraki, Japan Shinsuke Ikeda
Nagoya, Japan Hideaki Kiyoshi Kato
Ibaraki, Japan Fumio Ohtake
Kobe, Japan Yoshiro Tsutsui

Acknowledgements

We thank Emi Kurimune and Azusa Ohishi of the Research Center of Behavioral Economics, Osaka University, for all the hard work, and Springer Japan for their support for this project. We appreciate the financial support from the Joint Usage/Research Center Project of the Institute of Social and Economic Research, Osaka University (project title: Overview and Future Issues of Behavioral Economics in Japan, 2014 and 2015), which was granted from the Ministry of Education, Culture, Sports, Science and Technology.

Contents

About the Editors

Shinsuke Ikeda is a professor at the Institute of Social and Economic Research (ISER), Osaka University, and serves as the director of the Research Centre of Behavioral Economics in ISER. He got a B.Com. of Kobe University in 1980 and a Ph.D. (Doctor) of Osaka University (economics) in 1997. He was the former president of the Association of Behavioral Economics and Finance. He published articles on behavioral economics, macroeconomic dynamics, and asset pricing in Journal of Finance, Journal of Health Economics, Journal of International Economics, Journal of Monetary Economics, International Economic Review, etc. His work on behavioral economics is incorporated into the book The Economics of Self-Destructive Choices, Springer, to appear in 2015.

Hideaki Kiyoshi Kato is a professor of finance at Graduate School of Economics, Nagoya University, Nagoya, Japan. Before joining Nagoya University, he taught at Kobe University. He received his Ph.D. degree from the University of Utah in 1985. He has published several books and more than 30 articles in the leading finance journals such as Review of Financial Studies, Management Science, Journal of Financial Economics, Journal of Financial and Quantitative Analysis, International Review of Finance, Japan and the World Economy, Pacific Basin Finance Journal, Journal of Portfolio Management, Journal of Financial Research, Journal of Banking and Finance, and Journal of Futures Markets on subjects including the market efficiency and anomalies, stock options, investor behavior, dividend policy, equity offerings, and stock index futures. He is currently an associate editor of Pacific Basin Finance Journal and International Review of Finance.

Fumio Ohtake is Osaka University distinguished professor and a professor in the Institute of Social and Economic Research at Osaka University and an executive vice president of Osaka University. He earned his M.A. and a Ph.D. from Osaka University in 1985 and 1996, respectively, and a B.A. from Kyoto University in 1983. He is the president of the Association of Behavioral Economics and Finance, and an executive director of the Japanese Economic Association. His research topics are behavioral economics, labor economics, income distribution, and household behavior. He is also a recipient of the 2005 Nikkei Prize for Excellent Books in Economic Science; the 2005 Suntory Prize for Social Science and Humanities; the 2005 Economist Prize; the 2006 Ishikawa Prize of the Japanese Economic Association; and the 2008 Japan Academy Prize.

Yoshiro Tsutsui is a professor of economics at Konan University. He had previously taught at Nagoya City University and Osaka University. He was awarded a Ph.D. (economics) from Osaka University. He was the first president of the Association of Behavioral Economics and Finance, and the president of Japan Society of Monetary Economics. His primary areas of teaching and research are behavioral economics and banking and finance. Currently, his research includes happiness study, time discounting, international linkage of stock prices, and regional banking and finance. His publications appeared in journals including Review of Economics and Statistics, Journal of Financial and Quantitative Analysis, Journal of Banking and Finance, Journal of Risk and Uncertainty, Regional Science and Urban Economics, and Journal of Research in Personality. In 1988, his book, The Financial Markets and Banking Industry: Economic Analysis of Industrial Organization (Toyokeizai-Shinpo Sha, in Japanese) was awarded the Nikkei Prize for Excellent Books in Economic Science.

Part I
Intergenerational Interactions

Chapter 1
An Equilibrium Model of Child Maltreatment

Hideo Akabayashi

Abstract We propose a dynamic equilibrium model of human capital development of a child that can explain why a parent-child relationship might lead to child maltreatment. Assuming that a parent cannot observe a child's human capital accumulation or effort, and that the child's time preference develops endogenously, an unstable path of the parent's beliefs regarding the child can persist in equilibrium when the parent faces a high degree of uncertainty in inferring the child's human capital. The parent with an initial high estimate of the human capital then tends to underestimate the child's effort, which results in persistently punitive—abusive—nteractions.

Keywords Human capital production • Parental intervention • Family education • Child development • Time preference

1 Introduction

The purpose of this chapter is to propose a dynamic equilibrium model of a child's human capital formation and the parents' style of interactions with the child and thereby explain complicated phenomena in modern families, such as child maltreatment (abuse).[1] This is probably the first rational choice model of child maltreatment in economic literature that is consistent with recent views on child maltreatment.

The original article first appeared in the Journal of Economic Dynamics and Control 30: 993–1025, 2006. A newly written addendum has been added to this book chapter.

[1] The terms "maltreatment" and "abuse" are used interchangeably throughout this chapter, although the former is usually intended to cover a wider range of behavioral patterns such as abuse and neglect, both emotional and physical.

H. Akabayashi (✉)
Faculty of Economics, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
e-mail: hakab@econ.keio.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_1

Parental interactions with the child change over the course of a child's development. My calculations from the National Longitudinal Survey of Youth—Child Supplement in 1988 suggest that, on an average, a mother spanks her 5-year-old boy once a week and praises him 6 times a week, whereas the frequencies reduce to 0.7 times a week and 4 times a week, respectively, when he turns 7. Generally, there is a tendency for a mother to intervene less often as a child grows older. This may be explained by an older child's improved ability to control him or herself with better foresight, thus, the frequency with which the parent needs to intervene is reduced. We can say that the average path of interactions is *stable* in this sense, and a dynamic model should help us explain this formally.

However, there is a growing concern regarding the causes of deviations from such a path in some families, especially in extreme cases of child maltreatment. Until recently, such deviations, characterized by parents' increasing interventions and a child's developmental delay, have frequently been ascribed psychopathological explanations. Nonetheless, very few studies have succeeded in differentiating between abusers and nonabusers on the basis of traditional measures of personality disturbances (Wolfe 1987, p. 45). In recent explanations, it is common to view child abuse "along the hypothetical continuum that establishes the polar opposites of abusive and healthy parenting styles" and as a "process between parent and child, within both the familial context and the larger social structure" (Wolfe 1987, p. 40 and p. 48). In this sense, most professionals no longer believe that child abuse is pre-programmed in "crazy parents," but view it as an outcome of the continuous breakdown of normal parent-child relationships. The model presented in this chapter allows us to interpret such a deviation as an unstable equilibrium path of parent-child interactions.[2]

The principal-agent framework is used to describe a family consisting of an altruistic parent and a growing child. The key assumptions are: (1) a child's human capital develops through his or her own effort under parental influence and interventions, (2) a child's rate of time preference is a decreasing function of the human capital, (3) the parent cannot directly observe the child's human capital and the parent's observation errors can be reduced by spending additional time with the child, and (4) the parent updates her beliefs regarding the child's human capital level using available information.

The dynamic equilibrium process of the parent's beliefs about the child's human capital indicates that the parental beliefs may diverge. It is then suggested that the parent with a high initial expectation about the child's ability tends to maintain an unreasonably high expectation about the child's behavior, which leads to a persistently negatively biased assessment of the child's effort. The parent's optimal interactions with the child tend to be punitive rather than positive, thereby providing an explanation of child maltreatment.

[2]Therefore, child abuse as a direct result mental illness, frustration, stress from divorce and unemployment, unattractiveness of the child, and the parents' own upbringing is beyond the scope of this chapter. Sexual abuse is not explained either.

Theoretical results and implications, including the role of parental expectations, are found to be mostly consistent with the recent views on child maltreatment in psychological literature. Although the model presented here is not intended to be a comprehensive theory of child maltreatment or parent-child interactions, the results suggest that interpretations based on rational choice and equilibrium are highly useful for better understanding those phenomena.

A static model of parent-child relationships as the principal-agent model was first proposed by Akabayashi (1996), which focused on the conflicts between parents and children stemming from the fact that children's efforts are unobservable and that children tend to be myopic. It was shown that, in equilibrium, altruistic parents choose to provide incentives (e.g., praise and punishment) in order to influence the child's development, formalizing psychologists' views of parent-child relationships.[3] This chapter considers the dynamic interrelationships between the development of a child's characteristics, the accumulation of human capital, and parental expectations in order to determine how a child's characteristics develop and how parental actions toward the child change over time.

2 Issues of Child Maltreatment

There has been an increasing interest in the prevention and treatment of child abuse. The most recent incidence rate of child maltreatment is 42 per 1,000 children (Sedlak and Broadhurst 1996). Abused children not only have been found to have emotional and behavioral problems, such as higher rates of aggression, acting out, and hyperactivity, but also problems in cognitive development and social competence.[4] To the extent that these developmental elements predict their future socio-economic status, how a child is reared must be regarded as an important input to the production of human capital.

There are numerous theories that researchers use for explaining why parents might abuse children.[5] It is probably impossible to construct a single model that explains the various aspects and causes of child abuse, because it is a "multi-

---

[3]The conflict between an altruistic parent and a selfish child has been widely discussed in the family economics literature (Bergstrom 1997). Weinberg (2001) developed a similar model that includes the effects of parental income on the choice between pecuniary and non-pecuniary methods of punishment. Using state level panel data, Paxson and Waldfogel (2002) empirically found that child maltreatment is correlated with the father's absence, poverty, and unemployment.

[4]Appelbaum (1977) found a delay in language development among abused children. Erickson et al. (1989) states, "disproportionate numbers of abused children have been found to perform below the average range on IQ tests." Wolfe (1987) states, "from preschool age and beyond, studies have found that abused children are significantly more likely than their peers to show delays related to cognitive development and deficits in academic performance and intellectual functioning."

[5]"Psychological Theories of Child Maltreatment" gives a brief explanation of competing theories of child abuse.

dimensional" (Wolfe 1987, p. 59) phenomenon and the points of emphasis depend on the approach selected. Recent approaches, labeled as the *social-cultural* approach or the *social-interactional* approach, view child abuse as a consequence of a "process" of interactions between parents, children, and their socio-economic situation, rather than as a psychopathological consequence of the parents' predetermined characteristics. In this chapter, while placing child abuse in the broad context of the developmental consequences of the parent-child relationship, the focus is on the role of a parent's expectation of the child's development in causing an abusive relationship.

The importance of the parent's expectation of a child in an abusive relationship has been extensively documented in literature. For instance, Zigler and Hall (1989, p. 64) wrote, "Parents who have unrealistically high expectations of their child are more likely to abuse than are parents who have a good understanding of the sequence of child development." Wolfe (1987, p. 87) states, "Practitioners observed that many abusive incidents involved senseless attempts by the parent to force a child to behave in a certain manner that was beyond the child's developmental limitations."

Focusing on this aspect, here child abuse is defined as "a dynamic parent-child relationship where the parent unreasonably overestimates the child's ability, tends to form a negatively biased view of the child's behavior, and maintains or excessively increases negatively-biased ("punitive") interactions."[6] We define the parent's "interactions" as encompassing all physical, verbal, and psychological interactions with the child.[7] Although, initially, a parent's estimate of the child's characteristics may be inaccurate, this estimate generally converges to the true value as the parent collects information about the child. However, even after many observations of the child's behavior, some parents continue to possess an unreasonable belief about the child's characteristics and tend to build a negatively biased view of the child's self-control or effort. To highlight this idea, we consider the child's rate of time preference as a key determinant of the child's developmental characteristics, and we assume that this rate of time preference is related to the child's maturity or "basic human capital."[8] The study investigates the reasons for

[6]As will be clear in the following sections, "negatively biased" interactions mean only that utility transfer from a parent to a child tends to be smaller than the expected amount based on the predetermined incentive schedule and not that parents transfer a negative amount of utility.

[7]Therefore, this definition does not necessarily imply "physical" abuse. Some researchers are beginning to define child abuse in such a sense, which is broader than that commonly used. Wolfe, for example, states that, "Child abuse, according to this [social-psychological] perspective, can be viewed as an extreme disturbance of child rearing, which is to say it is not necessarily an individual disorder or psychological disturbance. Abusive families are ones in which the usual balance between positive and negative interactions and between discipline and emotional bonding has not been achieved" (1987, p. 18).

[8]Here "basic" human capital broadly represents the degree of maturity that encompasses personality and cognitive abilities, rather than skills or knowledge that are directly productive and observable in the market. In the subsequent sections, a mature child means a child with a high level of basic human capital. Henceforth, "basic" is dropped for simplicity.

the biased parental belief regarding the child not converging to the true value and remaining negatively-biased along with the manner in which this occurs.

3 The Model

3.1 Law of Motion of a Child's Human Capital

Let us suppose a family consists of one parent and one child and denote the child's human capital at the beginning of period t by h_t, where $t = 1,\ldots,T+1$. $T+1$ is the period when a child becomes independent of the parent and starts relying only on the value of his or her own human capital accumulated over the previous periods. A child's initial human capital or "potential ability," h_1, is assumed to be given and positive. A child's human capital at subsequent periods is assumed to be determined by the child's human capital level in the immediate past, the level of effort, the parent's time spent with the child, and the family environment including the parent's human capital level. We assume that the parent and the child know that *the law of motion of human capital* is described by the following linear process[9]:

$$h_{\tau+1} = (1-\delta)h_\tau + \varphi s_\tau H^\gamma + \phi a_\tau, \quad \text{for} \quad \tau = 1,\ldots,T, \tag{1.1}$$

where H is the parent's human capital (assumed to be positive and constant over time), $s_\tau \in (0,1]$ is her (normalized) time spent with a child, and a_τ is the child's effort level. We assume that δ is strictly positive and less than 1 so that the first term represents the depreciation of human capital. The second term represents the parent's investment in the child's human capital (or "education"), which is a function of the time spent by the parent and her human capital level. The third term represents the child's own investment in human capital (or "learning by own effort"). φ and ϕ are presumably positive marginal effects of "education" and "effort" on human capital, respectively. We assume that ϕ is less than $1-\delta$ so that the effect of effort is smaller than the effect of past human capital. By applying (1.1) repeatedly, we have $h_{T+1} = (1-\delta)^{T-t+1}h_t + \sum_{\tau=t}^{T}(1-\delta)^{T-\tau}(\varphi s_\tau H^\gamma + \phi a_\tau)$, for arbitrary $t < T+1$.

[9]Additivity and linearity in choice variables in the production process are imposed to avoid unnecessary complication. These restrictions ignore a productive complementarity between the child's current human capital and the parent's time spent with the child or the child's effort. The assumption that both the parent and the child know the process may also seem unrealistic; however, this is the standard assumption in the classical Kalman filter literature, and increasing the number of unobservables would unnecessarily complicate the analysis. The point is that the current setup is sufficient for explaining essential features that characterize child maltreatment *without* further complications.

Therefore, the human capital level when a child becomes independent of the parent (period $T+1$) can be expressed as a function of the human capital at an arbitrary period t and inputs of the child's effort as well as the parent's time spent at and after t.

3.2 Observation Equation of the Child's Behavior

We assume that, while the child knows his or her own h_t, the parent cannot directly observe her child's human capital or effort but can observe the child's performance at period τ, y_τ. The child's performance is determined by his human capital, effort, and a random shock in that period, according to the following linear *observation equation*:

$$y_\tau = h_\tau + a_\tau + \nu_\tau, \quad \text{for} \quad \tau = 0, 1, \ldots, T, \tag{1.2}$$

where ν_τ is a random variable distributed as $N(0\,,\,\sigma^2_{\nu\tau})$ for all τ. The random shock, ν_τ, includes shocks to the child's performance as well as the parent's measurement error. The first two terms indicate that a more mature child with more effort tends to behave better. The nature of the third stochastic term depends on the duration of the parent's observation. We assume that $\sigma^2_{\nu t}$ is decreasing in s_τ, the parent's time spent with the child, because spending more time with a child would presumably reduce the parent's measurement error in the observations. The error term cannot be eliminated even if the parent spends the maximum possible time with the child, because the child may still make unintended mistakes. We assume that $\sigma^2_{\nu\tau} \equiv K/s_\tau$, where K represents the parent's monitoring ability, possibly correlated with her human capital level.[10]

The information structure is specified as follows. The parent's *information set* at t is defined as the set of all information available at period t, denoted by $I_t \equiv \{y_t, y_{t-1}, \ldots, y_1\}$. We denote the parent's subjective expectation and mean squared forecasting error of h_t based on I_s by $\hat{h}_{t|s}$ and $\sigma^2_{ht|s}$, respectively. To simplify our notation, we define $\hat{h}_t \equiv \hat{h}_{t|t-1}$ and $\sigma^2_{ht} \equiv \sigma^2_{ht|t-1}$ which represent the best one-step-ahead predictor and mean squared error, respectively. Let the pair $(\hat{h}_t, \sigma^2_{ht})$ denote the parent's *belief at period t*. We assume that the parent has a prior belief $(\hat{h}_1, \sigma^2_{h1})$ at the time of the child's birth.

[10]The introduction of this endogenous observation error makes the expected utility function quasi-linear in s_t, allowing the achievement of a strictly positive solution for s_t.

3.3 Parent's Incentive Schedule

We assume that a parent considers a child's happiness as her own happiness (altruism) and can create and transfer "services" to the child. The child derives utility from these services. Let d_τ denote the amount of services created and transferred at period τ. These services are assumed to consist of two components: the time spent with the parent (s_τ), and parent's interactions (kiss, hug, spank, etc.). The first component is directly productive since it appears in (1.1), while the second component is assumed to have only psychological effects. At each period, the parent sets the time spent with the child and promises a schedule based on which she interacts with the child. Since the parent cares about the child's future human capital, her actual choice of interaction depends on the promised schedule and her estimate of the child's effort given the available observations, $\mathrm{E}[a_\tau|I_\tau]$. After interacting with the child, the parent revises her belief regarding the child's human capital.[11] More specifically, we consider only the following linear *incentive schedule,* by which the parent produces the argument of the child's utility function measured in hours multiplied by a measure of the parent's human capital,[12]

$$d_\tau = (s_\tau + b_\tau E[a_\tau|I_\tau])H^\gamma, \quad \text{for} \quad \tau = 0, 1, \ldots, T. \tag{1.3}$$

Among the components of d_τ, $s_\tau H$ is the service created by spending time with the child, $b_\tau \mathrm{E}[a_\tau|I_\tau]H$ is the service from the parent's interactions contingent on the new observation, y_τ, and b_τ is called the *slope of incentive*.[13] This represents the parent's marginal change of interactions with the child measured in the equivalent unit of time when her estimate of the child's effort changes. Note that we assume that the child's "effective" incentive that is created is a multiple of H, the parent's human capital level. Therefore, it is reasonable to term $b_\tau \mathrm{E}[a_\tau|I_\tau]$ the *parent's observed interaction*. Since b_τ is shown to be positive in equilibrium, the observed

[11]Therefore, the promise is binding even though the new observation may let the parent revise her belief. Otherwise, the child cannot weigh the expected reward against the painful effort since he cannot foresee the evolution of the parent's expectation.

[12]The linear incentive schedule, which is assumed as a technological constraint to parents, greatly simplifies our analysis while maintaining the important implications. This class of models was first analyzed by Holmström and Milgrom (1987) and was used by Gibbons and Murphy (1992). The latter analyzed the effect of CEOs' career concerns on the equilibrium incentive payments, assuming a linear incentive schedule and exponential preferences. However, our model differs fundamentally in that there is period-by-period development of capital with inputs from the principal, which is absent in the other two papers. An obvious problem in this class of models is the absence of an income effect, which means that we cannot study the direct effect of the parent's financial conditions on the child's development.

[13]One referee suggested formulating an incentive scheme in which time with the child, s_τ, enters multiplicatively with the parent's interactions, b_τ. Such a construction may be more realistic in some sense, but the cost of it would be that the model becomes nonlinear from the beginning, making it highly intractable.

interaction is large ("praise") when the parent observes good performance and forms a high estimate of the child's effort, and it reduces ("punishment") when poor performance is observed. Given this structure, a set of the two variables from period t onward, $\{s_\tau, b_\tau\}_{\tau=t}^{T}$, completely defines the parent's *plan of parenting at period t*. The assumptions that the parent can choose some part of the child's utility and that the parent cares about the child create a connection between the parent and the child, which is the foundation upon which this model is built.

Using (1.2), we have $\mathrm{E}[a_\tau|I_\tau] = \mathrm{E}[y_\tau - h_\tau - \nu_\tau|I_\tau] = y_\tau - \hat{h}_\tau = h_\tau + a_\tau + \nu_\tau - \hat{h}_\tau$; (1.3) is then rewritten as:

$$d_\tau = (s_\tau + b_\tau(h_\tau + a_\tau + \nu_\tau - \hat{h}_\tau))H^\gamma, \quad \text{for} \quad \tau = 0, 1, \ldots, T. \tag{1.4}$$

Therefore, given a series of the current and past observations (I_τ), the parent's incentive provision is based on the difference between the behavior observed today and the best estimate of the child's human capital. Clearly, given an observation y_τ at period τ, if the parent had a high expectation of the child's human capital level (high $\hat{h}_\tau$) at the beginning of that period, she tends to have a low estimate of the child's effort (low $\mathrm{E}[a_\tau|I_\tau]$) and tends to "punish" him or her (low d_τ), and vice versa.

3.4 Preferences

First, we assume that the rate of time preference is a decreasing function of human capital (Becker and Mulligan 1997). Let us denote the child's rate of time preference and the parent's rate of time preference by $\rho_{ct}(\equiv \rho(h_t))$ and $\rho_p(\equiv \rho(H))$, respectively. We also assume that $\lim_{h\to\infty} \frac{d}{dh}(1/(1+\rho_{ct})) = 0$ and that there is a value of h, h^c, such that $1/(1+\rho_{ct})$ is concave in $h_t \in (h^c, \infty)$. These are natural assumptions since the discount factor is bounded from above. The assumption on the limit may be restated as: "the discount factor tends to be inelastic with respect to human capital as the level of human capital increases," like many characteristics that tend to be fixed as a child becomes an adult. For instance, $(1/(1+\rho_{ct})) = 1/(1+\exp(-\eta h_t))$ with $\eta > 0$ satisfies these requirements for $h_t > 0$, which will be used later.

A child is assumed to be myopic in three ways. First, the child's rate of time preference is generally greater than the parent's rate, because the child is less mature (as measured by the child's level of human capital). Second, although the child knows about the law of motion of his or her own human capital, that future tastes might change with the evolution of human capital is not known to the child, and therefore, the child considers the current rate of time preference as given in deciding the future effort allocation plan.[14] Finally, due to the lack of knowledge of changing preferences, the child does not know how his or her choice today may influence the parent's future parenting choices through her improved knowledge of the child's preferences. Clearly, the child's decision might be time-inconsistent, and the child might regret and revise the plan. In contrast, the parent is less myopic in the sense of having a lower rate of time preference (ρ_p) and has the knowledge that the child's rate of time preference changes as the child grows.

Further, we assume that effort is painful to the child and provide disutility $-v(a_\tau)$, where $v(\cdot)$ is a positive, increasing and convex function. In particular, we assume $v(a) = (a - \underline{a})^2/2\psi$, where $\underline{a}$ is an individual fixed characteristic representing the child's least painful level of effort. We term this level as the *child's natural level of effort*. $1/\psi$ determines the child's marginal disutility of effort. The child's one-period utility is determined by the sum of this disutility of effort and the incentive schedule provided by the parent, namely $d_\tau - v(a_\tau)$.

Finally, we assume that both parent and child have exponential preferences toward risk in their life-cycle utility: the parent maximizes the expected value of $U(\cdot) \equiv -[\exp\{-R(\cdot)\}]$ while the child maximizes the expected value of $u(\cdot) \equiv -[\exp\{-r(\cdot)\}]$, where $(\cdot)$ takes each agent's sum of utility over the life cycle as its argument and R and r are the parameters governing attitudes toward risk.

4 Optimal Interactions and Equilibrium

4.1 Child's Decision Problem

A series of decisions in one period takes place as follows. Given a belief about a child's human capital $(\hat{h}_t, \sigma_{ht}^2)$ at the beginning of period t, the parent decides upon a plan of parenting $\{s_\tau, b_\tau\}_{\tau=t}^T$. Given this, the child chooses a plan of efforts $\{a_\tau\}_{\tau=t}^T$. Next, the child's performance is observed according to (1.2). The parent determines the amount of the services to be provided to the child via (1.4) and revises her belief. Finally, the child's human capital develops according to (1.1).

Consider the child's problem at period t. The child's optimization problem is

[14]This implies that a child has knowledge of his or her production function (or at least, the marginal productivity of his or her effort), but no knowledge of his preference formation function. This may sound inconsistent, but if we were to assume that a child had no knowledge of how his or her human capital evolved, then the child would have no motivation to suffer the disutility associated with investing in his or her own human capital. We would then obtain an unrealistic result that the endogeneity of his or her rate of time preference would have no effect on the child's choice. The most realistic way to proceed would be to introduce a child's "learning" of his or her own human capital production function, a complication that we will not pursue here.

$$\text{Max} \quad \mathrm{E}\left[u\left(\sum_{\tau=t}^{T}\left(\frac{1}{1+\rho_{ct}}\right)^{\tau-t}(d_\tau - v(a_\tau)) + \left(\frac{1}{1+\rho_{ct}}\right)^{T-t+1} B\cdot h_{T+1}\right)\middle| I_{t-1}\right]$$

$$\text{subject to } (1.1), (1.4) \text{ for } \tau = t, \ldots, T, \text{ given } \{s_\tau, b_\tau\}_{\tau=t}^{T} \text{ and } \mathrm{h_t},$$

where $B\cdot h_{T+1}$ (B is a positive constant) determines the child's utility from his or her own human capital at $T+1$. Then, the first order condition for a_t[15] is

$$b_t H^\gamma - (a_t - \underline{a})/\psi + \left(\frac{1}{1+\rho_{ct}}\right)^{T-t+1}(1-\delta)^{T-t}B\phi = 0. \tag{1.5}$$

While deciding the level of effort using (1.5), the child compares the immediate marginal pain (the second term) with the sum of the immediate marginal return from the parent's love (the first term) and the subjectively discounted future return from his or her human capital stock upon becoming an adult at $T+1$ (the third term). Notice that the child's decision regarding today's effort is independent of his or her future decisions or future human capital levels, due to additivity and the child's myopia over changing preferences. His or her planned future efforts as of today might differ from efforts actually chosen in the future, because the rate of time preference changes and the way in which it will change is unknown today. Furthermore, the child might make a wrong guess about the parent's future actions. This inconsistency does not pose a problem in interpreting the child's decision *today* because it depends only on his or her human capital and parental incentives *today*.[16] By defining the child's *subjective marginal return to investment* at age t, $D_t(h_t) \equiv \left(\frac{1}{1+\rho_{ct}}\right)^{T-t+1}(1-\delta)^{T-t}B$, (1.5) is solved for the optimal effort at t in response to the parental incentive[17]:

$$a^* = \underline{a} + \psi b H^\gamma + \phi\psi D_t(h). \tag{1.6}$$

The child's natural level of effort ($\underline{a}$) and the parental incentive (the second term) have positive effects on his effort. Since $D_t(h)$ is increasing in h and in t independently, an older or more mature child tends to make more effort. The reason

[15] See Appendix "The First Order Condition of the Child's Decision Problem" for the derivation of the first order condition of the child's decision problem.

[16] Thus, the usual "ratchet effect" is absent from our setting. See Appendix "The First Order Condition of the Child's Decision Problem".

[17] Hereafter, the time subscript is suppressed whenever there is no ambiguity to economize on the notation, except in mathematical appendices and in the case of $D_t(h)$, $D_t(H)$, and Q_t (defined later), which explicitly depend on time.

that age has an independent effect is that the need for effort becomes more apparent as the child becomes older (finite-horizon effects). A larger level of effort would be chosen if the child's marginal disutility of effort $(1/\psi)$ is smaller, or if effort is more productive in the accumulation of human capital (larger ϕ). Notice that a^*, which is known to the child, is an unobservable stochastic variable to the parent even though the parent controls the slope of incentive (b), because the uncertainty regarding h still remains. Thus, the child's choice of effort partly depends upon his "maturity," regarding which the parent can update her belief from past observations.

From (1.6), the observation equation (1.2) can be rewritten as $y^* = h + \phi\psi D_t(h) + \underline{a} + \psi b H^\gamma + \nu$. We can see that, given the parent's choice of the "effective incentive slope," (bH), the child's observed performance is positively correlated with his human capital level for at least two reasons. The first term represents the exogenous effect of h on the child's performance. The second term represents the endogenous effect of human capital on the child's performance, because the choice of effort depends on the child's rate of time preference which, in turn, depends on the child's human capital.

4.2 *Parent's Decision Problem*

At period t, the parent chooses a plan of parenting, $\{s_\tau, b_\tau\}_{\tau=t}^{T}$, to maximize her expected utility from family consumption and the child's happiness, given the child's response function.[18] Let π be the wage rate of one efficiency unit of the parent's human capital, and let α describe the parent's degree of altruism toward the child, both of which are assumed to be time-invariant. Assuming the parent has one unit of time to spend either with the child or working, she will spend $1 - s_t$ units of time working in the market. The parent's problem is to

$$\text{Max} \quad \mathrm{E}\left[U\left(\sum_{\tau=t}^{T} \left(\frac{1}{1+\rho_p}\right)^{\tau-t} [c_\tau + \alpha(d_\tau - v(a_\tau^*))] \right.\right.$$
$$\left.\left. +\alpha \left(\frac{1}{1+\rho_p}\right)^{T-t+1} B \cdot h_{T+1} \right) \right| I_{t-1} \Bigg]$$

$$\text{subject to } c_\tau = \pi(1-s_\tau)H^\gamma, (1.1), (1.4), \text{ and } (1.6), \text{ for } \tau = t, \ldots, T.$$

Notice that, for the same reason as stated before, the parent's decision today is independent of the parent's (or the child's) future decisions. Therefore, we can

[18] An underlying assumption is that the parent evaluates the child's stochastic utility using her own risk preference. In our formulation, there is no explicit incentive compatibility constraint that conditions the exit of the child, since the parent's altruism and risk-aversion impose a natural limit on her behavior.

focus on the choice of the "current" parenting plan, $\{s_t, b_t\}$, and the child's current response, a_t^*.[19] Choosing a large s_t is costly because it can be achieved only if the parent spends less time at work. Although there is no explicit market cost in choosing a large b_t, a risk averse, altruistic parent has a reason to avoid this, because she would prefer less variability in her interactions with the child. This is clearly seen from the following first order condition for b_t (with time subscript being suppressed again).[20]

$$\begin{aligned} bH^{\gamma} &= \frac{1}{R\alpha(\sigma_h^2 + K/s)}(\underline{a} + b\psi H^{\gamma} + \psi\phi D_t(H)) \\ &= \frac{1}{R\alpha(\sigma_h^2 + K/s)}(\hat{a}^* + \psi\phi[D_t(H) - D_t(\hat{h})]), \end{aligned} \tag{1.7a}$$

where $D_t(H) \equiv \left(\frac{1}{1+\rho_p}\right)^{T-t+1}(1-\delta)^{T-t}B$ defines the parent's *subjective marginal return to investment* and $\hat{a}^*$ is the parent's estimated effort of the child based on the estimate of human capital. Thus, the parent needs to set a steeper incentive slope if she wants to induce greater effort ($\hat{a}^*$) or if she estimates a larger difference in her own and the child's subjective marginal return to investment, other things being equal. As long as H is larger than $\hat{h}$ and $\underline{a}$ is positive, the optimal slope must be positive. The first order condition for s_t is

$$\pi = \alpha + \alpha\varphi D_t(H) + \frac{R\alpha^2 H^{\gamma} K b^2}{2s^2}. \tag{1.7b}$$

Here, the left hand side is the opportunity cost of being with the child. The right hand side is the marginal return to time spent with the child and consists of the following three components: The first term is the immediate marginal happiness derived from being with the child, the second term is the future marginal return to increasing the child's human capital, and the last term is the return to improved information about the child's behavior. The last term appears because being with the child makes it easier for the parent to monitor the child, which reduces the risk of punishing a good child. To clarify this point further, we rearrange (1.7b) to obtain a proportional relationship between s and b as follows[21]:

[19] As we show in Appendix "The First Order Condition of the Child's Decision Problem", regardless of the parenting plan chosen at time t, the future plan does not influence the child's current effort. Thus the child does not act to influence the parent's future belief or actions. This property comes from our linear technology and greatly simplifies our analysis. We do not argue that this formulation is the only possible way parents may plan actions, given the time-inconsistent preference structure. However, we find that this formulation is highly convenient and useful for our purpose.

[20] See Appendix "The First and Second Order Conditions of the Parent's Decision Problem" for the derivation of the following first and second order conditions.

[21] We assume that $\pi - \alpha - \alpha\varphi D_t(H) > 0$ for s_t to be always positive.

$$s = \left(\frac{2H^{\gamma}}{R\alpha^2 K} \left[\pi - \alpha - \alpha\varphi D_t(H) \right] \right)^{-1/2} bH^{\gamma} \equiv Q_t^{-1} bH^{\gamma}. \tag{1.7b'}$$

We can now see that there is an endogenous complementarity between the time spent with the child and the parent's interactions. This is because, if the parent spends more time with the child, she can observe the child's behavior with less error and can set stricter criteria for judging the child's performance. It is also easy to see the following implications for the parent's substitution between "time spent" and "interactions." First, if the wage rate (π) is higher, she tends to shift away from time spent with the child and toward a stricter discipline. Second, the reverse shift would occur if the parent is more altruistic (larger α). Finally, if K or R is larger, she also tends to spend more time with the child for reducing the risk of making a mistake in judging the child's effort.

4.3 Equilibrium System Equation and the Parent's Expectation Process

The parent's and the child's optimal actions are determined jointly by (1.7a) and (1.7b'). Assuming that the second order condition is always satisfied ($R\alpha\sigma_h^2 - \psi > 0$)[22] and that an interior solution can be achieved, we have the following reduced form solutions for the slope of incentive, bH^{γ}, and parental time with the child, s:

$$b^* H^{\gamma} = \frac{1}{R\alpha\sigma_h^2 - \psi} (\underline{a} + \psi\phi D_t(H) - KR\alpha Q_t), \tag{1.8a}$$

$$s^* = \frac{1}{Q_t (R\alpha\sigma_h^2 - \psi)} (\underline{a} + \psi\phi D_t(H) - KR\alpha Q_t). \tag{1.8b}$$

Applying (1.8a)–(1.6), we obtain the reduced form solution for the child's effort, a

$$a^* = \underline{a} + \phi\psi D_t(h) + \frac{\psi}{R\alpha\sigma_h^2 - \psi} (\underline{a} + \psi\phi D_t(H) - KR\alpha Q_t). \tag{1.8c}$$

Thus, better information about the child's human capital (smaller σ_h^2) has a positive effect on each of the following: the slope of incentive, the time spent with the child, and the child's effort. Care must be taken in interpreting the result that the parent's choice variables are independent of the estimated *level* of the child's human capital ($\hat{h}$). This is, of course, because of our additivity assumption on the human capital

[22]See Appendix "The First and Second Order Conditions of the Parent's Decision Problem" for the derivation of the second order condition.

production function, and if we allow any complementarity between the inputs of the human capital production, $\hat{h}$ should affect the parent's choice. Notice also that b may not be monotonically related to the parent's human capital, H; although a more educated parent needs less severe interactions on the left-hand side of (1.8a), such a parent will tend to have a higher $D_t(H)$, which would make her choose a larger b.

A more risk-averse (larger R) parent, being afraid of punishing a good child, tends to choose a smaller b and the induced effort tends to be small. Although the parent is tempted to shift from "intervention" to "being with the child," she finally chooses a smaller s, because the return to reducing observation error reduces when a less strict intervention plan is chosen.[23] In this way, the endogenous complementarity of the parent's interactions (b) and the time spent with the child (s) tend to generate another force in the choice of the plan of parenting. Due to this complementarity, some of the comparative statics results are now ambiguous. For example, a more altruistic (larger α) parent would be willing to stay with the child longer (larger s), thereby reducing the observation error; hence the improved accuracy in the observations would allow the parent to opt for a stricter discipline (larger b). At the same time, the parent tends to shift the parenting plan away from the use of intervention. It turns out that the effects of the degree of altruism are ambiguous, depending on the relative strengths of these two opposite forces. Among the child's other characteristics, a larger ψ (less marginal disutility of effort) and a larger $\underline{a}$ (natural level of effort) have positive effects on b, s, and a.

Substituting (1.8) into (1.1) and (1.2) yields a state-space representation of the dynamic equilibrium system. The *equilibrium law of motion* (system equation) is defined as

$$h' = (1-\delta)h + \phi a^* + \varphi s^* H^\gamma = F(h) + G(\sigma_h^2), \tag{1.9a}$$

where $h' \equiv h_{t+1,} F(h) \equiv (1-\delta)h + \phi^2\psi D_t(h)$, and $G(\sigma_h^2) \equiv \phi[\underline{a} + \frac{\psi}{R\alpha\sigma_h^2 - \psi}(\underline{a} + \psi\phi D_t(H) - KR\alpha Q_t)] + \varphi s^* H^\gamma$. The equilibrium law of motion of human capital has two major components: $F(\cdot)$ is related to the child's current human capital, including its endogenous effect on effort. $G(\cdot)$ is related to the parent's background (H), the level of uncertainty regarding the child's human capital (σ_h^2), and preference parameters. Both components are time-dependent functions due to the finite horizon nature of the model.[24] Similarly, we can derive the *equilibrium observation equation* as

$$y = h + a^* + \nu = A(h) + C(\sigma_h^2) + \nu, \tag{1.9b}$$

[23] See Appendix "Mathematical Details of the Parenting Plan" for mathematical details of the discussion in this paragraph.

[24] Again, the explicit time subscript in F and G as well as A, C, and Φ (to be defined shortly) is omitted to save on the notations.

where $A(h) \equiv h + \phi\psi D_t(h)$ and $C(\sigma_h^2) \equiv \underline{a} + \frac{\psi}{R\alpha\sigma_h^2 - \psi}(\underline{a} + \psi\phi D_t(H) - KR\alpha Q_t)$

The equilibrium observation equation also consists of two major components: $A(\cdot)$ is the contribution of the current human capital, including its endogenous effect on the child's effort (the second term in the definition of $A(\cdot)$). $C(\cdot)$ is a function of the other factors that affect the child's effort. Both functions are time-dependent for the same reason as stated before.

Considering this nonlinear equilibrium system (1.9), we assume that the parent forms and updates the expectation about the child's human capital with linear approximation as follows. Given a belief $(\hat{h}, \sigma_h^2)$ at the beginning of period t and a new observation on behavior, y, the parent first updates the belief of the child's human capital today from $\hat{h}(\equiv \hat{h}_{t|t-1})$ to $\hat{h}^u(\equiv \hat{h}_{t|t})$ using the Bayesian updating rule. She then uses the optimal recursive projection formula (Kalman filter) to construct the one-step-ahead projection of the child's human capital and its error variance $(\hat{h}', (\sigma_h^2)')$ that becomes her belief at the beginning of the next period. We call the stochastic process of the parent's belief, $\{\hat{h}_t, \sigma_{ht}^2\}$, constructed in this way, the *parent's expectation process*.

To construct the expectation process, we assume that the parent uses the following algorithm.[25] First, she uses $\hat{h}$ as the first guess to linearly approximate $A(\cdot)$ and updates it using a new observation y. The updating rule is essentially an average of the previous belief and the information obtained from the new observation weighted by the degree of uncertainty, written as follows:

$$\hat{h}^u = \hat{h} + \frac{A'(\hat{h})\sigma_{ht}^2}{A'(\hat{h})^2\sigma_h^2 + \sigma_v^2}(y - A(\hat{h}) - C(\sigma_h^2)), \tag{1.10}$$

where $A'(\hat{h}) = 1 + \phi\psi D_t'(\hat{h})$ is the first-order Taylor coefficient from (1.9b). She then uses the updated value as the second approximation, obtains a better approximation of $A(\cdot)$, and updates it by applying (1.10) again. After iterating on this procedure, the parent reaches an estimate $\hat{h}^u$ and uses it with the linear approximation of (1.9a) to estimate the human capital at the beginning of the next period. We thus have the parent's expectation process as

$$\hat{h}' = F(\hat{h}^u) + G(\sigma_h^2), \tag{1.11a}$$

$$(\sigma_h^2)' = \Phi \cdot \sigma_h^2, \tag{1.11b}$$

where $\Phi \equiv \frac{F'(\hat{h}^u)^2\sigma_v^2}{A'(\hat{h}^u)^2\sigma_h^2 + \sigma_v^2}$ and $F'(\hat{h}^u) \equiv (1-\delta) + \phi^2\psi D_t'(\hat{h}^u)$ is the first-order Taylor coefficient from (1.9a).

---

[25]See Appendix "The Derivation of the Parent's Expectation Process (The Kalman Filter)" for the details of the algorithm used to derive the parent's expectation process. This approximation is called the extended Kalman filter. A linearized filter of a nonlinear system is not necessarily optimal; however, an iteration algorithm introduced here improves its accuracy. See Hamilton (1994) and Anderson and Moore (1979) for more details on the algorithm considered here.

The time-dependent coefficient, Φ, characterizes the stability of the parent's expectation process. Since this is a finite-horizon problem and the rate of time preference is endogenous, the process is state-dependent. There is no stationary state and we cannot expect the error variance to converge mechanically as predicted by the theory of the time-invariant Kalman filter. However, if Φ is less than 1, the parent's belief converges in the sense stated in the following lemma[26]:

Lemma 1 *The error variance σ_h^2 of the parent's estimate of the child's human capital monotonically decreases over time if $\Phi < 1$ is satisfied.*

Proof Obvious from (1.11). □

From the definition of Φ, it is clear that a sufficient condition for Φ to be less than 1 is that $F'(\hat{h}^u)$ is less than 1.[27] Since $D_t(\hat{h}^u)$ is concave in $\hat{h}^u$ due to the assumption regarding the discount factor function, the condition in the lemma is satisfied if the parent's expectation is kept adequately high to make $D_t'(\hat{h}^u)$ sufficiently small. Let us define h_1^* and h_2^* to be the two levels of human capital that solve $F'(h) = 1$. If $\hat{h}^u$ is strictly higher than h_2^*, the child is perceived to be sufficiently mature and therefore the child's rate of time preference is insensitive to the change in his or her level of human capital. Then, the lemma states that, starting from any initial σ_h^2, the parent's uncertainty decreases over time. Clearly, such an h_2^* must be larger if ψ or ϕ is larger or if δ is smaller. However, if $F'(\hat{h}^u)$ is not less than 1, Φ can be greater than 1. In fact, we can show that the following proposition holds:

Proposition 1 *Suppose that $\rho_{ct} = \exp(-\eta h_t)$ with $\eta > 0$ and that there exists a level of h for which $F'(h) > 1$. Then, (i) there exists a combination of parameter values such that there exists a level of $\bar{\sigma}^2$ and the associated range of $\hat{h}^u$, $R(\bar{\sigma}^2) = (h_L, h_H)$, such that for any $\sigma_h^2 > \bar{\sigma}^2$ (equivalently, σ_v^2 is larger than $\sigma_h^2(1-(1-\delta-\phi)^2)/\phi^2$) and any $\hat{h}^u \in R(\bar{\sigma}^2)$, Φ is greater than one, (ii) given an updated belief $(\hat{h}^u, \sigma_h^2)$ and the parameter values that satisfy the condition described in (i), the parent's expectation process becomes divergent, and (iii) h_L is decreasing and h_H is increasing in $\bar{\sigma}^2$.*

Proof See Appendix "Proof of Proposition 1".

Proposition 1 indicates that the necessary condition for the parent's belief to be divergent is that the parent-child pair satisfies $\sigma_h^2 > \min(\bar{\sigma}^2) \equiv \bar{\sigma}_m^2 > 0$. Although this condition may not look intuitive, it is equivalent to $\sigma_v^2 > \sigma_h^2(1-(1-\delta-\phi)^2)/\phi^2)$ (see Appendix "Proof of Proposition 1"), which implies that the uncertainty due to the observation error is larger (by a certain factor) than the uncertainty regarding the child's current human capital. Recalling that the actual observation is the sum

[26] Hereafter we use "convergence" in the sense stated in the lemma. "Divergence" is defined and used similarly.

[27] Thus, in the absence of endogenous discounting, the parental expectation process will always converge.

of these components,[28] it is no surprise that, in such a situation, an additional observation will not improve the knowledge about the child.

5 Interpretation

5.1 Interpreting the Stable and Unstable Expectation Processes

We have seen that the nonlinear equilibrium system equations (1.11) are state-dependent, due to the endogenous development of the child's rate of time preference and the finite horizon. The reason for this property can be explained as follows. A child's human capital develops endogenously as the child matures, because it is enhanced by the child's effort, which depends on the child's rate of time preference (see (1.9a) and the definition of $F(\cdot)$). Therefore, despite the regression-to-the-mean nature of the *given* law of motion of human capital (1.1), the *equilibrium* law of motion of human capital may not exhibit regression to the mean. The same reasoning applies to the parent's expectation process. When the parent updates her belief about the child's maturity upwards, she also revises the expectation of the child's effort because she thinks, "*the kid seems to get smarter, so he must be more responsible*." The expectation may self-generate a higher expectation of the child's development with increased uncertainty due to the endogenous system coefficient, Φ, in (1.11). Φ can still be less than 1 if the revision of the expectation of the effort is sufficiently small. This is likely, as implied by the lemma and Proposition 1, (i) if the child is sufficiently grown up and the child's behavior is not greatly affected by a small change in his or her level of human capital, or (ii) if the parent has adequately good initial knowledge about the child and therefore spends a long time with the child in order to maintain low observation uncertainty (and low uncertainty from the unobservable effort).

When the conditions described in Proposition 1 are satisfied, the equilibrium development process of human capital becomes endogenously explosive, at least locally, and so does the associated parental expectation process. While providing the full statistical characteristics of this locally explosive process is beyond the scope of this chapter,[29] we will discuss two typical cases in which different initial

[28]The observation equation (1.2) also includes the child's "effort" term, but it adds the same information as that of human capital.

[29]Literature on the Bayesian learning process has recently investigated the nature of optimal control with learning of unknown (but fixed) parameters as "optimal experimentation" (see Wieland 2000). Literature on the Kalman filter has studied the explosive nature of the filter when specification errors exist in the simulated system (Fitzgerald 1971). Other authors (Basawa and Scott 1983; Domowitz and Muus 1988) have studied the likelihood estimation of parameters of non-ergodic processes, but not in the context of the state-space system. To the author's knowledge, neither literature has investigated the nature of the Kalman filter of explosive (or more generally non-ergodic) processes.

conditions may generate dramatically different equilibrium paths. The two cases are then visualized with a numerical simulation in order to enhance the understanding of qualitative discussions on the stability of the equilibrium path and the emergence of child maltreatment—a persistent negatively biased belief and interaction toward the child.

(a) *Case of converging belief—normal family*

Figure 1.1 illustrates the phase diagram of the parent's expectation process based on Proposition 1. The figure treats only the domain where the child is in the middle of development and the parent estimates a low level of $\hat{h}^u$. Therefore, in the figure, $\hat{h}^u$ is increasing over time except for perturbations by random shocks. We will focus on what occurs around the curve $\Phi = 1$, since it provides us with the most important and interesting interpretations in characterizing the dynamics.

Suppose that a parent is very sure that the child has a high level of human capital with a small error variance, as shown by $\hat{\mathbf{A}}$ in Fig. 1.1. The parent believes that the child has "grown-up" characteristics, and therefore $D_t(\hat{h}^u)$ is insensitive to $\hat{h}^u$ and the effect of uncertainty about human capital (the second term in (1.9a)) is small. Since Φ is likely to remain less than 1 even with some shocks to the parent's observations (the arrow $\mathbf{S_a}$ in Fig. 1.1), the parent's belief is likely to be stable and to converge monotonically. In particular, if the initial uncertainty is less than $\bar{\sigma}_m^2$, the process would never diverge due to any shock. The parent's knowledge about the child improves as each new observation becomes available, and the expectation

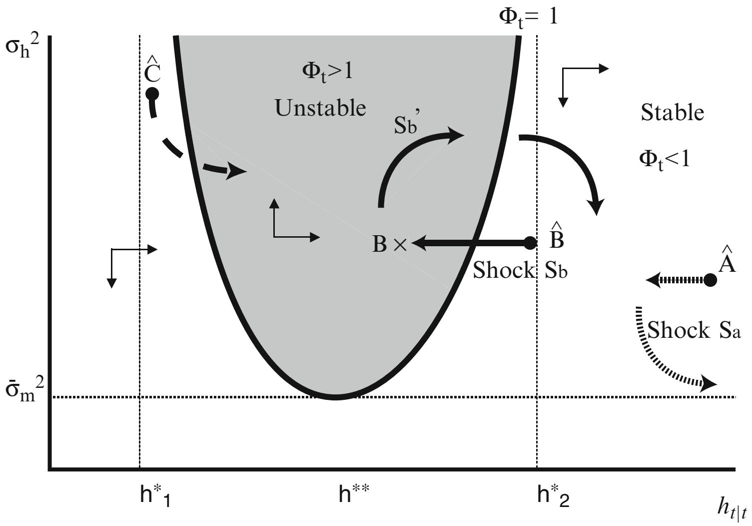


Fig. 1.1 Dynamics of the parent's expectation process

tends to become unbiased after many observations.[30] As in (1.8a), the slope of incentive, b, is increasing over time as the parent collects more information about the child. The child's induced effort is also increasing over time and approaches the first-best level as the child matures and the parent becomes confident about the child's human capital.

We now show that the probability of "punishment" decreases over time if the parent's belief converges. First, we construct *the equilibrium distribution of the parent's interaction* from the distribution of the parent's observed interactions, $b_t \mathrm{E}[a_t|I_t]$, defined in Sect. 3.3. Second, we define a statistic of the parent's observed negative interactions that measures the probability of "non-punishment." Finally, we examine how this statistic changes when the parent's belief is converging.

The parent's observed interaction in equilibrium, e_t^*, is defined and evaluated as follows:

$$
\begin{aligned}
e_t^* &\equiv b_t^* \mathrm{E}[a_t^*|I_t] = b^*(h - \hat{h} + a^* + \nu) = b^*[h - \hat{h} + \nu + \underline{a} + \psi b^* H^\gamma + \psi\phi D_t(h)] \\
&\cong b_t^*[h - \hat{h} + \nu + \underline{a} + \psi b^* H^\gamma + \psi\phi D_t(\hat{h}) + (h - \hat{h})\psi\phi D_t'(\hat{h})] \\
&= b^*[(h - \hat{h})(1 + \psi\phi D'(\hat{h})) + \nu + \underline{a} + \psi b^* H^\gamma + \psi\phi D_t(\hat{h})],
\end{aligned}
$$

where (1.6) is used to obtain the third line from the second and $D_t(h)$ is linearly approximated to derive the fourth line from the third. Since ν and h are Gaussian and uncorrelated with each other, when the parent has a belief $(\hat{h}, \sigma_h^2)$ at the beginning of the period and if this belief is unbiased, the equilibrium distribution of the parent's interaction, e_t^*, is defined as $N(\mu_e, \sigma_e^2)$, where

$$
\begin{aligned}
\mu_e &= b^*(\underline{a} + \psi b^* H^\gamma + \psi\phi D_t(\hat{h})), \\
\sigma_e^2 &= b^{*2}[(1 + \psi\phi D_t(\hat{h}))^2 \sigma_h^2 + \sigma_\nu^2].
\end{aligned}
\tag{1.12}
$$

The expected value of the parent's interaction is higher when the parent expects a high effort level due to a high expectation of the child's human capital (large $\hat{h}$) or chooses strict discipline (large slope of incentive, b^*). This also has a scale effect on both the expected value and the standard error.

Suppose that we recognize that parental behavior becomes "punitive" when the parent's interaction falls below a certain threshold level.[31] Since e_t^* is Gaussian, the probability that the parent does not become punitive is described by a "non-punishment" statistic, $t_e \equiv \mu_e/\sigma_e$. When t_e is large, it is unlikely that the parent's interaction falls below the threshold level. Using (1.12) and (1.7b'), this is evaluated as:

---

[30]Since the process is not stationary due to the finite horizon and the filter is applied to a linearized process, the expression "unbiased after many observations" must be interpreted as an approximation where the effects of both the initial and the terminal conditions can be ignored.

[31]The choice of the threshold level does not affect the qualitative nature of the following discussion.

$$t_e = \mu_e/\sigma_e = \frac{b^*(\underline{a} + \psi b^* H^\gamma + \psi\phi D_t(\hat{h})}{b^*[(1+\psi\phi D_t'(\hat{h}))^2\sigma_h^2 + \sigma_\nu^2]^{1/2}} = \frac{(\underline{a} + \psi b^* H^\gamma + \psi\phi D_t(\hat{h})}{[(A'(\hat{h}))^2\sigma_h^2 + (K/b^* H^\gamma Q_t)]^{1/2}}. \tag{1.13}$$

Since b^* is decreasing in σ_h^2 from (1.8a), it is found that $\frac{\partial t_e}{\partial(\sigma_h^2)} < 0$. This is because, as the parent becomes more certain about the child's human capital, she spends more time with the child and chooses stricter discipline to induce a greater effort that makes a low level of interaction less likely. Thus, we have proved the following proposition:

Proposition 2 *When the parent's belief is unbiased, as the uncertainty about the child becomes smaller, t_e becomes larger and the parent employs low level of interactions less often.*

Therefore, when the parent's belief converges over time, the probability of parental interaction below a certain level ("punishment") tends to decrease over time. The average observation in most "normal families" corresponds to this case. Along such an equilibrium path, a parent increases and maintains her "fair" control over the child and the probability of actual punishment decreases. This can be interpreted as what psychologists have called the "authoritative" parenting style (Baumrind 1967).

In the above discussion we have looked at only the partial effect of the change in the level of uncertainty on the parent's interaction. Equation (1.13) also shows the effect of the estimated human capital level. In the case where the parent's belief is convergent, the child is likely to develop his or her human capital rapidly with increasing effort and time inputs (see (1.8)). When the child's human capital is developing quickly and the parent is estimating the child's human capital in an unbiased manner, punitive interactions will be even less likely because the child will achieve a large discount factor more quickly. Therefore, the above result need not be altered fundamentally. Additionally, it is important to assume that the estimate of the child's human capital is unbiased to show this result. If the estimate is biased, it affects the parent's estimate of the child's effort, and the distribution of the parent's interactions. In particular, if it is positively biased, the parent would underestimate the child's effort and a negatively biased interaction is more likely than in the case of an unbiased belief. In fact, such a negative bias may prevail when the belief is diverging—an even worse combination—as shown in the next case.

(*b*) *Case of diverging belief—pathological family*

Using Fig. 1.1, we now illustrate how a unrealistically high expectation may lead to negatively biased interactions—child maltreatment.[32] To clarify the process of child maltreatment in our model, Fig. 1.2 shows a sequence of actions that leads to maltreatment with the relevant equation numbers. As shown by $\hat{\mathbf{B}}$ in Fig. 1.1,

[32]A stable equilibrium path may follow if the true value *happens* to be close to the expectation, although this is unlikely when the parent's uncertainty about the child is large. We focus on the case of very high, rather than very low, expectation because of its empirical relevance to child abuse.

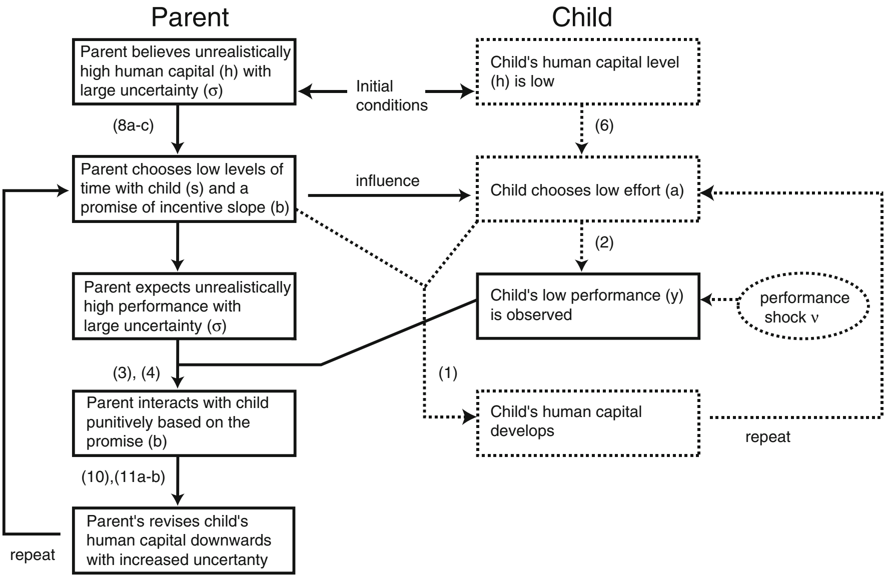


Fig. 1.2 A sequence of actions when the risk of child maltreatment increases
Note: The dotted boxes and arrows indicate that the states and the actions are not observed to the parent. The numbers in the parentheses indicate the relevant equations

suppose that, initially, or after observing a couple of the child's "lucky" performances, the parent's expectation of the child's human capital becomes unreasonably high relative to the true level of human capital while there is still a high level of uncertainty. The parent chooses small b and s according to (1.8a)–(1.8b). Since the child's true human capital level is low, he chooses the effort level according to (1.6) that is lower than the parent's expectation. The parent then observes the child's unexpectedly poor performances, administers punishment increasingly, and revises her belief of human capital downwards (the arrow $\mathbf{S_b}$ in Fig. 1.1). The child's human capital develops according to (1.1), but its speed is slow since parental time and child's effort are small. As the parent lowers her expectation, the expectation process becomes less stable because the child's effort is more sensitive to changes in the level of the child's human capital than before. If Φ becomes greater than 1, the beliefs regarding the child's human capital tend to diverge endogenously and the parent loses confidence in the child's characteristics and unobserved effort. ("*What is that kid thinking!*") As long as the belief lies in the unstable domain, the parent's expectation ($\hat{\mathbf{B}}$) cannot converge to the true value ($\mathbf{B}$) and easily reverts to and remains in the stable domain (the arrow $\mathbf{S_b'}$). The belief will be stable even though it is an upward biased false belief.[33] In this way, the parent tends to overestimate

[33]The notion of non-ergodicity may help in explaining this situation, although we cannot define it correctly due to the finite-horizon nature of the model. A system is called "non-ergodic" if the

the child's human capital. She tends to be disappointed by the child's performance and tends to punish him more often until the true value of human capital actually progresses into the stable domain.

As the true level of human capital develops and grows beyond h_2^*, the parent's expectation tends to stabilize because the child's characteristics become less sensitive to changes in the level of human capital, and it becomes easier for the parent to have a correct expectation. There is still some risk of getting into a bad cycle if the child experiences many unfortunate shocks that push the parental expectation downward beyond the $\Phi = 1$ curve, but it becomes increasingly less likely as the child grows.

When the parent remains unsure about the child's characteristics, Proposition 2 predicts that the probability that the parent's interactions become negatively-biased remains large if there is no bias in the parent's belief. If there is a positive bias in the belief, the result is worse, as is predicted above: the evaluation of the child's effort becomes negatively-biased, and the probability of punitive interactions is even larger than in the case without bias. Thus, during the process of a diverging belief, it is highly likely that we will observe a higher frequency of punitive interactions over time unless shocks to the child's performance quickly lead the parent's expectation and the actual development of the child toward a stable state.

The expectation may not stabilize very quickly if the development of the child's human capital is slow or the parent's initial uncertainty is large for some reason (bad luck, bad environment, etc.). For example, if the expectation process starts at $\hat{\mathbf{C}}$ in Fig. 1.1, the process inevitably passes through the unstable domain and takes a long time to escape from a situation of false beliefs about the child. After a series of positive observation shocks, the parent is easily trapped in the belief that the immature child is already mature and is likely to have a negatively biased opinion afterward. ("*This kid must be mature enough not to do such a stupid thing!*") Then, a divergent belief implies a delay in the child's development, because the child's effort and time spent with the parent decrease if uncertainty regarding the child's human capital is increasing (see (1.8a)). These are frequently observed actual phenomena in families with abuse (Wolfe 1987, p. 34).

The endogeneity of and the parental uncertainty over the child's time-preference play key roles in the emergence of child maltreatment. Since a child's effort depends on his or her own time-preference, the transmission of uncertainty from the child's preferences to effort and human capital development implies that parental uncertainty regarding the child's hidden characteristics may be magnified over time. In the absence of endogenous discounting, there is no uncertainty over the child's effort, new information on the child's performance always improves the knowledge

time average is not the same as the ensemble average. In such a case, a different initial condition may generate a path with a different time average. A simulation in finite horizon can show that the diverging parent's expectation process is locally explosive, and the (local) time average can become significantly different from the ensemble average.

of the child's human capital, and parental expectations and interactions are stable and normal under the standard production process.

(c) Numerical simulations

In order to give a visual illustration of the converging and diverging beliefs, Fig. 1.3 shows results from a numerical simulation of the solution of child human capital development and parental behavior and expectations.[34] Figure 1.3a shows the equilibrium development of h_t and $\hat{h}_t$ over time with the parental estimation error, σ_h, when the initial expectation is higher than the true value, that is $h_1 < \hat{h}_1$ at period 1. Until about period 20, the expectation process exhibits an unstable path, the parental beliefs tend to be persistently positively biased, and the child's human capital develops slowly. As the child's human capital develops beyond period 20, both the estimation error and the parental expectation become converging, and the speed of human capital development increases. The driving forces of this dynamics are depicted in Fig. 1.3b. Each solid curve and dotted curve shows the evolution of Φ_t and t_e, the stability criteria and the non-punishment statistic, respectively. It is confirmed that, as predicted by the lemma, if and only if Φ_t is greater than one, the expectation process diverges, and that under such a circumstance, t_e may decrease, i.e., the parental utility transfer may become lower and punitive over time. Figure 1.3c illustrates the dynamics of parental behavior, b_t and s_t. As predicted in Sect. 4.3, when parental uncertainty increases over time, the levels of b_t and s_t decrease.[35] However, once the parental expectation enters into a stable cycle, the levels of parental incentive and time to be spent with the child increase over time, and the development of child human capital takes off as shown in Fig. 1.3a. Once the development enters this phase, as suggested by the evolution of t_e in Fig. 1.3b, the risk of being maltreated becomes extremely low.

5.2 *The Characteristics of a Parent and a Child at Risk of Abuse*

We have seen that, when $\Phi = \frac{F'(\hat{h}^u)^2}{A'(\hat{h}^u)^2(\sigma_h^2/\sigma_v^2)+1} = \frac{F'(\hat{h}^u)^2}{A'(\hat{h}^u)^2(s^* \cdot \sigma_h^2/K)+1} < 1$, the parent's expectation process converges and the probability of punitive interactions decreases over time. Therefore, a family with parameters that reduce Φ_t is unlikely to fall into

[34]The parameter values used in the simulation are listed in Appendix "Numerical Simulation of the Equilibrium Dynamics".

[35]Therefore, our model predicts that the process of child maltreatment is associated with "decreasing" time spent with child. One referee points out that our formulation does not explain why an abusive parent bothers to take a time-consuming way to punish the child. I believe that child maltreatment can be time-saving if, for example, it means "parental ignorance or abundance." On the other hand, there is little question about the fact that the monitoring of and learning about the child is time-consuming.

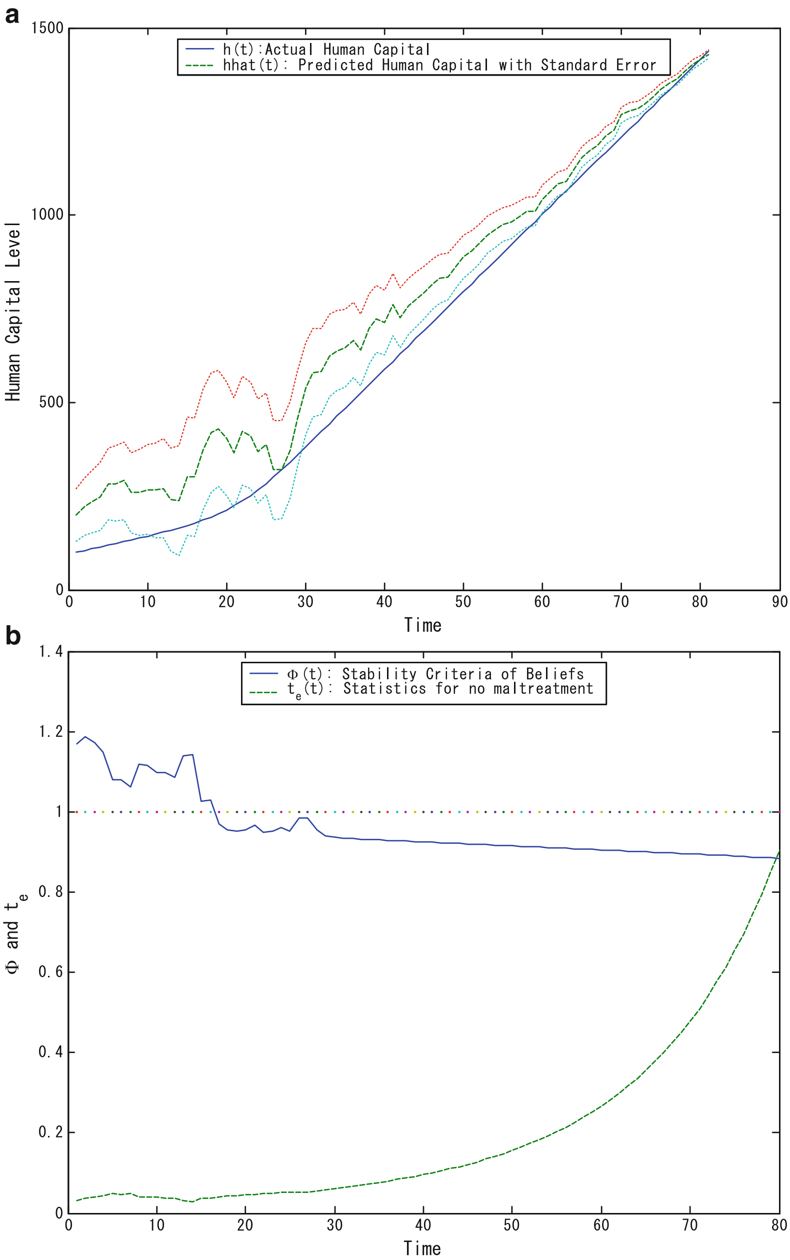


Fig. 1.3 A simulated dynamics of parental interactions, child development, and parental beliefs. (**a**) Dynamics of actual and predicted child human capital development. (**b**) Dynamics of Φ_t (stability criteria of parental beliefs) and t_e, (non-punishment statistic). (**c**) Dynamics of parental interactions

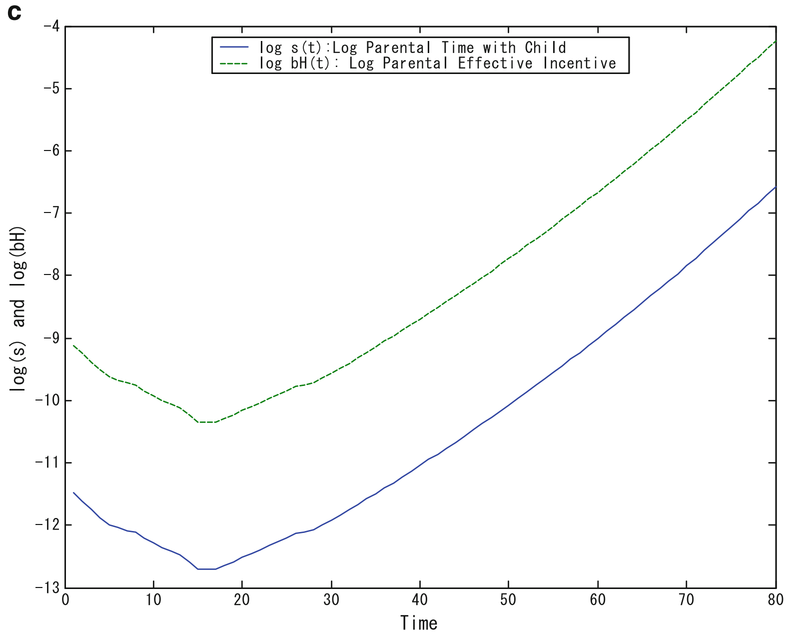


Fig. 1.3 (continued)

an unstable equilibrium path, or "a cycle of abuse." Since the parameters that have a positive effect on s^* will have such a property, most of the results in the following proposition are straightforward from the discussion in Sect. 4.3.

Proposition 3 *All else equal, a family is less likely to follow a path with persistently punitive interventions if (i) the productivity of time in the child's human capital (φ) is large, (ii) the wage rate (π) is low, (iii) the parent is good at observing the child (small K), (iv) the child is old (T-t is small), or (v) the child's natural level of effort is high (large $\underline{a}$).*

Proof (i), (ii), (iv), and (v) are clear from the expression of s_t^* in (A1) in Appendix "Mathematical Details of the Parenting Plan". To prove (iii), notice that $\Phi_t = \frac{F'(\hat{h}_{t|t})^2}{A(\hat{h}_{t|t})^2(s_t^* \cdot \sigma_{ht}^2/K)+1} = \frac{[(1-\delta)+\phi^2\psi D_t'(\hat{h}_{t|t})]^2}{1+[1+\phi\psi D_t'(\hat{h}_{t|t})]^2 \frac{\underline{a}+\phi\psi D_t(H)-KR\alpha Q_t}{KQ_t(R\alpha-\psi/\sigma_{ht}^2)}}$, Φ_t tends to be small, and the parent's equilibrium process is likely to be stable for a small K. □

Among the above results, (i), (ii), (iv), and (v) would encourage the parent to stay with the child longer, and thereby work to reduce the variance of the observation error (σ_v^2). (iii) would make the parent choose to spend less time with the child, but knowledge about the child would improve due to the superior observation ability.

A couple of remarks follow. First, some of the above results may seem counterfactual. For example, we usually observe that parents tend to spend less time with older children in normal families. However, if we interpret schooling as an extension of education at home, the implication of (iv) is not inconsistent with the fact that older children spend a longer time at school under teachers' supervision. (ii) might also seem to contradict the observations, because some researchers have documented that unemployment and poverty are likely to increase the risk of abuse. For example, the recent empirical results of Paxson and Waldfogel (2002) suggest that state-level share of children living in poverty increases child maltreatment. As is discussed in Sect. 4.2, our prediction is driven by the fact that, in the absence of income effect, a higher wage rate generates a substitution away from being with children and toward paid work, resulting in increased parental uncertainty.[36] Without income effect, a parent with a higher value of time would like to save time in favor of a steeper slope of incentive, b. If we allowed an income effect, it should operate in the direction to decrease the labor supply of female workers as in the standard labor supply model and to increase time spent with the child that leads to improved information.[37] Interestingly, Paxson and Waldfogel empirically found positive effects of per capita income on maltreatment rate under some econometric specifications that control state poverty rate. Although their empirical results are not directly comparable to our theoretical predictions, it implies the complicated nature of the effects of income and wage on child maltreatment. We also have to remember that poverty is related to a low level of human capital on the part of the parent, which can have several other effects that do not operate through income as discussed below. Further, unemployment can lead to parents' "frustration," which is not modeled here. Therefore, we have to be cautious in interpreting (ii).

Second, we may interpret some of the above effects as resulting from the parent's "psychiatric" characteristics. For example, suppose that a drug-addicted or alcoholic parent has abnormally low observation ability (very large K). Then the parent's perception of the child's effort becomes inaccurate, and the parent's expectation tends to be unstable. Not only does this contribute to high risk of child abuse but it also causes a delay in the child's development. Since Q_t is increasing in K, we can see from (1.8) that bH^{γ}, s, and a are all smaller. The child is less motivated to make an effort, and the child's development is delayed as a consequence. The uncertainty about the child's human capital will tend to remain large for a long time, and the parent will tend to maintain a negatively biased view of the child.

Third, as we have seen in the previous paragraph, the stabilizing effect is not the only way through which these characteristics affect the risk of child abuse. For example, a larger $\underline{a}$ not only has a stabilizing effect on the parent's expectation

[36] Paxson and Waldfogel (2002) also find that the existence of single working mothers tends to increase the likelihood of maltreatment. To the extent that single working mothers lack sufficient time for monitoring children, their results conform to our prediction.

[37] Note that a higher b alone does not necessarily mean "risk of maltreatment." For example, as a parent has more information (lower uncertainty) about the child's human capital, she can choose a higher b—stricter discipline. If a higher b is combined with unbiased (or at least, converging) belief about human capital, the model predicts that the child's development is accelerated. We sometimes observe that busy and educated parents tend to compensate children for the lack of time at home by more promises, strict rules, and rewards (gifts).

process but also a positive effect on the child's development (both s and a become larger). Obviously such a child is at less risk of facing abuse because he can reach a high level of human capital faster.

Finally, it may be surprising to note that the parent's level of human capital (H) has theoretically mixed effects. However, if H is larger, (i) the speed of a child's development is faster because the parent's teaching is more effective, (ii) K is likely to be small (less observation errors), and (iii) such a family tends to have fewer children (more time to spend with each child), then it would be reasonable to think that the child's development tends to be faster and that there will be less risk of a divergent parental belief and child abuse. Moreover, it must be noted that the effects of some other key parameters are, surprisingly, uncertain. For instance, the effect of the degree of altruism, α, is unclear because, as we discussed in Sect. 4.3, a more altruistic parent does not necessarily choose to spend more time with the child. Parental love may lead to a more relaxed level of control (smaller b) so that the parent does not choose to improve the accuracy of observations by spending large amounts of time with the child. As a result, the parent's expectation process can become divergent.

Our model provides several insights into the prevention of child maltreatment. First, although the parent always maximizes her subjective expected utility, if we could exogenously remove the abnormal bias and error in the parent's expectation of the child with low costs, a family's welfare could be improved. This is particularly important when an unstable relationship lasts long relative to the finite length of the parent-child relationship. Obvious intervention methods implied would include (i) providing the parent with the correct knowledge about child development or a shock to change her belief or (ii) providing monitoring and service (child care) to reduce the monitoring error (σ_ν^2). These have been recently mentioned by several researchers in the literature. For example, Wolfe (1991) wrote, "Many [young, socially disadvantaged parents] lack knowledge about infant development and therefore have inappropriate expectations for their infants' behavior..." (p. 90) and advocated a training program for parents at risk of abuse that includes "setting reasonable expectations for children's emotional and behavioral development" (p. 120).

Second, a low natural level of the child's effort in our model ($\underline{a}$) may be interpreted as the child's developmental characteristics that are genetically determined. There is some evidence that children with cognitive or emotional problems tend be abused more often than otherwise, consistently with Proposition 3(v). Any policy that identifies and monitors such children at early stage of development is predicted to prevent the occurrence of maltreatment.

Finally and more broadly, the model emphasizes the importance of learning about a child's characteristics on the part of the care providers for the healthy development of the child, and this may apply not only parents but also other adults who take care of children. It is often argued in psychology literature (Berk 1991, p. 426) that a long-term relationship between the child and the care-giver is important for the quality of child care. Furthermore, the recent literature identifies that teacher training is one of few consistent determinants of child care quality (Blau 2001, p. 131). Our model is consistent with such general idea about the quality of child care.

6 Conclusion

In this chapter, we attempted to shed light on the role of expectations in the interaction between a parent and a child, which may lead to child maltreatment in equilibrium. We constructed a model that explains the dynamic interrelationship between the development of the child's human capital, endogenous rate of time preference, and the parent's interactions and beliefs. The equilibrium dynamics are state-dependent, because a child's chosen effort both affects and is affected by the development of the child's characteristics. The model predicts that if the initial uncertainty regarding the child's human capital is large, the parent's expectation process of the child's human capital might diverge and be negatively biased due to the endogenous nature of the parent's expectation formation process. It is shown, both analytically and numerically, that the divergent beliefs may lead to the parent's unrealistically high expectations and persistently punitive interactions with the child, which explain the emergence of child maltreatment or abuse. The model also identifies the characteristics of families at risk of diverging beliefs and persistent negative interactions that can be empirically testable.

The intension is not to deny the existence of child maltreatment due to mental illness nor is it to insist that we can ignore other potential factors that might influence parent-child interactions. Other factors not addressed here, such as substance abuse and prior maltreatment as a child, would definitely be important. Nonetheless, our model can add several insights to the existing psychological literature of child maltreatment. First, our model provides rational perspectives that place maltreatment in a wide context of parenting styles. Our equilibrium concept proves useful in explaining many important characteristics of child maltreatment, which have been empirically established in recent literature. Second, our focus on time and information resources of parents can suggest policies that are unconventional but still implementable easily. For example, our theory suggests that increasing parental time, not necessarily parental wages, would prevent child maltreatment. Finally, our model has a potential that can be applied not only to child maltreatment but also to any pathological relationships between a principal and an agent sharing common interests, such as domestic violence between husband and wife.

Appendices

Psychological Theories of Child Maltreatment

The earliest model, the *psychopathological* explanation, lays emphasis on the parents' underlying emotional disturbance and distinguishes abusers from nonabusers by their personality dimensions. However, researchers have found that there is not much evidence to support mental illness as a major contributor to child abuse. In addition, they have not been successful in finding consistent characteristics of abusers. The problem with this approach is that "by believing abusers to be 'crazies,' remedial and preventive efforts are directed away from the general population, where abuse is most likely to occur" (Zigler and Hall 1989). Subsequently, researchers have shifted their focus from purely psychopathological characteristics to psychological processes. The *social-cultural* approach, which was extended to the ecological model by Belsky (1980), has adopted a broader socio-environmental perspective. This approach emphasizes the stress induced by the social structure in which a parent lives. By focusing on the interactional process between the parent and child, the recently influential *social-interactional* approach has tried to explain why only some parents with such characteristics and conditions as those described in the socio-cultural model become abusive (Wolfe 1987).

The First Order Condition of the Child's Decision Problem

The child's Lagrangian is

$$L = E\left[-\exp\left[-r\left(\sum_{\tau=t}^{T}\left(\frac{1}{1+\rho_{ct}}\right)^{\tau-t}\{(s_\tau + b_\tau(h_\tau + a_\tau + \nu_\tau - \hat{h}_\tau))H^\gamma \right.\right.\right.$$
$$\left.\left.\left. -\frac{(a_\tau - \underline{a})^2}{2\psi}\} + \left(\frac{1}{1+\rho_{ct}}\right)^{T-t+1} B\cdot h_{T+1}\right)\right] | I_{t-1}\right].$$

The first order condition for a_t is

$$\frac{\partial L}{\partial a_t} = \frac{\partial}{\partial a_t} E\left[-\exp\left[-r\left(b_t a_t H^\gamma - \frac{(a_t-\underline{a})^2}{2\psi}\right.\right.\right.$$
$$+\sum_{\tau=t+1}^{T}\left(\frac{1}{1+\rho_{ct}}\right)^{\tau-t} b_\tau(h_\tau - \hat{h}_\tau)H^\gamma\left(\frac{1}{1+\rho_{ct}}\right)^{T-t+1} B\cdot\{(1-\delta)^{T-t}h_t$$
$$\left.\left.\left.+\sum_{\tau=t}^{T}(1-\delta)^{T-\tau}(\varphi s_\tau H^\gamma + \phi a_\tau)\}\right)\right]\right| I_{t-1}\Big] = 0.$$

The calculation is straightforward, except for the third term in $\exp[-r(\cdot)]$ (discounted sum of $b_\tau(h_\tau - \hat{h}_\tau)H^\gamma$). To evaluate this, take the case where $\tau = t + 1$. Care must be taken in that $\hat{h}_{t+1}$ depends on the parent's estimate of the child's effort at t after observing a new performance. We therefore have,

$$\begin{aligned}
&b_{t+1}(h_{t+1} - \hat{h}_{t+1}) \\
&= b_{t+1}\{(1-\delta)h_t + \phi a_t + \phi s_t H^\gamma - ((1-\delta)\hat{h}_t + \phi E[a_t|I_t] + \phi s_t H^\gamma)\} \\
&= b_{t+1}\{(1-\delta)(h_t - \hat{h}_t) + \phi(a_t - (h_t + a_t + \nu_t - \hat{h}_t))\} \\
&= b_{t+1}\{(1-\delta)(h_t - \hat{h}_t) - \phi\nu_t\}.
\end{aligned}$$

Without the knowledge of the way in which the parent's choice of b_{t+1} depends on changing preferences, this part is beyond the child's control and its partial derivative with respect to a_t is zero. Thus, the usual "ratchet effect" is absent. This reasoning applies to other terms where $\tau > t + 1$ due to linearity. We have,

$$\frac{\partial L}{\partial a_t} = \left[b_t H^\gamma - \frac{a_t - \underline{a}}{\psi} + \left(\frac{1}{1+\rho_{ct}}\right)^{T-t+1} B \cdot (1-\delta)^{T-t}\phi \right] \cdot L = 0.$$

Then (1.5) follows. Note that we used only the law of motion of human capital, the parent's observation equation and the incentive schedule as the child's knowledge to obtain this result. Also there is no "career concern" (Gibbons and Murphy 1992) here since incentive is provided based on the estimated effort, rather than human capital.

The First and Second Order Conditions of the Parent's Decision Problem

Using (1.6), the Lagrangian of the parent at period t is written as

$$\begin{aligned}
L = E\Bigg[U \Bigg(&\sum_{\tau=t}^{T} \left(\frac{1}{1+\rho_{ct}}\right)^{\tau-t} \{\pi(1-s_\tau)H^\gamma \\
&+ \alpha(s_\tau + b_\tau(h_\tau + a_\tau^* + \nu_\tau - \hat{h}_\tau))H^\gamma \\
&- \frac{(a_\tau - \underline{a})^2}{2\psi}\} + \alpha\left(\frac{1}{1+\rho_{ct}}\right)^{T-t+1} B \cdot h_{T+1} \Bigg) \Bigg| I_{t-1} \Bigg] \\
= E\Bigg[U \Bigg(&\sum_{\tau=t}^{T} \left(\frac{1}{1+\rho_p}\right)^{\tau-t} \{\pi(1-s_\tau)H^\gamma
\end{aligned}$$

$$
\begin{aligned}
&+ \alpha(s_\tau + b_\tau(h_\tau + \underline{a} + \psi b_\tau H^\gamma + \phi\psi D_\tau(\hat{h}_\tau) + \nu_\tau - \hat{h}_\tau))H^\gamma \\
&- (\psi b_\tau H^\gamma + \phi\psi D_\tau(\hat{h}_\tau))^2/2\psi\} \\
&+ \alpha B \left(\frac{1}{1+\rho_p}\right)^{T-t+1} \times \left[(1-\delta)^{T-t+1} h_t\right. \\
&\left.\left.\left.+ \sum_{\tau=t}^{T} (1-\delta)^{T-\tau}(\varphi s_\tau H^\gamma + \phi a_\tau)]\right]\right) \right| I_{t-1} \Bigg] \\
&= E\left[U\left(\sum_{\tau=t}^{T}\left(\frac{1}{1+\rho_p}\right)^{\tau-t}\{\pi(1-s_\tau)H^\gamma\right.\right. \\
&+\alpha(s_\tau + b_\tau(h_\tau + \underline{a} + \frac{\psi}{2} b_\tau H^\gamma + \nu_\tau - \hat{h}_\tau))H^\gamma - \alpha\frac{\psi}{2}\phi^2 D_\tau(\hat{h}_\tau)^2\Big\} \\
&+ \alpha B \left(\frac{1}{1+\rho_p}\right)^{T-t+1} \times \left[(1-\delta)^{T-t+1} h_t\right. \\
&\left.\left.\left.+ \sum_{\tau=t}^{T} (1-\delta)^{T-\tau}(\varphi s_\tau H^\gamma + \phi(\underline{a} + \psi b_\tau H^\gamma + \phi\psi D_\tau(\hat{h}_\tau)))]\right]\right) \right| I_{t-1} \Bigg].
\end{aligned}
$$

Using the moment conditions of Gaussian error terms and the first order Taylor expansion of $D_t(\hat{h}_t)$, it is evaluated as

$$
\begin{aligned}
L \simeq &-\exp[-R[\pi(1-s_t)H^\gamma + \alpha\{s_t + b_t(\underline{a} + \tfrac{\psi}{2} b_t H^\gamma)\}H^\gamma - \alpha\tfrac{\psi}{2}\phi^2 D_t(\hat{h}_t)^2 \\
&\qquad +\alpha D_t(H)\left(\hat{h}_t + \phi(\underline{a} + \psi b_t H^\gamma + \phi\psi D_t(\hat{h}_t)) + \varphi s_t H^\gamma\right)]] \\
&\cdot \exp\left(\tfrac{1}{2}[R^2\alpha^2[b_t^2(\sigma_{\nu t}^2 + \sigma_{ht}^2)H^{\gamma 2} + \tfrac{\psi^2}{4}\phi^4 Var(D_t(\hat{h}_t)^2)\right. \\
&\qquad \left.+D_t(H)^2\sigma_{ht}^2(1 + \phi^2\psi D_t'(\hat{h}_t))^2]]\right) \\
&\cdot\mathrm{E}\left[\exp[-R[\sum_{\tau=t+1}^{T}\left(\tfrac{1}{1+\rho_p}\right)^{\tau-t}[\pi(1-s_\tau)H^\gamma\right. \\
&\qquad +\alpha\{(s_\tau + b_\tau(h_\tau + a_\tau^* + \nu_\tau - \hat{h}_\tau))H^\gamma - \tfrac{(a_\tau^* - \underline{a})^2}{2\psi}\}] \\
&\qquad\quad \left.+\alpha D_t(H)\left(\sum_{\tau=t+1}^{T}(1-\delta)^{t-\tau}(\phi a_\tau^* + \varphi s_\tau H^\gamma)\right)]]|I_{t-1}\right].
\end{aligned}
$$

Then, we have the following first order condition:

$$
\begin{aligned}
\frac{\partial L}{\partial b_t} &= -R\alpha H^\gamma[(\underline{a} + b_t\psi H^\gamma + \psi\phi D_t(H) - R\alpha b_t H^\gamma(K/s_t + \sigma_{ht}^2)] \cdot L = 0, \\
\frac{\partial L}{\partial s_t} &= [-R(-\pi H^\gamma + \alpha H^\gamma + \alpha D_t(H)\varphi H^\gamma) - \frac{R^2\alpha^2[b_t^2(K/s_t^2)H^{\gamma 2}]}{2}] \cdot L = 0,
\end{aligned}
$$

which lead to (1.7a) and (1.7b). The second order condition is

$$\begin{aligned}\frac{\partial^2 L}{\partial b_t^2} &= [-R\alpha H^\gamma \left(\psi H^\gamma - R\alpha H^\gamma (K/s_t + \sigma_{ht}^2)\right)] \cdot L \\ &= R\alpha H^{2\gamma} \left(R\alpha (K/s_t + \sigma_{ht}^2) - \psi\right) \cdot L < 0, \\ \frac{\partial^2 L}{\partial s_t^2} &= \frac{R^2\alpha^2 b_t^2 K H^{\gamma 2}}{s_t^3} \cdot L < 0, \\ &\frac{\partial^2 L}{\partial b_t^2}\frac{\partial^2 L}{\partial s_t^2} - \left(\frac{\partial^2 L}{\partial s_t \partial b_t}\right)^2 \\ &= [-R\alpha H^{2\gamma} \left(\psi - R\alpha (K/s_t + \sigma_{ht}^2)\right)] \cdot \frac{R^2\alpha^2 b_t^2 K H^{2\gamma}}{s_t^3} \cdot L^2 \\ &\quad -[R^2\alpha^2 H^{2\gamma} b(-K/s^2)]^2 \cdot L^2 \\ &= R^3\alpha^3 H^{4\gamma} b^2 (K/s^3) \cdot (R\alpha\sigma_{ht}^2 - \psi) \cdot L^2 > 0.\end{aligned}$$

Therefore, $R\alpha\sigma_{ht}^2 - \psi > 0$ is sufficient in order to satisfy the second order condition.

Mathematical Details of the Parenting Plan

Using $Q_t = \left(\frac{2H^\gamma}{R\alpha^2 K}\left[\pi - \alpha - \alpha\varphi D_t(H)\right]\right)^{1/2}$, (1.8a) can be evaluated as follows:

$$\begin{aligned}b_t^* H^\gamma &= \frac{1}{R\alpha\sigma_{ht}^2 - \psi}(\underline{a} + \psi\phi D_t(H) - KR\alpha Q_t) \\ &= \frac{1}{R\alpha\sigma_{ht}^2 - \psi}(\underline{a} + \psi\phi D_t(H) - (2RKH\left[\pi - \alpha - \alpha\varphi D_t(H)\right])^{1/2}),\end{aligned}$$

$$s_t^* = \frac{1}{Q_t(R\alpha\sigma_{ht}^2 - \psi)}(\underline{a} + \psi\phi D_t(H) - KR\alpha Q_t) \tag{1.14}$$

$$= \frac{1}{R\alpha\sigma_{ht}^2 - \psi}(\frac{\underline{a} + \psi\phi D_t(H)}{Q_t} - KR\alpha) \tag{1.15}$$

$$\begin{aligned}&= \frac{\alpha}{R\alpha\sigma_{ht}^2 - \psi}[\frac{\underline{a} + \psi\phi D_t(H)}{\left(\frac{2H^\gamma}{RK}\left[\pi - \alpha - \alpha\varphi D_t(H)\right]\right)^{1/2}} - KR] \\ &= \frac{\alpha KR}{R\alpha\sigma_{ht}^2 - \psi}[\frac{\underline{a} + \psi\phi D_t(H)}{(2H^\gamma RK\left[\pi - \alpha - \alpha\varphi D_t(H)\right])^{1/2}} - 1].\end{aligned}$$

Therefore it is seen that R has a negative effect on both $b_t^* H^\gamma$ and s_t^* while the effect of α is ambiguous.

The Derivation of the Parent's Expectation Process (The Kalman Filter)

Rewriting the equilibrium equations (1.9) here,

$$\begin{aligned} h_{t+1} &= F(h_t) + G(\sigma_{ht}^2), \\ y_t &= A(h_t) + C(\sigma_{ht}^2) + \nu_t, \end{aligned}$$

where

$$\begin{aligned} F(h_t) &= (1-\delta)h_t + \phi^2\psi D_t(h_t), A(h_t) = h_t + \phi\psi D_t(h_t), \\ G(\sigma_{ht}^2) &= \phi\left[\underline{a} + \frac{\psi}{R\alpha\sigma_{ht}^2 - \psi}\left(\underline{a} + \psi\phi D_t(H) - KR\alpha Q_t\right)\right] + \varphi s_t^* H^\gamma, \\ C(\sigma_{ht}^2) &= \underline{a} + \frac{\psi}{R\alpha\sigma_{ht}^2 - \psi}\left(\underline{a} + \psi\phi D_t(H) - KR\alpha Q_t\right). \end{aligned}$$

To derive the parent's expectation process, the Kalman filter is applied to the linearized version of the equilibrium system equation. First, we will consider the linear approximation of $F(\cdot)$ and $A(\cdot)$ with respect to h_t in the neighborhood of its estimate. We use $\hat{h}_{t|t}$ to approximate $F(\cdot)$ and $\hat{h}_t$ to have an initial approximation of $A(\cdot)$ as follows:

$$\begin{aligned} h_{t+1} &= F(\hat{h}_{t|t}) + F'(\hat{h}_{t|t})(h_t - \hat{h}_{t|t}) + G(\sigma_{ht|t}^2), \\ y_t &= A(\hat{h}_t) + A'(\hat{h}_t)(h_t - \hat{h}_t) + C(\sigma_{ht}^2) + \nu_t, \end{aligned} \tag{1.16}$$

where $F'(\hat{h}_{t|t}) = (1-\delta) + \phi^2\psi D_t'(\hat{h}_{t|t})$ and $A'(\hat{h}_t) = 1 + \phi\psi D_t'(\hat{h}_t)$ are the first-order Taylor coefficients of (1.9). After a sufficient iteration of updating $\hat{h}_t$ using (1.10), we have the following relations:

$$\begin{aligned} \hat{h}_{t|t} &\approx \hat{h}_{t|t} + \frac{A'(\hat{h}_{t|t})\sigma_{ht}^2}{A'(\hat{h}_{t|t})^2\sigma_{ht}^2 + \sigma_{\nu t}^2}(y_t - A(\hat{h}_{t|t}) - C(\sigma_{ht}^2)), \\ \sigma_{ht|t}^2 &= \sigma_{ht}^2 - \frac{(A'(\hat{h}_{t|t})\sigma_{ht}^2)^2}{A'(\hat{h}_{t|t})^2\sigma_{ht}^2 + \sigma_{\nu t}^2}. \end{aligned}$$

The parent uses this updated belief to construct her belief in the next period believing that (1.16) is the true system equation. Then, we have

$$\hat{h}_{t+1} = F(\hat{h}_{t|t}) + G(\sigma_{ht}^2),$$

$$\sigma^2_{ht+1} = F'(\hat{h}_{t|t})^2 \sigma^2_{ht|t} = F'(\hat{h}_{t|t})^2 \left(\sigma^2_{ht} - \frac{(A'(\hat{h}_{t|t})\sigma^2_{ht})^2}{A'(\hat{h}_{t|t})^2 \sigma^2_{ht} + \sigma^2_{\nu t}} \right)$$
$$= \frac{F'(\hat{h}_{t|t})^2 \sigma^2_{\nu t}}{A'(\hat{h}_{t|t})^2 \sigma^2_{ht} + \sigma^2_{\nu t}} \sigma^2_{ht},$$

which is the same as (1.11a).

Proof of Proposition 1

The stability condition in the lemma is $\Phi_t = \frac{F'(\hat{h}_{t|t})^2}{A'(\hat{h}_{t|t})^2(\sigma^2_{ht}/\sigma^2_{\nu t})+1} < 1$. Since $F' < 1$ automatically implies the stability, for a divergence to occur, it is necessary to have $F' > 1$. Therefore, the divergence requires the following inequality to hold given a belief $(\hat{h}_{t|t}, \sigma^2_{ht})$:

$$\sigma^2_{\nu t}/\sigma^2_{ht} > \frac{[1 + \phi\psi D'_t(\hat{h}_{t|t})]^2}{[1 - \delta + \phi^2\psi D'_t(\hat{h}_{t|t})]^2 - 1} \equiv Z_t(\hat{h}_{t|t}).$$

The question is whether there exists a combination of parameters with which there is a non-empty set of beliefs that can cause the expectation process to diverge. Let the two real roots of the equation, $F'(h) = (1 - \delta) + \phi^2\psi D'_t(h) = 1$, be h^*_1 and h^*_2 , which exist if $\rho_{ct} = \exp(-\eta h_t)$ with $\eta > 0$. It can be verified that for the range (h^*_1, h^*_2), $Z_t(h)$ is continuous, its minimum is attained at $h^{**} \in (h^*_1, h^*_2)$, and the value of $Z_t(h^{**})$ is $(1 - (1 - \delta - \phi)^2)/\phi^2$. Clearly $\lim_{h \to h^*_1 + 0} Z_t(h) \to +\infty$ and $\lim_{h \to h^*_2 - 0} Z_t(h) \to +\infty$. Therefore, for the above inequality to hold, it is necessary that the following inequality holds for a given value of σ^2_h and the given parameter values:

$$(K/s^*_t)/\sigma^2_{ht} = \sigma^2_{\nu t}/\sigma^2_{ht} > Z_t(h^{**}) = (1 - (1 - \delta - \phi)^2)/\phi^2. \tag{1.17}$$

If (1.17) holds, since both sides of (1.17) are independent of h and $Z_t(h)$ is continuous in h for $h \in (h^*_1, h^*_2)$, there exists a range of $\hat{h}_{t|t}$, $R(\sigma^2_h) = (h_1, h_2)$, such that any value of $\hat{h}_{t|t}$ in $R(\sigma^2_h)$ will cause the beliefs to diverge. It can be verified that $\sigma^2_{\nu t}/\sigma^2_{ht} = \frac{KQ_t(R\alpha\sigma^2_{ht} - \psi)}{\underline{a} + \phi\psi D_t(H) - KR\alpha Q_t}$ is increasing in σ^2_h since an increase in uncertainty of human capital discourages the monitoring of the child. Therefore, if the given set of parameter values satisfies $\lim_{\sigma^2_{ht} \to \infty}(\sigma^2_{\nu t}/\sigma^2_{ht}) = \frac{KQ_t R\alpha}{\underline{a} + \phi\psi D_t(H) - KR\alpha Q_t} > (1 - (1 - \delta - \phi)^2)/\phi^2$, we can find a value of σ^2_h and the associated range of $\hat{h}_{t|t}$ for the divergence. It is possible because the LHS of the inequality infinitely increases in K as long as the denominator is positive (the requirement for s^* to be positive) for any combination of the other parameters. This proves (i) and (iii). Using the result of the lemma, (ii) follows naturally.

Numerical Simulation of the Equilibrium Dynamics

Figure 1.3a–c are generated with the following parameter values and functional forms.

$T = 80, \phi = 0.7, \varphi = 0.01, \psi = 0.7, \delta = 0.001, \underline{a} = 7, K = 1.5, R = 2, \alpha = 0.9, \pi = 2, B = 50, H = 40{,}000, \gamma = 0.5, h_1 = 100, \hat{h}_1 = 200, \sigma_{h1}^2 = 5{,}000, \rho_i(h) = \exp(-0.02 \cdot h), i = p$ or c.

Addendum: "An Equilibrium Model of Child Maltreatment" as a Model of Child Development with Parent-Child Interaction[38]

This addendum discusses the relationship between Akabayashi (2006) and other models of human capital development in the family in the economics literature. Special attention is placed on models that consider a child's human capital development as an interactive process between a parent and their child.

Economists have used the human capital model to explain the relationship between parental income, educational investment for children, and their adult outcomes (Becker 1967). Empirical works that show the positive relationship between these variables are abundant. For example, an important work by Leibowitz (1974) showed that the IQ level of school children tended to correlate with the time invested by parents during preschool. Until the mid-1990s, however, the human capital theory only concerned monetary and time investments, and lacked elements of parents' psychological investment in their children. The unobserved aspects of family influences on children were called the "family background," potentially determined by the family's culture and traditions, genes, and environment, and it was assumed as given or "transmitted" by a law of motion (Becker and Tomes 1979, 1986).

In developmental psychology, a parent's behavior toward their child, or "parenting style," has been considered the crucial part of "investment" to foster the healthy development of the child (Baumrind 1967; Maccoby and Martin 1983). There is a wide variation in parenting style across different ethnicities, regions, and cultures (Bhatt and Ogaki 2012; Chua 2011; Steinberg et al. 1992). There is also an extreme parenting style, "child abuse and maltreatment," that is known to have

[38] This addendum has been newly written for this book chapter.

a lasting harmful effect on the child (Wolfe 1987). Akabayashi (2006) attempted to incorporate the psychological interactions between a parent and a child in the traditional human capital investment model that can explain a broad range of parent-child interactions, including child maltreatment, as an equilibrium. Influenced by an increasing use of the contract theory in macroeconomics and family economics in the 1990s (Bergstrom 1989; Bernheim et al. 1995), the Akabayashi model uses the technique to explain the parental choice of incentives toward a child. It is straightforward to interpret "praising good behavior and punishing bad behavior of a child" as a form of optimal incentives in contract theory; interestingly, the psychological counterpart is called the "optimum parenting style" (Garcia and Gracia 2009).

The Akabayashi model can be thought of as a generalization of a rational model of human capital development within a family that allows endogenous preferences and asymmetric information about hidden abilities, and thus a unique contribution to the theoretical background of the Nature vs. Nurture debate.

In 1994, the publication of *The Bell Curve* (Herrnstein and Murray 1994) created controversy in the United States, claiming that the intelligence measure significantly predicts many adult outcomes, and it is genetically determined. Several social scientists critically responded to this book, including Heckman (1995) and Cawley et al. (1997). His main criticisms were twofold: the intelligence measure, which appears to be correlated with the later social/economic life, is likely to be determined in an early childhood environment, and there are many inputs in human capital production that are unobservable and yet potentially modified by the social and educational policies. Heckman started to work on child development problems both theoretically and empirically using the microdata of children that were becoming widely available at that time, especially the National Longitudinal Survey of Youth—Child Supplement (CNLSY). The Akabayashi model is consistent with Heckman's criticism toward *The Bell Curve* in that an important part of family background is not just given at birth but is emerging endogenously through the interactions between a parent and their child.

There have been growing interests in whether early childhood intervention is more or less effective than intervention in later life (Heckman and Krueger 2003). Heckman and his colleagues tried to reconcile a large body of evidence using the concept of "technology of skill formation" with multidimensional human capital and dynamic complementarity of different skills (Cunha and Heckman 2007). Through Heckman's research, it became increasingly recognized that early childhood investment was important in the development of non-cognitive abilities for later development of cognitive skills. Although the setting is different, the Akabayashi model, which highlighted the importance of a child's time preference as a crucial part of personality for more general human capital investment through self-learning, shares many aspects with the Cunha-Heckman model. To date, however, the Akabayashi model has remained a purely theoretical framework, while the Cunha-Heckman model has been estimated using the CNLSY (Cunha and Heckman 2008; Cunha et al. 2010).

Along with Heckman, there is a growing number of economists who have created models and estimated the development of children (Bernal and Keane 2010, 2011; Del Boca et al. 2014; Todd and Wolpin 2003, 2007). Most of the models, however, do not explicitly consider the role of a child's own choices on the parental decision or the child's human capital accumulation. It is only recently that a few papers have attempted to model the dynamics of the parent-child interaction and the parent's learning about their child's characteristics and the effectiveness of their investment in their child (Beenstock 2012; Heckman and Mosso 2014).

Weinberg (2001) is an early prominent example. His main focus is the substitution between psychological incentives and pecuniary incentives under the credit constraint. Since wealthy parents can use monetary rewards more easily than poor families, the Weinberg model predicts that low-income parents tend to use physical punishment more frequently than high income parents, and give incentives for human capital development more efficiently. The Weinberg model fills a missing part in the Akabayashi model, the effects of parental income on the choice of parental incentives.

Lizzeri and Siniscalchi (2008) built a model of the parent-child interaction where the parent could help the child to perform better. In their model, the parent can observe the child's actions, helping them to achieve both their current happiness and future learning. The two objectives conflict since the child cannot directly observe the parent's assistance, and excessive parental help would reduce the amount of learning by the child. The paper showed that the optimal parenting policy was "partial-sheltering," where the parent intervened in the child's development more if the parent felt the child was acting differently from them. The parental "correction" serves to make the parent and child behavior more similar.

In Hao et al. (2008), siblings play a repeated game with the parent so younger siblings can learn about the preferences of parents from their sibling's behaviors and the consequences toward their elder siblings (e.g. parental punishment). Using the reputation model in game theory, the paper showed that parents have incentives to penalize older children for their risky behavior in order to discourage their younger siblings from engaging in similar behavior. A related paper (Hotz and Pantano 2013) attempts to theoretically explain the existence of the "birth order effect" that is frequently found in empirical literature.

While several other researchers have worked on similar issues in recent years, the Akabayashi model has remained unique in fully allowing the dynamic development of the discount factor and human capital, with information uncertainty for the parent.

There is an increasing interest in the effectiveness of early childhood education in both academia and the government, and research attempting to explain the mechanism and find effective methods for child-rearing and investments has flourished in economics, psychology, and brain science. I believe that a growing body of evidence and a better understanding of the parent-child interaction and a child's human capital development will allow us to build better education and social policies in the future.

References

Akabayashi H (1996) On the role of incentives in the formation of human capital in the family. Dissertation, University of Chicago
Akabayashi H (2006) An equilibrium model of child maltreatment. J Econ Dyn Control 30(6): 993–1025
Anderson BDO, Moore JB (1979) Optimal filtering. Prentice-Hall, Englewood Cliffs
Appelbaum AS (1977) Developmental retardation in infants as a concomitant of physical child abuse. J Abnorm Child Psychol 5(4):417–423
Basawa IV, Scott DJ (1983) Asymptotic optimal inference for non-ergodic models. Springer, New York
Baumrind D (1967) Child care practices anteceding three patterns of preschool behavior. Genet Psychol Monogr 75(1):43–88
Bhatt V, Ogaki M (2012) Tough love and intergenerational altruism. Int Econ Rev 53(3):791–814
Becker GS (1967) Human capital and the personal distribution of income: an analytical approach. Woytinski lecture. University of Michigan, Ann Arbor
Becker GS, Mulligan CB (1997) The endogenous determination of time preference. Q J Econ 112(3):729–758
Becker GS, Tomes N (1979) An equilibrium theory of the distribution of income and intergenerational mobility. J Polit Econ 87(6):1153–1189
Becker GS, Tomes N (1986) Human capital and the rise and fall of families. J Labor Econ 4(3):S1–S39
Beenstock M (2012) Heredity, family, and inequality. MIT, Cambridge
Belsky J (1980) Child maltreatment. Am Psychol 35(4):320–335
Bergstrom T (1989) A fresh look at the rotten kid theorem–and other household mysteries. J Polit Econ 97(5):1138–1159
Bergstrom TC (1997) A survey of theories of the family. In: Rosenzweig MR, Stark O (eds) Handbook of population and family economics, vol 1. Elsevier Science, Amsterdam, pp 21–79
Berk LE (1991) Child development, 2nd edn. Allyn and Bacon, Newton
Bernal R, Keane MP (2010) Quasi-structural estimation of a model of childcare choices and child cognitive ability production. J Econom 156(1):164–189
Bernal R, Keane MP (2011) Child care choices and children's cognitive achievement: the case of single mothers. J Labor Econ 29(3):459–512
Bernheim BD, Shleifer A, Summers LH (1995) Strategic bequest motives. J Polit Econ 93(6):1045–1076
Blau DM (2001) The child care problem: an economic analysis. Russell Sage Foundation, New York
Cawley J, Conneely K, Heckman JJ, Vytlacil E (1997) Cognitive ability, wages and meritocracy. In: Devlin B, Fienberg S, Resnick D, Roeder K (eds) Intelligence, genes, and success: scientists respond to the bell curve. Springer, New York, pp 179–192
Chua A (2011) Battle hymn of the tiger mother. Penguin Books, New York
Cunha F, Heckman JJ (2007) The technology of skill formation. Am Econ Rev 97(2):31–47
Cunha F, Heckman, JJ (2008) Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. J Hum Resour 43(4):738–782
Cunha F, Heckman JJ, Schennach SM (2010) Estimating the technology of cognitive and noncognitive skill formation. Econometrica 78(3):883–931
Del Boca D, Flinn C, Wiswall M (2014) Household choices and child development. Rev Econ Stud 81(1):137–185
Domowitz I, Muus L (1988) Likelihood inference in the nonlinear regression model with explosive linear dynamics. In: Barnett WA, Berndt ER, White H (eds) Dynamic econometric modeling: proceedings of the third international symposium in economic theory and econometrics. Cambridge University Press, Cambridge, pp 53–72

Erickson MF, Egeland B, Pianta R (1989) The effects of maltreatment on the development of young children. In: Cicchetti D, Carlson V (eds) Child maltreatment: theory and research on the causes and consequences of child abuse and neglect. Cambridge University Press, Cambridge, pp 647–684

Fitzgerald RJ (1971) Divergence of the Kalman filter. IEEE Trans Autom Control 16(6):736–747

Garcia F, Gracia E (2009) Is always authoritative the optimum parenting style? Evidence from Spanish families. Adolescence 44(173):101–131

Gibbons R, Murphy KJ (1992) Optimal incentive contracts in the presence of career concerns: theory and evidence. J Polit Econ 100(3):468–505

Hamilton JD (1994) Time series analysis. Princeton University Press, Princeton

Hao L, Hotz VJ, Jin GZ (2008) Games parents and adolescents play: risky behavior, parental reputation and strategic transfers. Econ J 118(528):515–555

Heckman JJ (1995) Lessons from the bell curve. J Polit Econ 103(5):1091–1120

Heckman JJ, Krueger AB (2003) Inequality in America: what role for human capital policies? MIT, Cambridge

Heckman JJ, Mosso S (2014) The economics of human development and social mobility. NBER working paper No. 19925

Herrnstein RJ, Murray CA (1994) The bell curve: intelligence and class structure in American life. Free Press, New York

Holmström B, Milgrom P (1987) Aggregation and linearity in the provision of intertemporal incentives. Econometrica 55(2):303–328

Hotz VJ, Pantano J (2013) Strategic parenting, birth order and school performance. NBER working paper No. 19542

Leibowitz A (1974) Home investments in children. J Polit Econ 82(2):S111–S131

Lizzeri A, Siniscalchi M (2008) Parental guidance and supervised learning. Q J Econ 123(3):1161–1195

Maccoby EE, Martin JA (1983) Socialization in the context of the family: parent-child interaction. In: Mussen PH (ed) Handbook of child psychology, vol 4. Wiley, New York, pp 1–101

Paxson C, Waldfogel J (2002) Work, welfare, and child maltreatment. J Labor Econ 20(3):435–474

Sedlak AJ, Broadhurst DD (1996) Third national incidence study of child abuse and neglect: final report. U.S. Department of Health and Human Services, Washington, DC

Steinberg L, Dornbusch SM, Brown BB (1992) Ethnic differences in adolescent achievement: an ecological perspective. Am Psychol 47(6):723–729

Todd PE, Wolpin KI (2003) On the specification and estimation of the production function for cognitive achievement. Econ J 113(485):F3–F33

Todd PE, Wolpin KI (2007) The production of cognitive achievement in children: home, school, and racial test score gaps. J Hum Cap 1(1):91–136

Weinberg BA (2001) An incentive model of the effect of parental income on children. J Polit Econ 109(2):266–280

Wieland V (2000) Learning by doing and the value of optimal experimentation. J Econ Dyn Control 24(4):501–534

Wolfe DA (1987) Child abuse: implications for child development and psychopathology. Sage Publications, Newbury Park

Wolfe DA (1991) Preventing physical and emotional abuse of children. Guilford Press, New York

Zigler E, Hall NW (1989) Physical child abuse in America: past, present, and future. In: Cicchetti D, Carlson V (eds) Child maltreatment: theory and research on the causes and consequences of child abuse and neglect. Cambridge University Press, Cambridge, pp 38–75

Chapter 2
Tough Love and Intergenerational Altruism

Vipul Bhatt and Masao Ogaki

Abstract This chapter develops and studies a tough love model of intergenerational altruism. We model tough love by modifying the Barro-Becker standard altruism model in two ways. First, the child's discount factor is endogenously determined, so lower childhood consumption leads to a higher discount factor later in life. Second, the parent evaluates the child's lifetime utility with a constant high discount factor. Our model predicts that parental transfers will fall when the child's discount factor falls. This is in contrast with the standard altruism model, which predicts that parental transfers are independent of exogenous changes in the child's discount factor.

Keywords Tough love • Endogenous time discounting • Parenting

1 Introduction

How different generations are connected to each other is an important economic issue with implications for individual economic behavior such as savings, investment in human capital, and bequests, which in turn affect aggregate savings and growth. These interactions are also important from a policy perspective since they determine how families respond to public policies aimed at redistributing resources among family members. A commonly used paradigm to study such linkages is the

---

The original article first appeared in the International Economic Review 53:791–814, 2012. A newly written addendum has been added to this book chapter.

V. Bhatt
Department of Economics, James Madison University, 421 Bluestone Drive, ZSH 433, Harrisonburg, VA 22807, USA
e-mail: bhattvx@jmu.edu

M. Ogaki (✉)
Department of Economics, Keio University, 612 Mita-Kenkyu-shitsu, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
e-mail: mogaki@econ.keio.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_2

standard altruism model proposed by Becker (1974) and Barro (1974) in which the current generation derives utility from its own consumption and the utility level attainable by its descendant(s). Using this framework, Barro found that there will be no net wealth effect of a change in government debt.

A striking prediction of the standard altruism model is that when the child becomes impatient, transfers from the parent to the child do not change when the child is borrowing constrained (as we show in Sect. 3). This implication of the model is not consistent with recent empirical evidence on pecuniary and non-pecuniary parental punishments (see Bhatt 2011; Hao et al. 2008; Weinberg 2001 for empirical evidence). For example, imagine that a child befriends a group of impatient children and suddenly becomes impatient because of their influence. As a result, the child starts to spend more time playing with the new friends and less time studying. In the worst cases, the child starts to smoke, drink, or consume illegal drugs (see Ida and Goto (2009) for empirical evidence that shows the association of low discount factors and smoking). At least some parents are likely to respond by imposing pecuniary punishments, such as reducing allowances, or non-pecuniary punishments, such as grounding. Another feature of the standard altruism model is that it precludes parents from directly influencing their children's time preferences. However, there is empirical evidence that parents attempt to shape their children's economic behavior and attitudes, including time preferences, as reviewed below. In many recent theoretical contributions, preferences of children are not exogenous, but are shaped by the attitudes and actions of their parents and other role models. For example, in the literature on cultural transmission of preferences, Bisin and Verdier (2001) proposed a general model with endogenous cultural transmission mechanisms wherein parents take actions to affect children's traits, which as a special case can correspond to time preferences. In some other models, even though parents do not take actions with the deliberate intention of affecting their children's preferences, they end up doing so indirectly. For example, Fernandez et al. (2004) used a dynamic model where mothers who work play an important role in the transmission of attitudes favoring the participation of women in the labor force to their sons. We will further discuss this issue by presenting empirical evidence for the parents' role in children's endogenous preference formation in the next section.

The main contribution of this chapter is to propose a new theoretical model of parent–child interaction that incorporates a mechanism through which parents can affect their children's time preference formation. We develop a tough love model of intergenerational altruism, in which the parent is purely altruistic to the child, but exhibit tough love: he allows the child to suffer in the short run with the intent of helping her in the long run. The main prediction of our tough love altruism model is that transfers from the parent will fall when the child's discount factor falls exogenously. This is in sharp contrast to the prediction of the standard altruism model where the parent does not respond to such a change in the child's discount factor. An interpretation of this result is that parents with the tough love motive use pecuniary incentives to mold their children's time preferences. Since exogenous changes in the child's discount factor that make her impatient are likely to cause behavior that calls for parents' corrective actions, the prediction of the tough love altruism model is more consistent with the empirical evidence on parental punishments, as well as with the role of parents in shaping children's preferences as compared with that of the standard altruism model.

In the simple setting of a three-period economy with a single parent and single child with perfect information and borrowing constraints, we model parental tough love by combining two ideas that have been studied in the literature in various contexts. First, the child's discount factor is endogenously determined, so low consumption at a young age leads to a higher discount factor later in life. This is based on the endogenous discount factor models of Uzawa (1968), except that the change in the discount factor is immediate in Uzawa's formulation, whereas a spoiled child with high consumption progressively grows impatient in our formulation.[1] Second, the parent evaluates the child's lifetime utility with a constant discount factor that is higher than that of the child. Since the parent is the social planner in our simple model, this feature is related to recent models in which the discount factor of the social planner is higher than that of the agents.[2] In our model, these two features lead the parent to exhibit tough love behavior in which he takes into account the influence of income transfers to the child based on the latter's discount factor.

An argument for the plausibility of endogenous discounting can be found in Becker and Mulligan (1997). They model an individual whose discount factor depends on the remoteness or vividness of imagined future pleasures. Becker and Mulligan's model involves investment in human capital to increase the vividness of imagination. For the direction of the effect of wealth on the discount factor, this argument can be used to support both Fisher's conjecture that poor people are less patient (see Fisher 1930, p. 72 for details) and Uzawa's (1968) hypothesis that poor people are more patient. Because richer people tend to invest more, their model typically implies that poorer people are less patient. On the other hand, if a child experiences low consumption, it should be easier for the child to imagine future misery more vividly. This argument implies that a child who experiences low consumption will tend to grow more patient. The child may experience low consumption either because the parent is poor or because the parent is concerned about spoiling the child. In our review of empirical evidence in the next section, we find mixed evidence for both directions, which seems to imply that both of these forces are working in practice. For the purpose of our paper, we abstract from the human capital aspect and adopt the formulation that a child who is spoiled by high consumption in childhood grows to be less patient.[3]

[1]Recent theoretical models that adopt the Uzawa-type formulation include Schmitt-Grohé and Uribe (2003) and Choi et al. (2008).

[2]See Caplin and Leahy (2004), Farhi and Werning (2007), Phelan (2006), and Sleet and Sevin (2005, 2007).

[3]This chapter focuses on the parent's role in preference formation. A related work is Mulligan (1998) on the altruistic preference formation of the parent toward the child.

Turning to the plausibility of the parent using a higher discount factor than the child, an extreme case is a parent with a newborn baby. When the baby is born, it is very impatient and cries for food all the time but the parent does not give in to this persistent demand. This is likely because the parent evaluates the baby's utility over its lifetime with a higher discount factor as compared with the baby's very low discount factor. We think that it is likely that many parents continue to evaluate their children's lifetime utility when they are no longer babies. Mischel's (1961) results, which we mention in the next section, are consistent with our view. Parents may continue to do this until their children learn to be as patient as them.

As a model of parental punishments, our model is related to Weinberg's (2001) model.[4] He develops a static incentive model based on asymmetric information, whereas our model is dynamic without any uncertainty. The parent in Weinberg's model does not affect the child's preferences, whereas the parent in our model takes actions with explicit intent to affect the child's discount factor. In this chapter, we are emphasizing the role of the parent in molding the time preference of the child. In this regard, our model is closely related to those of Akabayashi (2006) and Doepke and Zilibotti (2008). In these models also, the parent takes actions in order to influence the child's discount factor. In Akabayashi's model, the child has endogenous discounting, and the parent evaluates the child's lifetime utility with a fixed discount factor. Together with asymmetric information about the child's ability, Akabayashi's model can explain abusive repeated punishments by parents under certain parameter configurations. In Doepke and Zilibotti's model, the parent uses the child's discount factor to evaluate the child's lifetime utility. They use their model of occupational choice to account for a number of observations about the British Industrial Revolution. The main difference from our model is that these authors adopt a Becker-Mulligan formulation of endogenous discounting so that children become more patient when their human capital is higher. In contrast, we adopt an Uzawa-type formulation for our model.

The remainder of the paper is organized as follows. Section 2 reviews the empirical evidence related to the key assumptions and implications of the tough love altruism model. Section 3 explains the structure and main findings of our model with only a consumption good, and contrasts the implications of the model with those of the standard altruism model. Section 4 proposes two alternative models of altruism in order to show that both features discussed above (the endogenous discount factor of the child and the parent's evaluation of the child's lifetime utility with a high constant discount factor) are necessary in order for transfers to decrease when a child exogenously becomes impatient. Section 5 introduces leisure in the tough love

[4]In a recent work, Slavík and Wiseman (2009, Tough love for lazy kids: dynamic insurance and equal bequests, unpublished manuscript) have also proposed a model of tough love. These authors emphasize the moral hazard problem faced by parents in order to construct a model with a dynamic insurance strategy that involves providing greater inter vivos transfers to poor children and dividing bequests equally. Their model does not involve endogenous time discounting.

altruism model with the objective of studying how parental transfers are affected by endogenous changes in the child's income caused by (exogenous) changes in her discount factor. Section 6 concludes.

2 A Review of Empirical Evidence

In this section, we review the empirical evidence related to the key assumptions and implications of the tough love altruism model.

Endogenous discounting is an important assumption of our tough love model and we first review the existing empirical evidence for this assumption. In the literature, there are two competing hypotheses that allow for endogenous discount factor by linking patience to wealth. First is Fisher's hypothesis that the rich are more likely to be patient, and second is the Uzawa's hypothesis that implies the discount factor is decreasing in wealth.

Becker and Mulligan (1997) cite empirical evidence for endogenous discounting consistent with the Fisher hypothesis. Similarly, using the Panel Study of Income Dynamics (PSID), Lawrance (1991) employed the Euler equation approach to estimate the endogenous discount factor model and found evidence in favor of the discount factor increasing in wealth.

However, one has to be careful in evaluating the empirical evidence for endogenous discounting because of two problems. First, we have the endogeneity problem in that patient people with high discount factors tend to accumulate financial and human wealth. Thus, we may find that rich people have higher discount factors than poor people even when the discount factor of an individual is decreasing in wealth as in Uzawa's model.[5] The endogeneity problem mentioned above is addressed in Ikeda et al. (2005). In their paper, they found that without accounting for the possible endogeneity between discount factors and wealth, the discount factor appears to be an increasing function of income/wealth. After taking into account the endogeneity problem, they find evidence in favor of the discount factor decreasing in wealth.[6] Another way to control for the endogeneity problem is to give different levels of consumption to the subjects before an experiment to see which subjects are more patient. Implementing this idea with human subjects is

[5]This issue is related to the literature on the importance of initial endowments on subsequent outcomes of a dynamic process (Heckman 1981, 1991). As suggested by Heckman, it is important to distinguish between heterogeneity (how persistent is the effect of initial endowments on outcomes), and state dependence (whether subsequent experiences attenuate or accentuate the effect of initial endowments). It is possible that a raw correlation between wealth and consumption growth reflects a causal influence of wealth on consumption growth (state dependence), or the fact that individuals differ in time preferences and more patient people accumulate more wealth (heterogeneity).

[6]They control the endogeneity problem by analyzing how the discount factor changes with the size of a prize obtained in another experiment.

difficult, so rats were used instead. The results were in favor of the view that the discount factor is decreasing in wealth as reported in Kagel et al. (1995, Ch. 7, Section 3).

The second problem in evaluating empirical evidence for endogenous discounting is that endogenous discounting and wealth-varying intertemporal elasticity of substitution (IES) can have similar implications in growing economies, and may be hard to distinguish from one another (Atkeson and Ogaki 1996). Hence, although the Lawrance (1991) estimation method based on the instrumental variable approach could potentially resolve the endogeneity problem, she did not allow the IES to vary with wealth. Ogaki and Atkeson (1997) allow both the IES and the discount factor to vary with wealth for a panel data of households in Indian villages. They find evidence in favor of the view that the discount factor is constant and that the IES is increasing in wealth. It is possible that the discount factor is decreasing in wealth for richer households, but Lawrence found the opposite result by not allowing the IES to change. Ogawa (1993) argues that empirical results from Japanese aggregate data are consistent with a combination of Fisher's and Uzawa's hypotheses.

Overall, we think that the empirical evidence is consistent with the view that reality is best described by a combination of the two hypotheses. In our view, a child who experiences low consumption will grow to be more patient because he/she can more vividly imagine future misery. At the same time, a wealthier parent is more likely to invest in the child's human capital to help the child see the future more vividly. In this chapter, we aim to develop a simple model that captures our intuition of tough love, which is that a parent allows suffering so that the child can learn to be more patient. Such a model will imply that transfers decrease when the child exogenously becomes impatient. For this purpose, we will assume that low childhood consumption leads to more patience (higher discount factor) in adulthood and will abstract from the human capital nature of endogenous discounting.

The tough love model presented in this chapter hypothesizes a strong parental role in shaping child behavior and preferences. The main prediction of the model is that parents with tough love motives will provide lower childhood consumption to their children in order to influence their discount factor. Ideally, we would like to present evidence for such a parent-child interaction in data. However, to our knowledge, there is no existing study that seeks to answer this question directly. This is partially a consequence of lack of data (survey or experimental) on parental motives, childhood consumption, and discount factor of parents and their children. As result, we attempt to approach this issue indirectly by reviewing empirical evidence on three related questions.

Our first question is whether or not there is empirical evidence for parents' behavior influencing their children's discount factors as well as other economic preferences and attitudes. A necessary condition for parents' behavior to be able to affect children's time discounting factors is that genetic factors do not completely determine time discounting. Using a unique data set of twins in Japan, Hirata et al. (2010) found empirical evidence in favor of this condition. Knowles and Postlewaite (2005, Wealth inequality and parental transmission of savings behavior, unpublished manuscript) used data from the PSID to examine the relationship between parental attitudes toward planning for the future and their children's saving rates. They found that for the oldest children, the parents' attitudes explain one-third of the variance in savings rates that remains after controlling for income and demographics. Similarly, Webley and Nyhus (2006) used De Nederlandsche Bank household survey (DHS) data and found evidence to support the hypothesis that parental orientations, especially those related to intertemporal choice, affect the economic behavior of their children in both childhood and adulthood. In Webley and Nyhus' analysis, they observed high degrees of association between children's savings and parental savings, household income and economic socialization of parents. In the psychology literature, there is evidence in favor of the influence of parents in the development of children's willingness to delay rewards. Mischel (1961) studied children in the West Indian islands of Grenada and Trinidad. In both cultures, he found a significant relationship between absence of the father within a household and greater preference for immediate reward reflecting impatient child behavior. Such an association, among other things, suggests a strong role for the father in handing down values of thrift to his child.

Our second question is whether or not there is direct empirical evidence that some parents take actions with the intention of affecting their children's behavior. This issue of the relationship between various parenting styles, identified by varying degrees of control, has been addressed more directly in the psychology literature than in the economics literature.[7] For instance, Carlson and Grossbart (1988) used survey data on the mothers of schoolchildren (kindergarten through sixth grade) and divided them into groups based on the parenting style, ranging from neglecting to rigidly controlling. They found evidence suggesting that authoritative parents grant less consumption autonomy to their children, have greater communication with their children about consumption-related issues, set higher consumer socialization goals and exhibit greater monitoring of children's consumption vis-Evis both permissive and authoritarian parents. Such a relationship holds in the data even after accounting for possible cultural differences. For example, Rose et al. (2003) used data for India, Australia and Greece, and found evidence suggesting that authoritative parents more closely monitor their children's consumption compared with other parenting styles.

Our third question is whether or not there is empirical evidence related to the main implication of our model that a parent reduces his child's childhood consumption when the discount factor that the parent uses to evaluate the child's life

[7] Baumrind (1966) identified three modes of parental control. The first mode is *permissive*, where parents act as a resource for their children and do not actively involve themselves in shaping the current and future behavior of the child. The second mode is *authoritarian*, where the parent uses a set standard of conduct that is theologically or religiously motivated and tries to shape and control the child's behavior with overt use of power. The third mode is *authoritative*, where the parent actively involves himself/herself in shaping the child's behavior and attitudes and uses reasoning and discipline to ensure a well-rounded long-term development of the child. The parent affirms the child's current behavior, separating right from wrong, and also sets standards for the child's future behavior.

time utility is higher than the child's discount factor. The data limitation is obviously a difficulty to find such evidence, but Kubota et al. (2011) provided empirical evidence related to this. They used a unique U.S. and Japanese survey data that contained hypothetical survey questions concerning parents' tough love attitudes and their time discount factors for their own financial decisions. Their empirical results suggest that parents with lower time discount rates (higher discount factors) are more likely to have tough love attitudes to reduce consumption of a medicine when the medicine has a side effect to weaken the child's immune system after the child grows up.

3 A Consumption Good Economy

The main purpose of this section is to develop and analyze a model of altruism in which the parent's transfers decrease when the child exogenously becomes impatient. For this purpose, we modify the standard altruism model in two ways: the child's discount factor is endogenous in that higher consumption in childhood causes her discount factor to be lower, and the parent evaluates the child's lifetime utility with a high constant discount factor. The modified model is called the tough love altruism model. In order to gain a clear understanding of the properties of the model, we consider the simplest setting and compare the tough love model with the standard altruism model in this section.

Imagine a three-period model economy with two agents, the parent and the child. For simplicity, we consider the case of a single parent and a single child. The three periods considered are childhood, work and retirement.[8] The model has seven features. First, the timing of the model is assumed to be such that the life of the parent and the child overlaps in the first two periods of the child's life. Hence, the parent has the child in the second period of his own life, which in turn corresponds to the first period of the child's life. Second, the parent not only cares about his own consumption, but is also altruistic toward the child. He assigns a weight of η to his own utility, where $0 < \eta < 1$. Third, the parent receives an exogenous income, denoted by y_p, in period 2 of his life. For simplicity, we assume that there is no bequest motive and also that the parent receives no income in the last period of his life but simply consume savings from the previous period. Fourth, the parent maximizes utility over the last two periods of life by choosing consumption and transfers to his child, denoted by C_2^p and T, respectively, in period 2 of life and

[8]For expositional ease, we begin by making the simplifying assumption that these three periods are of equal duration. Note that results presented in this section as well as in Sect. 4 are robust to varying durations for the three periods. Further, in Sect. 5 we relax this assumption and study the model with varying durations for childhood, work, and retirement.

consuming savings in the last period of life.[9] Fifth, the child is assumed to be a nonaltruist and derives utility only from her own consumption stream $\{C_t\}_{t=1}^{3}$.[10] We assume that the child's income in periods 1 and 2, denoted by y_1 and y_2, respectively, is given exogenously and she receives no income in the last period of life. Sixth, the child is assumed to be borrowing constrained in period 1. Lastly, there is no uncertainty in the economy.

3.1 Standard Altruism Model

We start our analysis with the standard altruism model. In this model, both the parent and the child use the same constant discount factor when evaluating the child's future utility. The parent's problem is:

$$\max_{C_2^p, T} \Big\{ \eta \Big[v(C_2^p) + \tilde{\beta} v(R(y_p - C_2^p - T)) \Big] + \tilde{\beta}(1-\eta) \Big[u(C_1^*) + \beta_2 u(C_2^*) + \beta_2 \beta_3 u(R^2(y_1 + T + \frac{y_2}{R} - C_1^* - \frac{C_2^*}{R})) \Big] \Big\}, \tag{2.1}$$

subject to:

$$C_1 = y_1 + T, \tag{2.2}$$

and:

$$\{C_1^*, C_2^*\} \equiv \arg\max_{C_1, C_2} \Big[u(C_1) + \beta_2 u(C_2) + \beta_2 \beta_3 u(R^2(y_1 + T + \frac{y_2}{R} - C_1 - \frac{C_2}{R})) \Big], \tag{2.3}$$

where $v(.)$ and $u(.)$ are standard concave period utility functions of the parent and the child, respectively. $\tilde{\beta}$ is the parent's own discount factor whereas β_t is the period t discount factor used to evaluate the child's future utility. R is the gross nominal interest rate.

We can simplify the parent's problem by making two modifications. First, we are interested in the case where the borrowing constraint is binding for the child and assume that the parameters are such that the constraint is binding. We substitute

[9]Given the timing of our model, this implies that transfers, T, are made only in period 1. Further, we assume that transfers are made from the parent to the child and there are no reverse transfers.

[10]In this simple consumption good economy, we view consumption as a composite good that may include leisure activities such as TV time, video game time etc. In Sect. 5, we extend this basic setup and introduce leisure as a second good.

out the borrowing constraint faced by the child in period 1 in the parent's problem described above. Second, we can reduce the dimensionality of the maximization problem by solving for optimal C_2^p as a function of transfers T and other model parameters and then substituting it out from the parent's maximization problem. We denote the resulting indirect utility function of the parent by $V(R(y_p - T), \tilde{\beta})$. After incorporating these modifications, we can rewrite the parent's optimization problem as:

$$\max_{T} \left\{ \eta V(R(y_p - T), \tilde{\beta}) + \tilde{\beta}(1-\eta)\Big[u(y_1 + T) + \beta_2 u(C_2^*) \right.$$
$$\left. + \beta_2\beta_3 u(R(y_2 - C_2^*))\Big] \right\}, \tag{2.4}$$

subject to:

$$\{C_2^*\} \equiv \arg\max_{C_2} \Big[u(C_2) + \beta_3 u(R(y_2 - C_2)) \Big]. \tag{2.5}$$

Let us focus on the child's optimization program. From the first-order condition for the child's problem described in Eq. (2.5), we obtain:

$$u_{C_2}(C_2) - \beta_3 R u_{C_2}(R(y_2 - C_2)) = 0, \tag{2.6}$$

where:

$$u_x(x) \equiv \frac{\partial u(x)}{\partial x}.$$

Assuming that the utility function satisfies conditions for the Implicit Function Theorem,[11] we can solve Eq. (2.6) for C_2 as a function of the model parameters and the state variables:

$$C_2^* = C_2(y_2, \beta_3, R). \tag{2.7}$$

The optimal period 2 consumption for the child is independent of the period 1 transfers of the parent and hence can be dropped from the parent's optimization program. Hence, we can rewrite the parent's problem described by Eqs. (2.4) and (2.5) as:

$$\max_{T} \Big[\eta V(R(y_p - T), \tilde{\beta}) + \tilde{\beta}(1-\eta) u(y_1 + T) \Big]. \tag{2.8}$$

---

[11] u(.) is continuously differentiable with a nonzero *Jacobian*.

From the first-order condition for the above problem and the implicit function theorem, we obtain:

$$T^* = T(y_p, y_1, \tilde{\beta}, R, \eta). \tag{2.9}$$

We now consider a comparative statics exercise for the standard altruism model wherein we decrease the child's discount factor β_3 and observe how this rise in the child's impatience is accommodated by the parent in terms of a change in period 1 transfers. From Eq. (2.9), optimal period 1 transfers by the parent in the standard altruism model are in fact independent of the child's discount factor. Hence, such an exogenous change in the child's discount factor will have no effect on the period 1 transfers made by the parent.[12] As discussed earlier, this implication of the model does not seem to be consistent with data where we find that both pecuniary and non-pecuniary punishments are used by parents to influence their children's behavior and outcomes.

3.2 Tough Love Altruism Model

We propose a tough love altruism model that provides for a channel through which parents can influence their child's economic behavior.[13] We introduce the tough love motive of the parent via asymmetric time preferences between generations and endogenous discounting. In this model, the parent uses a constant and high discount factor, denoted by $\beta_{t,p}$, to evaluate the child's lifetime utility. The child herself uses a discount factor that is endogenously determined as a decreasing function of period 1 consumption:

$$\beta_{t,k}(C_1) \quad ; \quad \frac{\partial \beta_{t,k}}{\partial C_1} < 0.$$

With the borrowing constraint faced by the child in period 1, her period t discount factor is given by $\beta_{t,k}(y_1 + T)$.

In this model, the parent solves the following optimization problem:

[12]Note that changes in the parent's own discount factor will affect transfers. However, here we are imagining a sudden change in social norms that affects only the child's discount factor with no effect on the parent's discount factor.

[13]The discussion presented here postulates a model of parental tough love in the context of a single generation. An interesting extension is to model tough love in a dynastic framework exemplified in the context of the standard altruism model by Barro (1974). In such a model, the child will be a repeater in the dynamic process and will pass on the discount factor she inherited from her tough love parent to her own offspring. We are investigating the implications of such a framework for tough love altruism in a separate paper that is a work in progress.

$$\max_T \left\{ \eta \, V(R(y_p - T), \tilde{\beta}) + \tilde{\beta}(1-\eta)\Big[u(y_1 + T) + \beta_{2,p} u(C_2^*) + \beta_{2,p}\beta_{3,p} u(R(y_2 - C_2^*)) \Big] \right\}, \tag{2.10}$$

subject to:

$$\{C_2^*\} \equiv \arg\max_{C_2} \Big[u(C_2) + \beta_{3,k}(y_1 + T) u(R(y_2 - C_2)) \Big]. \tag{2.11}$$

From the first-order condition for the child's problem described in Eq. (2.11) and the implicit function theorem, in principle we can solve for the optimal C_2 as a function of the model parameters and the state variables:[14]

$$C_2^* = C_2(y_2, \beta_{3,k}(y_1 + T), R). \tag{2.12}$$

3.2.1 Relationship Between Transfers and the Child's Discount Factor

One of the distinguishing predictions of our tough love altruism model concerns the relationship between the optimal parental transfers and the child's discount factor. In what follows, we provide an analytical result that formalizes this relationship in our model. For this purpose, we assume the following specification for the child's discount factor:

$$\beta_{t,k}(y_1 + T) = \beta_0 + \psi(y_1 + T) \quad ; \quad \psi'(y_1 + T) < 0.$$

In this specification, β_0 is introduced for the purpose of performing comparative statics for exogenous changes in the child's discount factor, which do not change the sensitivity of the discount factor to changes in period 1 consumption. We are interested in establishing a relationship between optimal parental transfers, T^* and β_0 for the parent with the tough love motive. Our intuition is that the parent with a tough love motive will reduce transfers in response to an exogenous decrease in the child's discount factor, i.e. $\frac{\partial T^*}{\partial \beta_0} > 0$. Using the optimization conditions for the parent and child problems, we next derive the expression for $sign\left(\frac{\partial T^*}{\partial \beta_0}\right)$.

[14]It can be easily shown that

$$\frac{\partial C_2^*}{\partial \beta_0} < 0 \quad \frac{\partial C_2^*}{\partial T} > 0$$

Proposition 1 *A necessary and sufficient condition for* $\dfrac{\partial T^*}{\partial \beta_0} > 0$ *is*

$$\left[1 + R(\beta_p - \beta_{3,k})\frac{u''(R(y_2 - C_2^*))}{u'(R(y_2 - C_2^*))}\frac{\partial C_2^*}{\partial \beta_0} - (\beta_p - \beta_{3,k})\frac{\partial^2 C_2^*}{\partial \beta_0 \partial T}\right] > 0. \quad (2.13)$$

Proof See section "A Proof for Proposition 1" in the appendix for a proof. □

This is the main result of this chapter. In order to facilitate intuitive interpretation of the condition in this proposition, we give the next proposition. It shows that the condition specified in Eq. (2.13) is related to the convexity of the marginal utility and the child's impatience (both the absolute level and the relative level compared with the parent's discount factor used to evaluate the child's lifetime utility).

Proposition 2 *The following three conditions are jointly sufficient for the condition in Proposition 1 to hold*

(i) $u'''(.) \geq 0$,
(ii) $\beta_p \geq \beta_{3,k}$, and
(iii) $\beta_{3,k} RG \leq 1$,

where

$$G = \left(\frac{u'''(C_3^*)}{u''(C_3^*)}\right)\frac{\partial C_2^*}{\partial \beta_0}.$$

Proof See section "A Proof for Proposition 2" in the appendix for a proof. □

The first condition in Proposition 2 implies convexity of the marginal utility function. This condition is satisfied by many functional forms that are used for the utility function in the consumption literature: it is satisfied with the strict inequality for the power utility function, and with equality for the quadratic utility function. The second condition implies that the child is relatively inpatient compared with the parent's norm in the sense that the child's discount factor is less than or equal to the discount factor used by the parent to evaluate the child's lifetime utility.[15] This condition is consistent with the first assumption of our tough love altruism model that the parent uses a high constant discount factor to evaluate the child's lifetime utility. The third condition is trivially satisfied if $u'''(C_3^*) = 0$. If $u'''(C_3^*) > 0$, the condition requires a certain level of impatience of the child, where the level is not

[15]Note that since $\frac{\partial T^*}{\partial \beta_0}$ is strictly positive for $\beta_p = \beta_{3,k}$, we have a positive relationship between T^* and β_0 even when $\beta_p < \beta_{3,k}$, as long as the difference is small in magnitude.

directly affected by the parent's norm. For example, if we assume a power utility function, then the third condition can be expressed as:

$$\beta_{3,k} R \leq \left(\frac{\sigma}{R}\right)^{\frac{1}{\sigma}}, \tag{2.14}$$

as shown in section "Power Utility Function" in the appendix. If $\sigma = R$, then this condition reduces to $\beta_{3,k} R \leq 1$. If $\sigma = 1$, then this condition reduces to $\beta_{3,k} \leq 1/R^2$, which implies $\beta_{3,k} \leq 0.69$ if $R = 1.2$. Assuming $R = 1.2$, the condition implies that β_3 is less than equal to 0.96, 1, and 1.07 when σ is equal to 1.5, 1.61, and 2, respectively.

Thus, the condition (2.14) is satisfied by all β_3 less than one if $\sigma \geq 1.61$ when $R = 1.2$. However, the condition may seem stringent for smaller values of σ. Here, it should be noted that the upper bound in the condition (2.14) is not meant to be sharp because the condition is sufficient but is not necessary. Transfers can have a positive relationship with β_0 with much larger values of $\beta_{3,k}$. To provide greater insight on this issue we numerically solve the optimization problem of the parent with tough love motive. The objective of this exercise is to numerically find the magnitude of the greatest upper bound at which a decrease in β_0 leads to higher parental transfers. For this purpose, we impose the following parameterization[16]:

$$u(C) = v(C) = \frac{C^{1-\sigma}}{1-\sigma}. \tag{2.15}$$

The discount factor is given by:

$$\beta(y_1 + T) = \beta_0 + \frac{1}{1 + a(y_1 + T)}, \quad \text{where} \quad a > 0 \text{ and } \beta_0 \leq 0. \tag{2.16}$$

We solve the problem described in Eqs. (2.10) and (2.11) numerically as a nonlinear root finding problem using the above parametric specification and a given set of parameter values.[17]

[16]Our simulation results are robust to alternative parametric specifications of the utility function and also to a wide range of model parameter values.

[17]We have chosen our parameter values to be consistent with consensus estimates reported in the literature. When such estimates are not available, we have used the optimality conditions of our model and used micro data to approximate the parameter values. For σ, we are using a value of 1.5, which implies an elasticity of intertemporal substitution of around 0.67. In the literature, many studies have used micro data and have estimated this parameter to be between 0.4 and 0.7 (see Hall 2009; Ogaki and Reinhart 1998). For deriving a value for η, given our parametric specification and under the assumption that β_p and β_k are approximately close to each other, from the parent's first-order condition we obtain:

$$\frac{C_p^{-\sigma}}{C_1^{-\sigma}} = \frac{1-\eta}{\eta}.$$

Table 2.1 Effect of an exogenous decrease in the child's discount factor

Global parameters							
$\eta = 0.66$; $R = 1.2$; $\tilde{\beta} = \beta_p = 0.99$; $y_1 = 1$; $y_2 = 10$; $y_p = 10$; $a = 0.02$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\beta_0 = 3.0$	$\beta_0 = 2.5$	$\beta_0 = 2.0$	$\beta_0 = 1.5$	$\beta_0 = 1.0$	$\beta_0 = 0.5$	$\beta_0 = 0.0$
Panel 1: $\sigma = 1.5$							
T^*	1.7345	1.7347	1.7348	1.7347	1.7344	1.7330	1.7282
$\beta_{3,k}$	3.9481	3.4481	2.9481	2.4481	1.9481	1.4482	0.9483
Panel 2: $\sigma = 0.7$							
T^*	0.7312	0.7329	0.7347	0.7365	0.7372	0.7341	0.7156
$\beta_{3,k}$	3.9665	3.4665	2.9665	2.4664	1.9664	1.4665	0.9668

We consider comparative statics for exogenous changes in the discount factor of the child as captured by a change in the parameter β_0. We first solve the model for the parametric specification given in (2.15) and (2.16) with a given set of model parameter values. This gives us the benchmark optimal transfers T^*. We then decrease β_0 and trace out the parental response in terms of transfers. The results of this exercise are summarized in Table 2.1.

As seen in the first panel of Table 2.1, when we decrease β_0 from 3.0 to 0.0, parental transfers first increase and then start to decrease with the peak around $\beta_0 = 2.0$, which corresponds to $\beta_{3,k} = 2.9481$. Hence, for the benchmark case, the greatest upper bound far exceeds $\beta_{3,k} = 1$. In order to check the robustness of this result, the second panel of Table 2.1 reports results for the case with $\sigma = 0.7$ with the other parameters unchanged from the benchmark case. Again we find that as decrease β_0 from 3.0 to 0.0, parental transfers first increase and then start to decrease with the peak around $\beta_0 = 1.0$, which corresponds to $\beta_{3,k} = 1.9664$. These numerical results indicate that the condition in Proposition 1 is satisfied by a wide range of reasonable parameter values even for low values of σ.

Proposition 1 formalizes the main prediction of our tough love altruism model that an exogenous decrease in the child's discount factor will lead to a fall in parental transfers under some regularity conditions. This prediction is in sharp contrast to that in the standard altruism model where parental transfers were independent of the child's discount factor.[18]

We used $\sigma = 1.5$ and data from the Consumer Expenditure Survey (CEX) on per capita annual consumption expenditure for individuals aged 25 or below to approximate C_1. We used per capita annual consumption expenditure for individuals aged 65 or above to approximate C_p. Then, the above optimality condition $\eta = 0.66$. Finally, for parameter a, we assumed a value of 0.02, although we also tried alternative values of 0.01 and 0.04 and found that the results are robust to these alternative values for parameter a.

[18]We have also studied the version of our model where we add the bequest motive for the parent. We find that the main result of the paper remains qualitatively unchanged. Consistent with our

3.2.2 Comparative Statics for Changes in the Child's Income and Family Income

We constructed the tough love model, so that transfers decrease when the child's discount factor exogenously falls. We now present comparative statics results for changes in the child's current and permanent income as well as for changes in family permanent income. The objective of this exercise is to challenge our tough love model by first illustrating its predictions with respect to aforesaid income changes and then comparing these predictions with the existing empirical evidence.

Effect of Changes in the Child's Current Income

One of the most important implications of the standard altruism model is the redistributive neutrality property (also called the transfer derivative restriction). The standard altruism model implies that an exogenous dollar decrease in the child's income coupled with a dollar increase in the parent's income will lead to a dollar increase in transfers from the parent to the child. Empirical evidence on the redistributive neutrality property is mostly negative and although many studies have found an inverse relationship between transfers and the recipient's current income, the magnitude is much smaller than one-for-one. For instance, Altonji et al. (1997) used Panel Survey of Income Dynamics (PSID) data and found that transfers only increase by 13 cents even when the recipient child is borrowing constrained.

Our tough love altruism model also implies redistributive neutrality. Because the parent optimizes the child's consumption level in the first period, if an exogenous factor changes the distribution of income for the parent and the child, the parent neutralizes the change by changing transfers.[19] However, this redistributive neutrality only holds for exogenous current income changes. In Sect. 5, we address this issue by allowing for leisure as a second good in the utility function. Consistent with empirical evidence, we find that our tough love altruism model predicts a less than one-for-one inverse relationship between parental transfers and the child's current income.

Effect of Changes in the Child's Permanent Income

In the literature on parent child interactions, an important issue relates to the compensatory nature of parental transfers, wherein one can argue that the lower

intuition of parental tough love, the parent with a tough love motive respond to a fall in his child's discount factor by reducing transfers and increasing bequests. In contrast, the parent in the standard altruism model increase transfers and decrease bequests in response to a fall in the child's discount factor.

[19]For brevity, we have not provided proofs for the redistributive neutrality property for all the models presented in this chapter. These analytical results are available from the authors upon request.

Table 2.2 Effect of a decrease in the child's permanent income on transfers

Global parameters		
$\eta = 0.66$; $\sigma = 1.5$; $R = 1.2$; $\tilde{\beta} = \beta_p = 0.99$; $\beta_0 = 0$; $y_1 = 1$; $a = 0.02$		
	(1)	(2)
	$y_2 = 10$; $y_p = 10$	$y_2 = 8$; $y_p = 11.67$
$T*$	1.7282	2.1419
Child's permanent income	11.0615	9.8085

is the ability of the child, the greater will be the resource transfer from parents. To the extent that the child's permanent income is a reasonable proxy for her ability, one way to address this issue is by predicting how a parent with tough love adjusts transfers in response to an exogenous change in the child's permanent income.

For conducting this experiment, we consider comparative statics for exogenous changes in the child's period 2 income (y_2). For this purpose, we first solve the model for the parametric specification given in (2.15) and (2.16) and a given set of model parameter values. This gives us the benchmark optimal transfers $T*$. We then decrease y_2 exogenously, while increasing the parent's income by the same amount. This adjustment is necessary to keep family permanent income constant. The results for a given set of model parameter values are summarized in Table 2.2.

As seen in Table 2.2, controlling for family permanent income, a decrease in a child's permanent income leads to an increase in parental transfers. Hence, our tough love altruism model predicts that the parent transfers more if the child is less able, measured by a decrease in the child's permanent income.

The empirical evidence on the relationship between the child's permanent income and transfers is mixed but more in favor of a negative relationship. For instance, Altonji et al. (1997) used PSID and found a negative relationship between transfers and the recipient's permanent income.[20] Hence, the prediction of our model is consistent with the empirical evidence and to the extent that permanent income reflects the child's ability, our model predicts greater transfers to the less able child.

[20]Using data from the Health and Retirement Survey (HRS), McGarry (1999) found a positive relationship between the recipient's permanent income and the amount of transfers. However, this study uses education of the respondent as a proxy for the permanent income, which may be missing important aspects of permanent income. Further, when estimating the relationship between the recipient's permanent income and amount of transfers, McGarry (1999) did not adjust the parent's permanent income in order to keep family permanent income constant. We believe that the Altonji et al. (1997) result is more robust as they used better measures of permanent incomes of the parent and the child and also controlled for family permanent income when evaluating the relationship between recipients' permanent income and their transfers.

3.2.3 Relationship Between Family Permanent Income and the Child's Discount Factor

In this subsection, using simulations, we illustrate the prediction of our tough love model for the effect of family income on the child's discount factor. We use comparative statics for an exogenous change in the family's permanent income. For this purpose, we first solve the model for the parametric specification given in (2.15) and (2.16) and a given set of model parameter values. This gives us the benchmark optimal transfers T^* and the child's discount factor $\beta(C_1^*)$. We then increase y_p implying a higher level for the family's permanent income. The results for this exercise are summarized for a given set of model parameter values in Table 2.3.

As seen in Table 2.3, an increase in family permanent income leads to an increase in parental transfers and a lower realized discount factor for the child. Hence, our tough love altruism model predicts a negative relationship between wealth and the discount factor.

Based on our earlier discussion (Sect. 2), there are two competing hypotheses regarding the relationship between wealth and discount factors: Fisher's hypothesis of a positive relationship and Uzawa's negative relationship hypothesis. Becker and Mulligan (1997) provided a theoretical framework for endogenous discounting, where an individual's discount factor is affected by investment in future-specific human capital intended to improve the vividness of his/her imagination. We also presented a brief discussion of the existing empirical evidence and it seems mixed at best. We believe that the mixed nature of empirical findings is consistent with the view that reality is probably best described by a combination of the two hypotheses. In our view, a child who experiences low consumption will grow to be more patient because he/she can more vividly imagine future misery. At the same time, a wealthier parent is more likely to invest in the child's human capital to help the child see the future more vividly. In this chapter, we aim to develop a simple model that captures our intuition of tough love, which is that a parent allows suffering so that the child can learn to be more patient. Consistent with this tough love intuition, in our model low childhood consumption leads to more patience (higher discount factor) in adulthood. Hence, the prediction of a negative relationship between family permanent income and the child's discount factor illustrated here is a manifestation of our tough love intuition. We believe that the addition of the human capital nature of endogenous discounting propounded by Becker and Mulligan (1997) may help

Table 2.3 Effect of an increase in family permanent income on the child's discount factor

Global parameters		
$\eta = 0.66$; $\sigma = 1.5$; $R = 1.2$; $\tilde{\beta} = \beta_p = 0.99$; $\beta_0 = 0$; $y_1 = 1$; $y_2 = 10$; $a = 0.02$		
	(1)	(2)
	$y_p = 10$	$y_p = 11.67$
T^*	1.7282	2.1420
Child's discount factor ($\beta(C_1^*)$)	0.9483	0.9409

our model to best capture the reality. However, for simplicity, we abstract from the human capital approach in our paper and leave such an extension for future work.

4 How Important Is Tough Love?

The main result of our tough love altruism model is that the parent will decrease transfers in response to an exogenous decrease in the child's discount factor. Our model modifies the standard altruism model in two ways. Do we need both of these modifications in order to obtain this result? In order to answer this question, we analyze two alternative models of altruism. First, we modify the standard altruism model by assuming that the parent evaluates the child's lifetime utility with a higher constant discount factor than that of the child. However, we do not introduce endogenous discounting in this model. This model is called the paternalistic altruism model. Second, we modify the standard altruism model by introducing endogenous discounting on the part of the child. However, we assume that the parent will use the child's endogenous discounting to evaluate the child's lifetime utility.

4.1 Paternalistic Altruism Model

In this model, both the parent and the child use constant discount factors to evaluate future utility. However, unlike the standard altruism model, here the discount factor used by the parent is higher than the child's discount factor, i.e. $\beta_{t,p} > \beta_{t,k}$, where $\beta_{t,p}$ is the discount factor used by the parent to evaluate the child's future utility and $\beta_{t,k}$ is the discount factor used by the child in period t. The parent's problem is given by:

$$\max_T \left\{ \eta\, V(R(y_p - T), \tilde{\beta}) + \tilde{\beta}(1-\eta)\Big[u(y_1 + T) + \beta_{2,p} u(C_2^*) + \beta_{2,p}\beta_{3,p} u(R(y_2 - C_2^*))\Big] \right\}, \tag{2.17}$$

subject to:

$$\{C_2^*\} \equiv \arg\max_{C_2} \Big[u(C_2) + \beta_{3,k} u(R(y_2 - C_2))\Big]. \tag{2.18}$$

As before, we solve the child's optimization problem first, which gives us the optimal period 2 consumption of the child:

$$C_2^* = C_2(y_2, \beta_{3,k}, R). \tag{2.19}$$

The optimal period 2 consumption of the child is independent of the period 1 transfers of the parent, so it can be dropped from the parent's optimization program. We rewrite the parent's problem described by Eqs. (2.17) and (2.18) as:

$$\max_T \Big[\eta V(R(y_p - T), \tilde{\beta}) + \tilde{\beta}(1-\eta)u(y_1 + T)\Big]. \tag{2.20}$$

From the first-order condition for the above problem, in principle we can solve for the optimal period 1 transfers as:

$$T^* = T(y_p, y_1, \tilde{\beta}, R, \eta). \tag{2.21}$$

We now consider an exogenous decrease in the child's discount factor, $\beta_{3,k}$. From Eq. (2.21), optimal period 1 transfers by the parent are independent of the discount factor of the child. Therefore, like the standard altruism model, in this model there is no effect of a decrease in the discount factor on the period 1 transfers. Thus, we cannot replicate the tough love altruism prediction of lower parental transfers in response to a decrease in the child's discount factor by only introducing paternalistic altruism. Next, we show that only adding endogenous altruism to the standard altruism model is also not sufficient to generate the positive relationship between parental transfers and the child's discount factor implied by our tough love altruism model.

4.2 Endogenous Altruism Model

In this model, as was assumed in the tough love altruism model, the discount factor used by the child is endogenously determined as a decreasing function of period 1 consumption:

$$\beta_{t,k}(c_1) \quad ; \quad \frac{\partial \beta_{t,k}}{\partial C_1} < 0.$$

With the borrowing constraint faced by the child in period 1, the discount factor is given by $\beta_{t,k}(y_1 + T)$. However, unlike the tough love altruism model, the parent also uses the above discount factor for evaluating the child's future utility. So the key difference is the assumption:

$$\beta_{t,p}(x) = \beta_{t,k}(x) = \beta_t(x).$$

The parent's problem in this model is given by:

$$\max_T \left\{ \eta\, V(R(y_p - T), \tilde{\beta}) + \tilde{\beta}(1-\eta)\Big[u(y_1 + T) + \beta_2(y_1 + T)u(C_2^*) \right.$$
$$\left. + \beta_2(y_1 + T)\beta_3(y_1 + T)u(R(y_2 - C_2^*))\Big] \right\}, \quad (2.22)$$

subject to:

$$\{C_2^*\} \equiv \arg\max_{C_2} \Big[u(C_2) + \beta_3(y_1 + T)u(R(y_2 - C_2))\Big]. \quad (2.23)$$

As before, we first solve the child's optimization problem, which gives us the optimal period 2 consumption:

$$C_2^* = C_2(y_2, \beta_3(y_1 + T), R). \quad (2.24)$$

We solve the problem described in Eqs. (2.22) and (2.23) numerically as a nonlinear root finding problem. The solution method, parameterization, and the parametric values used are identical to those we used for the comparative statics exercises for the tough love altruism model in Sect. 3.2.

We consider an exogenous decrease in the discount factor of the child achieved by decreasing the preference parameter β_0, and then trace out the effect of this change on the period 1 transfers, T. The results for the assumed set of model parameter values are summarized in Table 2.4. We find that as β_0 is reduced monotonically, parents in the endogenous altruism model will reduce transfers.

The results of this exercise seem to suggest that endogenous discounting is enough to obtain the main result of our tough love altruism model. With the given set of parameter values, this model also predicts a positive relationship between parental transfers and the child's discount factor. However, unlike the results of the tough love model, the result reported in Table 2.4 is very sensitive to the assumption made on σ. Table 2.5 below presents simulation results with $\sigma < 1$.

Now we find that as β_0 falls, transfers increase monotonically. Hence, with the endogenous altruism model, the direction of the relationship between parental trans-

Table 2.4 Endogenous altruism model

Global parameters				
$\eta = 0.66$; $\sigma = 1.5$; $R = 1.2$; $y_1 = 1$; $y_2 = 10$; $y_p = 10$; $a = 0.02$; $\tilde{\beta} = 0.99$				
Optimum	$\beta_0 = 0$	$\beta_0 = -0.4$	$\beta_0 = -0.6$	$\beta_0 = -0.8$
T^*	2.0206	1.9415	1.8995	1.8526
C_1^*	3.0206	2.9415	2.8995	2.8526
C_2^*	5.2495	6.1446	6.8350	7.9305
C_3^*	5.7006	4.6265	3.7981	2.4834
$\beta(C_1^*)$	0.9430	0.5444	0.3452	0.1460

Table 2.5 Endogenous altruism model

Global parameters				
$\eta = 0.66; \sigma = 0.7; R = 1.2;$ $y_1 = 2; y_2 = 10;\ y_p = 10;\ a = 0.02; \tilde{\beta} = 0.99$				
Optimum	$\beta_0 = 0$	$\beta_0 = -0.4$	$\beta_0 = -0.6$	$\beta_0 = -0.8$
T^*	0.0428	0.1955	0.2751	0.3483
C_1^*	1.0428	1.1955	1.2751	1.3483
C_2^*	4.8784	6.7003	7.8960	9.1850
C_3^*	6.1459	3.9596	2.5247	0.9780
$\beta(C_1^*)$	0.9796	0.5766	0.3751	0.1737

fers and the child's discount factor rests critically on the parametric assumptions. This is in sharp contrast to our tough love altruism model where this relationship is positive regardless of the parametric specification.

Thus to conclude, the results of this section show that in order to obtain the prediction that the parent's transfer decreases in response to an exogenous decrease in the child's discount factor, we need to introduce both endogenous discounting and paternalistic evaluation by the parent of the child's lifetime utility.

5 Tough Love Altruism Model with Leisure

We constructed the tough love model so that transfers decrease when the child's discount factor exogenously falls. We now turn to another type of comparative statics for the purpose of challenging our tough love model. Here, we examine the model's properties related to the relationship between the child's income and parental transfers. Until now, we have considered an economy where agents derive utility only from consumption. To examine the role of the child's income in a more realistic way, we now generalize our setup by allowing for leisure as a choice variable for the child. This is motivated by empirical evidence against the standard altruism model's redistributive neutrality property. As we have discussed earlier, the empirical evidence for redistributive neutrality is largely unfavorable. Although our tough love model also implies redistributive neutrality, such neutrality only holds for exogenous current income changes. We study below how endogenous changes in income caused by an exogenous change in the child's discount factor are related to transfers.

For this experiment, we extend our model in an important dimension. Until now, for notational simplicity, we assumed that the three periods of the child's life are of equal duration. In reality, we can expect them to vary. Allowing the duration to vary, we denote that of the childhood period by τ_1, that of the work period by τ_2, and that of the retirement period by τ_3.[21] We imagine the childhood period of the model to correspond with the period around high school and the early years of college in which children may engage in part-time work (e.g. 16–20 years of age) and set the duration to be 5 years.[22] The benchmark duration of the work period of the model is set to be 40 years, and corresponds to ages between 21 and 60 years. The benchmark duration for the retirement period is set to be 20 years, and corresponds to ages between 61 and 80 years.

We continue to assume perfect information. In our setup, this implies that the parent can fully observe the child's effort level. The remaining model assumptions are retained with transfers being made only in period 1 and with the child being borrowing constrained in period 1. The following notation is used. L_1 and L_2 denote the amount of leisure consumed by the child in period 1 and period 2, respectively. w_1 and w_2 denote the wage income of the child in the two periods. For simplicity, we assume that the child earns no wage income in period 3 and simply consumes her past savings. The parent's problem is:

$$\max_T \left\{ \eta\, V(R(y_p - T), \tilde{\beta}, \tau_1, \tau_2) + \tilde{\beta}(1-\eta)\Big[\tau_1\, u(w_1(1-L_1^*) + T, L_1^*) \right.$$
$$\left. + \beta_{2,p}\, \tau_2\, u(C_2^*, L_2^*) + \beta_{2,p}\beta_{3,p}\, \tau_3\, u(R(w_2(1-L_2^*) - C_2^*))\Big] \right\}, \tag{2.25}$$

subject to:

$$\{C_2^*, L_1^*, L_2^*\} \equiv \underset{C_2, L_1, L_2}{\arg\max}\Big[\tau_1\, u(w_1(1-L_1) + T, L_1) + \beta_{2,k}(w_1(1-L_1) + T)$$
$$\tau_2\, u(C_2, L_2) + \beta_{2,k}(w_1(1-L_1) + T)\beta_{3,k}(w_1(1-L_1) + T)$$
$$\tau_3 u(R(w_2(1-L_2) - C_2))\Big]. \tag{2.26}$$

We solve the above problem numerically as a nonlinear root finding problem and for that purpose we impose the following parametric specification:

$$u(C, L) = Log(C) + d\,\frac{L^{1-\gamma}}{1-\gamma} \;\; (Child's\ Utility\ Function), \tag{2.27}$$

[21]For simplicity, we abstract from the child's early life in which he does not face the work–leisure choice.

[22]Cunha et al. (2006) present a survey of empirical evidence that later interventions in adolescent years can affect noncognitive skills such as patience, self-control, temperament, time preferences, etc., while these interventions cannot affect cognitive skills.

$$v(C) = Log(C) \ \ (Parent's\ Utility\ Function). \tag{2.28}$$

The child's discount function is given by:

$$\beta(w_1(1-L_1)+T) = \beta_0 + \frac{1}{1+a(w_1(1-L_1)+T)} \tag{2.29}$$
$$\text{where} \quad a > 0 \text{ and } \beta_0 < 0.$$

There are two important parameters of the period utility function to which we need to assign values for solving the optimization problem numerically. The first parameter, γ, is the reciprocal of the IES for labor/leisure. For men, most estimates of the intertemporal labor-supply elasticity are between 0 and 0.5 (Altonji 1986; Blundell and MaCurdy 1999; French 2004; MaCurdy 1981). Ham and Reilly (2013) used an implicit contract model and found this elasticity to be 0.9. To be consistent with the literature, we report results for two values of γ. The first is $\gamma = 1.11$, which is consistent with the elasticity of 0.9. The second is $\gamma = 2$, which is consistent with the elasticity of 0.5.

The second parameter, d, captures the weight of leisure in the child's period utility function. In the real business cycle literature, the weight on leisure in the utility function is usually calibrated so that in the steady state the representative household spends about one-third of its total time working. In our model, given our parametric specification for the period utility, we obtain the following optimality condition that determines period t consumption-leisure choice:

$$d = \frac{W_t}{C_t} * L_t^{-\gamma}.$$

We set $L_t = 2/3$ and used the average value for labor income to consumption expenditure (for the period 1980–2009) from National Income and Product Account (NIPA) to approximate $\frac{W_t}{C_t}$. Then, using $\gamma = 1.11$, we obtain $d = 1.33$ and using $\gamma = 2$, we obtain $d = 1.9125$. We report simulation results for each combination of γ and d separately.

Table 2.6 summarizes the results of the simulations for a decrease in the parameter β_0.

We observe that as β_0 falls from 0 to -0.01, the parent with a tough love motive lowers transfers to the child. At the same time, there is also a fall in the child's income in the first period corresponding to the fall in β_0.

Thus, in our tough love altruism model, the parent's transfers and the child's income fall at the same time even though the child is borrowing constrained. Whether or not this feature of our model can explain the finding of Altonji et al. is an empirical problem that requires careful study of the PSID data. This depends, among other things, on how income changes are divided into endogenous and exogenous changes. However, the model does imply that the parent's transfers and

Table 2.6 Tough love altruism model with Leisure

Global parameters				
$\eta = 0.66; r = 1.02; a = 0.02; \tilde{\beta} = \beta_p = 0.99;$ $w_1 = 1; w_2 = 10; y_p = 10; \tau_1 = 5; \tau_2 = 40; \tau_3 = 20$				
	(1)		(2)	
	$\gamma = 1.11; d = 1.33$		$\gamma = 2; d = 1.9125$	
	$\beta_0 = 0$	$\beta_0 = -0.01$	$\beta_0 = 0$	$\beta_0 = -0.01$
$T*$	0.1869	0.1076	0.5198	0.4498
Child's first period income	3.1522	2.8155	1.0392	0.8275

the recipient's income can move in the same direction even when the recipient is borrowing constrained. This can potentially reconcile the apparent inconsistency between empirical results against the redistributive neutrality property and Laitner and Juster's (1996) result in favor of parents' altruism for children. They used Teachers Insurance and Annuity Association-College Retirement Equities Fund (TIAA-CREF) data and focused on bequests as the channel for parental altruism. They found that for the subsample of respondents characterized by willingness to leave a bequest, the projected amount of the bequest is largest for households with lowest assessments of their children's likely earnings in the future.

6 Conclusion

In the simple setting of a three-period economy with a single parent and single child, perfect information, and borrowing constraints, we develop a model of intergenerational altruism wherein the tough love motive for parents is a driving force behind the parent's behavior. In our tough love altruism model, the child's discount factor is endogenously determined, and the parent evaluates the child's lifetime utility with a constant discount factor that is higher than that of the child. With our modeling, we try to capture our intuition of tough love: in order to teach a child to be patient, the parent is willing to let the child suffer in the short run. In order to capture this intuition in a simple model, we abstract from the human capital nature of endogenous discounting.

The main prediction of our tough love model, for which we provide an analytical proof, is that an exogenous decrease in the child's discount factor lowers parental transfers. This prediction of our model is in contrast with that of the standard altruism model, in which the parent does not change transfers when the child becomes impatient. Since exogenous changes in the child's discount factor that make her impatient are likely to cause behavior that calls for the parent's corrective actions, the tough love altruism model is more consistent with empirical evidence on parental punishments as well as the role of parents in shaping children's preferences as compared with the standard altruism model.

Another contribution of our paper relates to the empirical evidence against the standard altruism model's redistributive neutrality property (also called the transfer derivative restriction). Our tough love altruism model also implies redistributive neutrality. However, this redistributive neutrality only holds for exogenous income changes, while in the data, we can have both endogenous as well as exogenous changes in income. In the version of the tough love altruism model with endogenous leisure choices for the child, we investigate how endogenous changes in income caused by an exogenous change in the child's discount factor are related to transfers. We find that an exogenous change in the discount factor to make the child more impatient can cause both lower income and lower transfers from the parent even when the child is borrowing constrained. This prediction of our model may be able to explain the empirical findings by Altonji et al. (1997) depending on how income changes are divided into endogenous and exogenous changes among other factors.

An important stylized fact for the U.S. economy is that the distribution of wealth is very concentrated and skewed to the right. Castaneda et al. (2003) emphasized that standard explanations based on household decision-making models with homogeneous preferences fail to account for this observed heterogeneity in wealth distribution. There is some evidence that heterogeneity in discount factors may be important in understanding differences in savings rates and hence in wealth accumulation. For example, Krusell and Smith (1998) found that incorporation of discount-rate heterogeneity markedly decreases the gap between model predictions and the observed wealth distribution. An interesting feature of our model is that it suggests a pecuniary channel through which parents with tough love motives can instill the virtue of patience in their children. Since not all parents will exhibit such tough love tendencies or the same degree of tough love, our model offers one rationale for individuals discounting the future at different rates depending on their parental tough love. To the extent that there is a link between heterogeneity in discount factors and heterogeneity in savings rates, this feature of our model has implications for the observed heterogeneity in wealth in the U.S.

In the future, it will be interesting to analyze the characteristics of parents who exhibit tough love in their children's upbringing. Bhatt and Ogaki (2009) suggest that the worldview of the parent may be an important factor. For example, how the parent views suffering may be important. If the parent views suffering as meaningless, it is harder for him to let the child suffer. If the parent views suffering as meaningful (e.g., educational), then it is easier for him to let the child suffer. A new direction for research in this field will be to utilize existing or new survey data on parents' view on suffering to infer their capacity to exhibit tough love as explained in Kubota et al. (2011, 2013). It will also be of interest to measure both the child's and the parent's discount factors, and use them to directly test the implications of our tough love model. Research efforts to conduct experiments for parent-child pairs for this purpose have already started.

Another interesting future research question is to explore channels that may compete with parental tough love in influencing their children's discount factor. For instance, parents may make efforts to select the peers and/or friends of their children. This can be achieved, presumably at some cost, for example, the cost of selecting a good neighborhood or a good school. It would be interesting to conceive of an alternative model where parents can incur such selection costs to ensure transmission of desired discount factors to their children. We can then analyze how large these costs have to be in order to substitute for tough love parenting methods suggested by our model.[23]

Finally, in this chapter, we have abstracted from the Becker–Mulligan type of human capital investment, which increases the discount factor for the child. In the future, it will be interesting to incorporate such an aspect into our tough love altruism model and investigate the impact of such modification on the predictions of our model regarding the relationship between parental transfers and the child's discount factor on one hand, and between parental transfers and their children's income (ability) on the other.

Acknowledgements We thank Shinsuke Ikeda for his initial collaboration, Dean Lillard for comments and discussions from the initial stage of this research, three anonymous referees, Fumio Hayashi, Nobuhiro Kiyotaki; and seminar and session participants at the 2008 North American Summer Econometric Society, Midwest Economics, and Midwest Macroeconomics Meetings; Aoyama Gakuin, Baskent, Bilkent, Doshisha, Fukuoka, Hiroshima, Hitotsubashi, Keio, Kobe, Koc, Kyoto, Nagoya, Nihon, Ohio State, Osaka, Osaka Prefecture, Sabanci, Waseda, Yale Universities; and the Universities of Cyprus and Tokyo for comments. This research of Masao Ogaki was partially supported by Grants-in-Aid for Scientific Research from Japan Society for the Promotion of Science.

Appendices

A Proof for Proposition 1

In this section of the appendix we provide an analytical proof of our main result specified in Eq. (2.13):

$$\frac{\partial T^*}{\partial \beta_0} > 0 \ \textit{iff} \ \left[1 + R(\beta_p - \beta_{3,k})\frac{u''(R(y_2 - C_2^*))}{u'(R(y_2 - C_2^*))}\frac{\partial C_2^*}{\partial \beta_0} - (\beta_p - \beta_{3,k})\frac{\partial^2 C_2^*}{\partial \beta_0 \partial T}\right] > 0.$$

We start the derivation of the above result by noting that the parent accounts for C_2^* when maximizing utility by choosing transfers, T. From the first-order condition for the parent's problem described in Eqs. (2.10) and (2.11), we obtain:

[23] We thank an anonymous referee for bringing this point to our notice as a potential future work.

$$V'(R(y_p - T^*), \tilde{\beta}) = \left(\frac{1-\eta}{\eta}\right)\frac{\tilde{\beta}}{R}\left(u'(y_1 + T^*) + \beta_p u'(C_2^*)\frac{\partial C_2^*}{\partial T} - \beta_p^2 R u'(R(y_2 - C_2^*))\frac{\partial C_2^*}{\partial T}\right),$$

where T^* denotes optimal parental transfers, $V'(.) = \dfrac{\partial V(.)}{\partial T}$, and $u'(.) = \dfrac{\partial u(.)}{\partial T}$.

Now consider an exogenous change in the child's discount factor (β_k) captured by a change in β_0. From the parent's first-order condition described above we obtain:

$$\frac{\partial T^*}{\partial \beta_0} = A * \left(\frac{\partial}{\partial \beta_0}\left[\beta_p u'(C_2^*)\frac{\partial C_2^*}{\partial T} - \beta_p^2 R u'(R(y_2 - C_2^*))\frac{\partial C_2^*}{\partial T}\right]\right),$$

where:

$$A = - \frac{\left(\frac{1-\eta}{\eta}\right)\tilde{\beta}}{V''(R(y_p - T^*), \tilde{\beta}) + \left(\frac{1-\eta}{\eta}\right)\tilde{\beta} u''(y_1 + T^*)}.$$

Given concavity of V(.) and u(.), we know that $A > 0$. Hence:

$$sign\left(\frac{\partial T^*}{\partial \beta_0}\right) = sign\left(\frac{\partial}{\partial \beta_0}\left[\beta_p u'(C_2^*)\frac{\partial C_2^*}{\partial T} - \beta_p^2 R u'(R(y_2 - C_2^*))\frac{\partial C_2^*}{\partial T}\right]\right).$$

Then $\dfrac{\partial T^*}{\partial \beta_0} > 0$ if and only if:

$$\frac{\partial}{\partial \beta_0}\left[\beta_p u'(C_2^*)\frac{\partial C_2^*}{\partial T} - \beta_p^2 R u'(R(y_2 - C_2^*))\frac{\partial C_2^*}{\partial T}\right] > 0.$$

Now using the first-order condition for the child's problem, we can rewrite the above condition as follows.

$$\frac{\partial}{\partial \beta_0}\left[-(\beta_p - \beta_{3,k}) R u'(R(y_2 - C_2^*))\frac{\partial C_2^*}{\partial T}\right] > 0$$

It is straightforward to show that the LHS of the above expression is given by:

$$R u'(R(y_2 - C_2^*))\frac{\partial C_2^*}{\partial T}\left(1 + R(\beta_p - \beta_{3,k})\frac{u''(R(y_2 - C_2^*))}{u'(R(y_2 - C_2^*))}\frac{\partial C_2^*}{\partial \beta_0} - (\beta_p - \beta_{3,k})\frac{\partial^2 C_2^*}{\partial \beta_0 \partial T}\right).$$

Since $\frac{\partial C_2^*}{\partial T} > 0$ and $u'(.) > 0$, the above expression is strictly positive if and only if:

$$\left[1 + R(\beta_p - \beta_{3,k})\frac{u''(R(y_2 - C_2^*))}{u'(R(y_2 - C_2^*))}\frac{\partial C_2^*}{\partial \beta_0} - (\beta_p - \beta_{3,k})\frac{\partial^2 C_2^*}{\partial \beta_0 \partial T}\right] > 0.$$

This establishes our claim in Eq. (2.14). Now, given positive marginal utility, concavity of $u(.)$, and $\frac{\partial C_2^*}{\partial \beta_0} < 0$, the sufficient conditions for the above expression to be strictly positive are:

(i) $\beta_p \geq \beta_{3,k}$ and
(ii) $\frac{\partial^2 C_2^*}{\partial \beta_0 \partial T} \leq 0$.

A Proof for Proposition 2

In this section we show that $\frac{\partial^2 C_2^*}{\partial \beta_0 \partial T} \leq 0$ will depend on the convexity of the marginal utility (as captured by the positive third derivative of the utility function) and the impatience level of the child.

We start our derivation with the partial derivative of the optimal second period consumption with respect to parental transfers:

$$\frac{\partial C_2^*}{\partial T} = \frac{\beta'(y_1 + T)Ru'(R(y_2 - C_2^*))}{[u''(C_2^*) + \beta_{3,k}R^2 u''(R(y_2 - C_2^*))]},$$

where

$$\beta_{3,k} = \beta_0 + \psi(y_1 + T).$$

Differentiating with respect to β_0

$$\frac{\partial^2 C_2^*}{\partial \beta_0 \partial T} = \frac{N}{D},$$

where

$$D = [u''(C_2^*) + \beta_{3,k}R^2 u''(R(y_2 - C_2^*))]^2 > 0.$$

and

$$N = [u''(C_2^*) + \beta_{3,k}R^2 u''(R(y_2 - C_2^*))][-\psi'(y_1 + T)R^2 u''(R(y_2 - C_2^*))\frac{\partial C_2^*}{\partial \beta_0}]$$

$$-\psi'(y_1+T)Ru'(R(y_2-C_2^*))[u'''(C_2^*)\frac{\partial C_2^*}{\partial \beta_0}+R^2u''()-\beta_{3,k}R^3u'''(R(y_2-C_2^*))\frac{\partial C_2^*}{\partial \beta_0}].$$

Since, D is always positive, the sign of $\frac{\partial^2 C_2^*}{\partial \beta_0 \partial T}$ depends on the sign of N. Now, the sign of N will be the same as the sign of:

$$sign([u'''(C_2^*)\frac{\partial C_2^*}{\partial \beta_0}+R^2u''(R(y_2-C_2^*))-\beta_{3,k}R^3u'''(R(y_2-C_2^*))\frac{\partial C_2^*}{\partial \beta_0}]).$$

If the above expression is negative then $N < 0$.

Hence, the condition for $N < 0$ is:

$$[u'''(C_2^*)\frac{\partial C_2^*}{\partial \beta_0}+R^2u''(C_3^*)\{1-\beta_{3,k}R\frac{u'''(C_3^*)}{u''(C_3^*)}\frac{\partial C_2^*}{\partial \beta_0}\}] < 0.$$

The above condition holds if,

(i) $u'''(.) \geq 0$ and
(ii) $\beta_{3,k}\ R\ G \leq 1$,

where

$$G = \left(\frac{u'''(C_3^*)}{u''(C_3^*)}\right)\frac{\partial C_2^*}{\partial \beta_0}.$$

Power Utility Function

In this section of the appendix we use the power utility function to interpret the following condition:

$$\beta_{3,k}\ R\ G \leq 1,$$

where

$$G = \left(\frac{u'''(C_3^*)}{u''(C_3^*)}\right)\frac{\partial C_2^*}{\partial \beta_0}.$$

We assume that the period utility function is given by:

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}.$$

Using the above specification of the utility function, from the child's optimization problem, we get:

$$C_2^* = \frac{Ry_2}{R + (\beta_{3,k}R)^{\frac{1}{\sigma}}}$$

$$C_3^* = R(y_2 - C_2^*) = (\beta_{3,k}R)^{\frac{1}{\sigma}} C_2^*$$

$$\frac{\partial C_2^*}{\partial \beta_0} = -\frac{R^2 y_2 (\beta_{3,k}R)^{\frac{1-\sigma}{\sigma}}}{\sigma[R + (\beta_{3,k}R)^{\frac{1}{\sigma}}]^2}$$

$$\frac{u'''(C_3^*)}{u''(C_3^*)} = -(\sigma + 1)(\beta_{3,k}R)^{-\frac{1}{\sigma}} \left(\frac{R + (\beta_{3,k}R)^{\frac{1}{\sigma}}}{Ry_2} \right)$$

Hence, for the power utility case we get:

$$G = \left(\frac{\sigma + 1}{\sigma} \right) \frac{1}{\beta_{3,k}[R + (\beta_{3,k}R)^{\frac{1}{\sigma}}]}$$

Using the above expression for G we can rewrite the inequality of interest as:

$$\beta_{3,k}R \left(\frac{\sigma + 1}{\sigma} \right) \frac{1}{\beta_{3,k}[R + (\beta_{3,k}R)^{\frac{1}{\sigma}}]} \leq 1$$

Rearranging, we get the following condition:

$$\beta_{3,k}R \leq \left(\frac{\sigma}{R} \right)^{\frac{1}{\sigma}}$$

Addendum: Recent Developments[24]

Two directions of research that are closely related to the tough love model have been developing since the working paper version (Bhatt and Ogaki 2008) of this chapter was published. One direction is for normative economics and the other is for positive economics. The first direction is to add an element of virtue ethics to welfarism that has been the basis of normative economics. The second direction is to empirically evaluate the tough love model.

[24]This addendum has been newly written for this book chapter.

In Bhatt et al. (2015a), we proposed an approach of normative economics for models with endogenous preferences that adds an element of virtue ethics to the traditional formulation based on the Bergson-Samuelson social welfare function (SWF). Our approach is based on the moral evaluation function (MEF) that evaluates different endogenous preferences in terms of moral virtue and on the social objective function (SOF), which is a function of SWF and MEF in order to express a balanced value judgment. This is in a sense a response to Sandel's (2013) call to introduce more value judgment into economics based on his ideas of political philosophy explained for wide audience in Sandel (2009, 2012). In the same issue of the *Journal of Economic Perspectives* as in Sandel (2013) and Bruni and Sugden (2013) argued that neoclassical economics has already incorporated virtue ethics if we think about market virtues. Our approach can incorporate market virtues, but can also incorporate other virtues such as altruism toward a disabled stranger as shown in Bhatt et al. (2015a). In Bhatt and Ogaki (2014, Rational addiction and optimal taxation: a reexamination based on the social objective function, unpublished), we applied our approach to the rational addiction model. If a society finds that preferences with drug addiction caused by past drug consumption is less virtuous, preferences affected by drug addiction are less virtuous than preferences that are not.

For the purpose of illustrating our approach by an example, Bhatt et al. (2015b) applies it to a version of Bhatt and Ogaki's tough love model. In this version, the original model is extended by bequest and bequest tax. In the model, the parent thinks that he should not spoil the child so that the child will grow to be patient, but is tempted to spoil the child because he enjoys watching his child having higher childhood utility. With a bequest motive, the parent can use the money saved by lowering childhood transfers in order to increase his bequest given to the child after she grows up. In this version of the model, the government has a policy tool of the bequest tax rate that can be used to influence the optimizing behaviors of the parent and the child. When the bequest tax rate is higher, the parent is more tempted to spoil the child because the money he saves by giving less transfers to the child will be taxed away when he gives the bequest.

If our informal discussions with many economists give a good guidance, many economists seem to think that moral value ethics is not desirable for public policy evaluations because they do not want the government to influence people's preferences. However, in our model, the government does influence the child's preferences as long as the bequest tax rate is not zero. The optimum tax rate is positive when the SWF is maximized. On the other hand, the optimum tax rate is zero when the SOF is maximized with $\alpha = 0.3$. Thus, introducing moral virtue ethics may result in a policy that does not affect people's preferences. This illustrates that introduction of moral virtue ethics does not necessarily mean that the government starts to influence people's preferences. Because any government policy may be influencing people's preferences even when the government does not intentionally do so, it seems important to examine how each policy is influencing people's preferences and think about how desirable such influences are for the society.

This line of thoughts leads us to the second direction of research. In order to examine whether or not any policy is influencing people's preferences, we need empirical work on models with endogenous preferences in which such policy can affect preferences. For the tough love model with the bequest tax rate, there already exit some empirical work. A starting point of any model with endogenous time discounting is that genetic factors do not completely determine time discounting. Using a unique data set of twins in Japan, Hirata et al. (2010) found empirical evidence in favor of this.

Kubota et al. (2011) examined how parents' tendencies for tough love behavior depend on various measures of time discounting for parents' own lending and borrowing over different time horizons. Using the Osaka University Global COE survey data for Japan and the United States, they found evidence that is consistent with the tough love model. They also found one empirical puzzle that proportionately more U.S. parents show tough love to young children before the school age than Japanese parents even after controlling for time discounting and other economic and demographic factors. The is especially puzzling because more patient parents tend to show tough love, and Japanese parents are estimated by Kubota et al. (2011) to be more patient than U.S. parents.

Kubota et al. (2013) examined a possible solution to this puzzle is cultural differences between the two countries. They used a framework in anthropology that a worldview is behind each culture. Using the Osaka University Global COE survey data for Japan and the United States, they found that differences in the distributions of the confidence in worldview beliefs can explain a substantial portion of the international differences in the parental attitudes.

Akabayashi et al. (2014) reported evidence that is more consistent with the tough love model than with the standard Barro-Becker model from experiments with parent-child pairs. In their time preference experiments, each parent-child pair makes individual and joint decisions about how much and when the child receives a payment (e.g., 1,000 yen now versus 1,001 yen in 2 months). The Barro-Becker model in which the parent does not know the child's time discount rate predicts that the only reason for the parent to make a different decision from the child when s/he makes an individual decision is the lack of knowledge about the child's time discount factor. Hence the joint decision is predicted to be the same as the child's individual decision. However, they find that about half of the pairs do not follow this prediction.

References

Akabayashi H (2006) An equilibrium model of child maltreatment. J Econ Dyn Control 30:993–1025

Akabayashi H, Kamesaka A, Nakamura R, Ogaki M, Tamura T (2014) An experimental study of intergenerational altruism with parent-child pairs. Institute for economic studies discussion paper 2014/8, Keio University

Altonji GJ (1986) Intertemporal substitution in labor supply: evidence from micro data. J Politi Econ 94:1121–1166
Altonji GJ, Hayashi F, Kotlikoff JL (1997) Parental altruism and inter vivos transfers: theory and evidence. J Politi Econ 105:1121–1166
Atkeson A, Ogaki M (1996) Wealth-varying intertemporal elasticities of substitution: evidence from panel and aggregate data. J Monet Econ 38:507–534
Barro JR (1974) Are government bonds new wealth? J Politi Econ 82:1095–1117
Baumrind D (1966) Effects of authoritative parental control on child behavior. Child Dev 37:887–907
Becker G (1974) A theory of social interactions. J Politi Econ 82:1063–1093
Becker G, Mulligan CB (1997) The endogenous determination of time preference. Q J Econ 112:1063–1093
Bhatt V (2011) Adolescent alcohol use and intergenerational transfers: evidence from micro data. J Fam Econ Issues 32:296–307
Bhatt V, Ogaki M (2008) Tough love and intergenerational altruism. Rochester center for economic research working paper series 544
Bhatt V, Ogaki M (2009) Tough love, intergenerational altruism, and worldviews. J Behav Econ Finance 2:1–6
Bhatt V, Ogaki M (2012) Tough love and intergenerational altruism. Int Econ Rev 53:791–814
Bhatt V, Ogaki M, Yaguchi Y (2015a) A reformulation of normative economics for models with endogenous preferences. JSPS Grants-in-Aid for Scientific Research (S) understanding persistent deflation in Japan working paper series no. 067, The University of Tokyo
Bhatt V, Ogaki M, Yaguchi Y (2015b) Normative behavioural economics based on unconditional love and moral virtue. Jpn Econ Rev 66(2):226–246
Bisin A, Verdier T (2001) The economics of cultural transmission and the dynamics of preferences. J Econ Theory 97:298–319
Blundell R, MaCurdy TE (1999) Labor supply: a review of alternative approaches. In: Aschenfleter Q, Card D (eds) Handbook of labor economics. Elsevier, Amsterdam, pp 1559–1695
Bruni L, Sugden R (2013) Market reasoning as moral reasoning: why economists should re-engage with political philosophy. J Econ Perspect 27:141–164
Caplin A, Leahy J (2004) The social discount rate. J Politi Econ 112:1257–1268
Carlson L, Grossbart S (1988) Parental style and consumer socialization of children. J Consum Res 15:77–94
Castaneda A, Diaz-Gimenez J, Rios-Rull JV (2003) Accounting for the U.S. earnings and wealth inequality. J Politi Econ 111:818–857
Choi H, Mark NC, Sul D (2008) Endogenous discounting, the world saving glut and the U.S. current account. J Int Econ 75:30–53
Cunha F, James H, Lochner L, Masterov DV (2006) Interpreting the evidence on life cycle skill formation. In: Hanushek EA, Stephen M, Ludger W (eds) Handbook of the economics of education. Elsevier, Amsterdam, pp 697–812
Doepke M, Zilibotti F (2008) Occupational choice and the spirit of capitalism. Q J Econ 123:747–793
Farhi E, Werning I (2007) Inequality and social discounting. J Politi Econ 115:365–402
Fernandez R, Fogli A, Olivetti C (2004) Mothers and sons: preference formation and female labor force dynamics. Q J Econ 119:1249–1299
Fisher I (1930) The theory of interest. Macmillan, New York
French E (2004) The labor supply response to (mismeasured but) predictable wage changes. Rev Econ Stat 86:602–613
Hall RE (2009) Reconciling cyclical movements in the marginal value of time and the marginal product of labor. J Politi Econ 117:281–323
Ham JC, Reilly KT (2013) Implicit contracts, life cycle labor supply, and intertemporal substitution. Int Econ Rev 54:1133–1158
Hao L, Hotz VJ, Jin GZ (2008) Games parents and adolescents play: risky behavior, parental reputation and strategic transfers. Econ J 118:515–555

Heckman JJ (1981) Heterogeneity and state dependence. In: Rosen S (ed) Studies in labor markets. University of Chicago Press, Chicago, pp 91–140

Heckman JJ (1991) Identifying the hand of past: distinguishing state dependence from heterogeneity. Am Econ Rev 81:75–79

Hirata K, Iiboshi H, Hayakawa K, Ikeda S, Tsutsui Y, Ohtake F (2010) Genetic inheritance of time-discounting behavior: a Bayesian approach using Markov chain Monte Carlo method. Paper presented at the 2010 Japanese Economic Association Spring Meeting, Chiba University, Chiba

Ida T, Goto R (2009) Simultaneous measurement of time and risk preferences: stated discrete choice modeling analysis depending on smoking behavior. Int Econ Rev 50:1169–1182

Ikeda S, Ohtake F, Yoshiro T (2005) Time discount rates: an analysis based on economic experiments and questionnaire surveys. Institute of Social and Economic Research Discussion Paper No. 638, Osaka University

Kagel JH, Battalio RC, Green L (1995) Economic choice theory: an experimental analysis of animal behavior. Cambridge University Press, Cambridge

Krusell P, Smith AA (1998) Income and wealth heterogeneity in the macroeconomy. J Politi Econ 106:867–896

Kubota K, Horioka CY, Kamesaka A, Ogaki M, Ohtake F (2011) Time discounting and intergenerational altruism. Paper presented at the 19th annual conference of the Nippon Finance Association, Waseda University, Tokyo

Kubota K, Kamesaka A, Ogaki M, Ohtake F (2013) Cultures, worldviews, and intergenerational altruism. Paper presented at the European regional science association 2013 congress, Palermo

Laitner J, Juster FT (1996) New evidence on altruism: a study of TIAA − CREF retirees. Am Econ Rev 86:893–908

Lawrance EC (1991) Poverty and the rate of time preference: evidence from panel data. J Politi Econ 99:54–77

MaCurdy T (1981) An empirical model of labor supply in a life-cycle setting. J Politi Econ 89:1059–1085

McGarry K (1999) Inter vivos transfers and intended bequests. J Publ Econ 73:321–351

Mischel W (1961) Father-absence and delay of gratification: cross-cultural comparisons. J Abnorm Soc Psychol 63:116–124

Mulligan CB (1998) Parental priorities and economic inequality. University of Chicago Press Economics Books, Chicago

Ogaki M, Atkeson A (1997) Rate of time preference, intertemporal elasticity of substitution, and level of wealth. Rev Econ Stat 79:564–572

Ogaki M, Reinhart CM (1998) Intertemporal substitution and durable goods: long-run data. Econ Lett 61:85–90

Ogawa K (1993) Economic development and time preference schedule: the case of Japan and East Asian NICs. J Dev Econ 42:175–195

Phelan C (2006) Opportunity and social mobility. Rev Econ Stud 73:487–504

Rose GM, Dalakas V, Kropp F (2003) Consumer socialization and parental style across cultures: findings from Australia, Greece, and India. J Consum Psychol 13:366–376

Sandel MJ (2009) Justice: what's the right thing to do? Farrar Straus & Giroux, New York

Sandel MJ (2012) What money can't buy: the moral limits of markets. Farrar Straus & Giroux, New York

Sandel MJ (2013) Market reasoning as moral reasoning: why economists should re-engage with political philosophy. J Econ Perspect 27:121–140

Schmitt-Grohé S, Uribe M (2003) Closing small open economy models. J Int Econ 44:257–287

Sleet C, Sevin Y (2005) Social credibility, social patience, and long-run inequality. Manuscript, Carnegie Mellon University

Sleet C, Sevin Y (2007) Credibility and endogenous societal discounting. Rev Econ Dyn 9:410–437

Uzawa H (1968) Time preference, the consumption function, and optimum asset holdings. In: Wolfe JN (ed) Value, capital, and growth: papers in honour of Sir John Hicks. University of Edinburgh Press, Edinburgh
Webley P, Nyhus E (2006) Parent's influence on children's future orientation and saving. J Econ Psychol 27:140–164
Weinberg AB (2001) An incentive model of the effect of parental income on children. J Politi Econ 109:266–280

Part II
Behavioral Macroeconomics

Chapter 3
Consumer Interdependence via Reference Groups

Hiroaki Hayakawa and Yiannis Venieris

Abstract In solving choice problems under bounded rationality, one relies on "heuristics" provided by social interdependence. Such "heuristics" consist in taking a particular social group as a reference group and in emulating its life-style by acquiring an associated cluster of complementary wants. A preference map generated by this reference-group-taking behavior exhibits smooth indifference curves which are convex to the origin with a "relevant range" over which the marginal rate of substitution is positive and diminishing. However, its implications on consumer choice and welfare economics are significantly different from those of traditional theory.

Keywords Consumer interdependence • Reference groups • Life-styles

1 Introduction

In spite of the various early contributions of several writers (Veblen 1899; Pigou 1913; Marshall 1920; Knight 1923), neoclassical consumer theory has been based on the joint assumptions that individual preference relations are independent and that tastes are exogenous. By large measure, these two abstractions can best be understood within the context of laissez faire and the unwillingness of economists to grapple with issues viewed as falling outside the traditional boundaries of their discipline. In the final analysis, it was argued, questions associated with interdependence and tastes could be safely ignored, as long as the resultant model retains considerable predictive power (Friedman 1962).

The original article first appeared in the Journal of Political Economy 85: 599–616, 1977. A newly written addendum has been added to this book chapter.

H. Hayakawa (✉)
University of Brunei Darussalam, Bandar Seri Begawan, Negara Brunei Darussalam
e-mail: hiroaki.hayakawa@ubd.edu.bn; hayakawah2@r4.dion.ne.jp

Y. Venieris
San Diego State University, San Diego, CA, USA

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_3

Still, to the extent that tastes are endogenous and individual relations interdependent, the predictions rendered by the neoclassical model can be improved considerably by avoiding the misspecification inherent in it (von Weizsdcker 1971). Based on this belief, two general types of modification of neoclassical theory have been proposed in the literature. In the first, tastes are assumed to be endogenous and the specification of their formation has led to the dynamization of the utility and demand functions (Gorman 1967; Peston 1967; Pollak 1970; Houthakker and Taylor 1970; Phlips 1971, 1972; von Weizsdcker 1971; Taylor 1973). In the second, consumer preferences are assumed interdependent, either explicitly through the incorporation of social and cultural propensities in the consumer's choice calculus (Duesenberry 1949; Leibenstein 1950; Clower 1951–52; Johnson 1952) or implicitly through the introduction of Veblnesque arguments (prices) into the utility function (Kalman 1968; Allingham and Morishima 1973). More recently, attempts have been made to synthesize these two approaches (Gaertner 1973; Krelle 1973) within a system of difference equations. Although the contributions listed above are important, they are based on the assumption that the economic actor possesses unlimited psychological stamina and computational capacity and that he operates in a frictionless environment. However, the flow of information is neither full nor free (Stigler 1961); seeking, collecting, and processing information about rapidly changing markets is costly. Furthermore, all human activities, including consumer choice, are subject to the most basic of all constraints time (Becker 1965; Linder 1970; Venieris et al. 1973). These considerations had been anticipated by Simon (1955, 1956, 1959), who argued that consumer behavior is best described by a model of adaptive behavior where ". . . this adaptiveness falls far short of the ideal of maximizing postulated in economic theory. Evidently, organisms adopt well enough to satisfice; they do not, in general, optimize . . . it may be useful, therefore, to ask: How simple a set of choice mechanisms can we postulate and still obtain the gross features of observed adaptive choice behavior?" (1956). As a social environment becomes even more complex and fluid, the individual is more likely to adopt simple, yet effective, choice mechanism(s) or process(es) to reduce the cost and uncertainty of his attempt to adapt rationally to his environment. Such a mechanism should be capable of helping him in solving choice problems by providing workable solutions (Simon 1955). After all, the cost of decision making should not exceed the net value of the object of choice.

In what follows we attempt to establish the existence of such a simple-choice mechanism or, as we shall call it, "low-cost" heuristics. In doing so, we shall relax the neoclassical assumptions of exogenous tastes and independent preference relations. Moreover, we shall investigate the implications of our findings for the theory of consumer behavior.

2 Life-Styles as Low-Cost Heuristics and Reference Group Taking

As individualistic as the economic actor may be, he, nevertheless, occupies a certain place in the social system. He is continuously under social pressure to behave in accordance with a complex system of reciprocal role expectations and obligations, and he is subject to social sanctions. His values, as Knight has pointed out (1923), reflect the cultural tradition to which he belongs. And this tradition – with its social, psychological, and economic dimensions – provides a useful and orderly way of defining the range of options involved in a variety of choice problems (Festinger 1957). It is this cultural tradition, and the social interdependence through which it affects the economic actor, that offers a source of low-cost heuristics. More specifically, the social environment of the economic actor has attributed a positive moral value to achievement and success. The structure of differentiated roles culminates in a system of pecuniary rewards and social prestige which results in efforts of the economic actor to seek status in these terms and emulate.[1] These efforts are largely expressed through consumption behavior since it is one of the more expedient and unambiguous ways of demonstrating one's success and, therefore, moral claim to social prestige.[2]

To the extent that individuals normally follow or avoid the consumption patterns of those surrounding them (because they represent certain classes or groups with which they want to identify), it follows that the insights of Veblen and Pigou can be most effectively abstracted for economic analysis in terms of a theory of interdependent choice via reference groups. For it is the social group, not the individual, that possesses properties which are helpful in solving choice problems by providing low-cost heuristics.

A consumer has a number of reference groups to which he can orient his behavior. The accumulated evidence suggests that his behavior is affected by others as long as they are representative of those social groups to which he

[1]This concept, of course, is not new; nor is it the product of modern social psychology. Veblen (1899) argued that pecuniary emulation and the concepts of conspicuous consumption and leisure are important factors which shape consumer behavior. Similarly, Pigou (1913) recognized the quest for social acceptance or distinction and formalized the concept of consumption externalities. More recently, Morgenstern (1948) studied the same functions and implied the possibility of coalitions among consumers. Leibenstein (1950), abstracting from this possibility, incorporated such externalities as bandwagon, snob, and Veblen effects into neoclassical demand theory and derived demand schedules for the cases when such externalities are present among individual consumers. Kalman (1968) studied the Veblenesque concept of pecuniary emulation and conspicuous consumption by introducing prices as additional arguments in the utility function. However, some of his results were erroneous and recently were corrected by Allingham and Morishima (1973).

[2]Actually, as the reader will have the opportunity to verify, our hypothesis is more general than this sentence seems to imply. It does not have to be limited to individuals in a Veblenesque or a Pigovian emulative system and applies equally well to individuals in any social system as long as their behavior is guided through orientation to particular reference groups.

wants to orient his behavior (Newcomb et al. 1965; Charters and Newcomb 1968; Festinger 1968; Hyman et al. 1968; Kelly 1968). The members of these groups are what social psychologists have called the *significant others* (Miller 1963), and are usually friends, neighbors or associates. The individual, using one or more criteria, can order these groups in terms of social status. Moreover, he reacts positively to some groups and adversely to others by varying the intensity of his reaction from one reference group to another. More formally, the consumer has a reaction function designed to measure the nature and the degree of his reaction to such ordered reference groups. Two points should be made with regard to this reaction function. First, reaction functions may differ between individuals; but, since the same fundamental psychological, economic, sociological, and communication principles operate within a given society, the general characteristic of emulation or aversion, and, therefore, the general shape of the reaction function is similar for most individuals. Second, the shape of an individual's reaction function is the result of not only his endemic traits but also a number of other factors.[3]

Wants are not distributed randomly throughout society but rather in clusters associated with social groups. In turn, social groups have their life-styles, and wants of their members are clustered to define these styles of life.[4] Heuristically, there are islands of clustered wants for different social groups. Social interaction with significant others provides an opportunity for learning about these clusters. It is also possible that consumers acquire knowledge not by one want at a time but rather as a collection of wants as they appear related through life-style activities.

In effect, the concept of life-style introduces a parallel between complementarities among commodities and particular life-styles. Some clusters of goods are more suited than others to serving a particular life-style. A social group with a certain life-style can be assumed to have borne the cost of evaluating the efficiency of various consumption technologies and the usefulness of the associated clusters of

---

[3]In this connection we may mention, for example, (a) his psychological motivation to seek status, (b) the social repercussions he experienced (or he thinks that he will experience) in terms of positive (negative) sanctions as a result of his conformity (deviation) from the norms of his or similar social groups, (c) the amount of useful factual information he possesses concerning different consumption patterns or choices, (d) the economic cost of deviating from the life-style of his social group as well as that of collecting and processing new information.

[4]Life-style is a catchall term. It is used to denote the way one disposes his leisure $_{\text{time,}}$ income, and wealth. Given sufficient time and consistency, "Objective circumstances create distinctive cultural likeness among persons similarly situated . . ." (Mayer 1955, p. 43). Among the factors which cause "cultural likeness" is the inclination for people in given economic situations to confine their social relationships to others of similar back- ground and wherewithal. Indeed, although life-styles may depend on various arguments such as the extent of education, profession, etc., factor analysis has indicated that the level of consumer's income appears to be among the more important ones (Roach et al. 1969). The authenticity and importance of these results have been repeatedly verified by a plethora of sociological and psychological studies. Earlier, for example, Hollingshead (1949) found that only one Elmtown adolescent in every four claimed a "best friend" from a class level different than his own, and Kahl (1957) argued that persons of similar prestige associate by far more with one another and that out of this interaction emerge subcultures peculiar to each prestige level.

goods. Suppose that a social group has existed for some time and its members are in continual contact with one another. They learn from experiences of one another, and they share common ideas, standards, information, education, and income. Simple social learning sets off an elementary process of incremental adjustment of life-style and of the consumption technology that mediates among properties, activities, and goods. Through these adjustments, a clearer definition of that life-style with an efficient cluster of goods will evolve over time. This cluster of goods will be called "complementary" because of the interconnections established by virtue of life-style activities and consumption technology.

Key points of the above discussion on consumer choice are: (1) a consumer is faced with imperfect conditions which prohibit the functioning of global rationality; (2) in order to cut costs, risk, and uncertainty in solving choice problems under limiting conditions, one relies on "heuristics"; (3) logical sources of such heuristics are social interdependence and cultural tradition of society; (4) social interdependence can be more effectively abstracted in terms of "interdependence via reference groups"; (5) a consumer, as a function of his social status, orients his behavior to a number of social groups with a specific reaction function; (6) there is a fairly objective ranking of social groups in terms of social status; (7) each social group has its own life-style (which is defined by a cluster of complementary wants with an efficient technology and an associated efficient cluster of goods serving the life-style); and (8) such clusters of wants are acquired by emulating consumers through social interaction with their significant others. The remainder of this chapter explores the content and significance of such a theory of consumer behavior. Although the main body of the analysis proceeds within a static framework of reference, some dynamic implications will also be considered.

3 Axioms of the Life-Style Hypothesis

We begin with the following two fundamental axioms:

Axiom 1. *There exist social groups with distinct life-styles and associated clusters of wants.*

Axiom 2. *A consumer, as a function of his social status, identifies himself with and emulates a social group as his reference group.*

Given these two axioms, we first introduce a set of assumptions abstracting various aspects of this type of consumer behavior. Then, logical implications of such assumptions in terms of exact characterization of the underlying preference structure are pursued in theorems in a rather heuristic manner.[5]

[5]Rigorous proofs for some of the theorems are provided in a mathematical appendix available from the authors on request. For a mathematical treatment and further results, see Hayakawa and Venieris (1974).

The notion of life-style is best dealt with in terms of characteristics. Therefore, direct objects of consumption are here assumed to be characteristics rather than goods themselves.[6] Because complementarity is an essential ingredient of life-style, that of a social group is represented by a ray emanating from the origin with desired proportions of characteristics defining the life-style.[7] We then assume as the core of our life-style hypothesis that in the consumption set of a consumer there is an optimal ray representing the life-style of an emulated social group. All remaining assumptions are related to this fundamental notion.[8]

Assumption 1 *In the consumption set (i.e., the space of characteristics) of a consumer, there is an optimal ray that he wishes to maintain.*

In this analysis, any life-style is represented by a ray from the origin with corresponding desired proportions of characteristics. Suppose that any ray from the origin in the characteristics space represents a certain life-style. Then, we have a 1–1 correspondence between rays from the origin and life-styles. The next assumption is that if the consumer is constrained to stay on a given ray (i.e., if he is to live a given life-style), the farther he is from the origin (i.e., the more intensively he lives the life-style), the better off he is.[9]

Assumption 2 *Given two intensities of a given life-style, the more intensive is strictly preferred to the less intensive.*

The next assumption asserts that if a consumer shifts his consumption from one bundle of characteristics to another, and if by so doing he is made worse of, then the direction of the shift is not desirable in the sense that no matter how much further he shifts his consumption in the same direction, he is worse off than at the second, less preferred bundle. Under continuity this assumption is equivalent to the assumption of convexity normally assumed in orthodox theory.

---

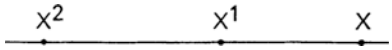
[6]Lancaster touches on the possibility of developing an alternative utility theory along this type of formulation, as he writes: ". . . We may note that the shape of the equilibrium efficiency frontier and the existence of the efficiency substitution effect can result in demand conditions with the traditionally assumed properties, even if the traditional, smooth, convex utility function does not exist. In particular, a simple utility function in which characteristics are consumed in constant proportions—the proportions perhaps changing with income—can be substituted for the conventional utility function" (1966, p. 152).

[7]The life-style of a social group and the related complementarity among characteristics and goods are derived from group norms in the distribution of consumption patterns of its members. Although such norms cannot be represented completely by a ray from the origin, we assume that there is a heavy concentration of consumption patterns along a relevant portion of one.

[8]A consumer may change his life-style over the long run as his social status changes, but in the short run he maintains a particular life-style. So, the optimal ray in his consumption set can be regarded as given for static analysis.

[9]This may be a strong assumption, because a consumer may not wish to live any life-style other than what he considers to be optimal. Nonetheless, it is plausible as an assumption.

Fig. 3.1 Directionality



Assumption 3a (Directionality). *Given consumption bundles of characteristics,* x^1 *and* x^2*, if* x^2 *is preferred to* x^1*, then any bundle x on an extended line segment through* x^1 *is less preferred to* x^1 (see Fig. 3.1).

Although it lacks intuitive plausibility, the following stronger condition may be assumed in place of Assumption 3a. Again, this assumption is, under continuity, equivalent to strong convexity.

Assumption 3b (Strong directionality). *Given two bundles of characteristics,* x^1 *and* x^2*, if* x^2 *is at least as good as* x^1*, then any bundle x on an extended line segment through* x^1 *is less preferred to* x^2*.*

Another assumption follows from the existence of positive sanctions for conformity to and negative sanctions for deviations from group norms. Group pressure is an important element of emulation which enters into the utility function of the consumer negatively. It is also an increasing function of deviation from group norms. Therefore, we should expect that the consumer is made worse off the more he deviates from his optimal life-style. This is stated as follows:

Assumption 4 *Given an intensity of the optimal life-style, the farther the consumer deviates vertically, that is, in the fastest possible manner, from it, the worse off he becomes.*

For further development, slice the space of characteristics by a two-dimensional linear space, and denote this slice by Π. The optimal ray representing the optimal life-style (to be denoted by R^* hereafter) divides this slice into two parts, Π_1 and Π_2. The assumption and theorems presented below are stated for only one of these two parts, Π_1, with the understanding that exactly the same rationale applies to the other part, Π_2, as well.

We further assume that for a given intensity of the optimal life-style (i.e., for a given point on the ray), there is a sufficiently different one for which any intensity would fail to make the consumer equally well off. As implied by Assumption 4, the further one deviates from an optimal life-style, the harder it becomes to compensate him for the welfare loss. If the deviation becomes excessive, then there is little way of compensating for welfare loss.

Assumption 5 *On* Π_1*, for an arbitrary intensity* x^0 *of the optimal life-style* R^**, there exists a completely unacceptable life-style R, or one for which any intensity is less preferred to* x^0*.*

Finally, the social group that the consumer emulates has a cluster of well-defined wants, and we may assume that the group has already borne costs of evaluating various combinations of characteristics in relation to the optimal life-style. Hence,

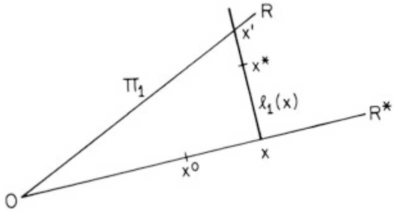


Fig. 3.2 The existence of x^* on $l_1(x)$ such that x^* is indifferent to x^0

for this discussion, it is assumed that for any intensity of an arbitrarily given life-style there is an intensity of the optimal life-style which is indifferent (equally preferred) to it.

Assumption 6 *For any intensity of an arbitrarily given life-style, there is an intensity of the optimal life-style which is indifferent to it.*

Assumptions 1–6 exhaust the assumptions in our life-style hypothesis. The consumer acquires a preference system from emulating the social group. By Assumption 6, such a preference system is in fact regular (i.e., reflexive, transitive, and complete). In addition to these assumptions, let it be assumed that the consumer's preference system is continuous. In the remainder of this chapter, four theorems are presented for a closer characterization of such a preference structure.

The first theorem states that if the consumer is initially living the optimal life-style R^* at a given intensity, then for any fastest deviation from the optimal life-style at greater intensity there is a unique life-style (other than the optimal one) with a unique intensity which gives him the same welfare level as the initial one. More precisely,

Theorem 1 *On Π_1, for an arbitrary intensity x^0 of the optimal life-style R^*, there exists on $l_1(x)$, a unique intensity x^* of a unique life-style (other than the optimal one) such that x^* is indifferent to x^0, where $l_1(x)$ is a line on Π_1 vertical to R^* at a point x on R^* more intensive than x^0* (see Fig. 3.2).

Assumption 2 states that for a given life-style the more intensive is preferred to the less intensive. Hence in Fig. 3.2 x is preferred to x^0. On the other hand, by Assumption 5 there exists a completely unacceptable life-style R. Therefore, x' is less preferred to x^0. Then, because of the continuity property, there must exist a point x^* between x' and x on $l_1(x)$, which is indifferent to x^0.

Suppose an arbitrary intensity x^0 is given on the consumer's optimal life-style R^*. Then let $S(x^0)$ be the set of angles θ_R (with respect to R^*) of all such life-styles R on Π_1 that contain an intensity preferred to x^0. More explicitly:

$$S\left(x^0\right) = \left\{ \theta_R \,\middle|\, \begin{array}{l} R \text{ is in } \Pi_1, \text{ and contains an intensity } x \\ \text{which is prefered to } x^0. \end{array} \right\}.$$

This set is nonempty because the optimal life-style has an intensity preferred to the given intensity x^0 of itself. It is also bounded from above by Assumption 5 (existence of a completely unacceptable life-style). Therefore, a least upper bound of the set exists. Let it be denoted by $\theta*$ and a life-style with the angular deviation of $\theta*$ by $R(x^0)$. Our next theorem refers to the inferiority of any life style on Π_1 whose angular deviation from $R*$ is greater than or equal to $\theta*$.

Theorem 2 *Suppose an arbitrary intensity x^0 of the optimal life-style $R*$ is given. Then, any life-style on Π_1 whose deviation is greater than or equal to $\theta*$ is less preferred to x^0 (i.e., any intensity of it is less preferred to x^0).*

By transitivity it is obvious that if there is a completely unacceptable life-style in relation to some given intensity x^0 of the optimal life-style, then any life-style whose deviation from the optimal one is greater is less preferred to x^0 Therefore, it suffices to show that the life-style $R(x^0)$ is less preferred to x^0. Suppose that $R(x^0)$ contains an intensity x' preferred or indifferent to the given intensity x^0 of the optimal life-style. If x' is preferred to x^0, then by continuity there must be a life-style with deviation greater than $R(x^0)$ which has an intensity preferred to x^0 in the neighborhood of x'. Obviously this contradicts the assumption that $\theta*$ is a least upper bound of the set $S(x^0)$. On the other hand, if x' is indifferent to x^0, then by Assumption 2 there is a higher intensity of the life-style $R(x^0)$ which is preferred to x'. Hence, an analogous argument can be applied to obtain a similar contradiction. Thus, the life-style $R(x^0)$ is less preferred to the given position x^0.

Theorem 1 showed that on Π_1, for an arbitrary intensity x^0 of the optimal life-style, there is a unique intensity $x*$ of a unique life-style on $l_1(x)$, where x is on the optimal life-style and more intensive than x^0. This theorem holds for Π_2. By shifting $l_1(x)$ and $l_2(x)$ along the optimal life- style $R*$ we can trace a two-dimensional slice of an indifference hypersurface through x^0. Denote an indifference hypersurface through x^0 by $I(x^0)$, a two-dimensional slice of it through Π by $I_\Pi(x^0)$, and that part of $I_\Pi(x^0)$, which lies on Π_1 by $I_{\Pi_1}\left(x^0\right)$. Then, Theorem 2 says that $I_{\Pi_1}\left(x^0\right)$ never crosses the life-style $R(x^0)$. In fact, we can show that $I_{\Pi_1}\left(x^0\right)$ is asymptotic to $R(x^0)$.

Theorem 3 *In $I_{\Pi_1}\left(x^0\right)$ (that part of a two-dimensional slice of $I(x^0)$ through Π that lies on Π_1) is convex and asymptotic to $R(x^0)$.*

Because of Assumptions 2 and 3 (a, b) and Theorem 2, the shape of $I_{\Pi_1}\left(x^0\right)$ is rather constrained. It must be convex under continuity, must not cross $R(x^0)$, must not be asymptotic to a life-style R with $\theta_R < \theta_{R\left(x^0\right)}$, and must not bend back. Moreover, the distance between $I_{\Pi_1}\left(x^0\right)$ and $R(x^0)$ must not approach a finite positive value. Therefore, $I_{\Pi_1}\left(x^0\right)$ must be convex and asymptotic to the life-style $R(x^0)$.

In Theorem 2, for an arbitrary intensity x^0 of the optimal life-style, the set $S(x^0)$ was defined as consisting of all angular deviations of such life-styles in Π_1 that possess an intensity preferred to x^0. And, a least upper bound of $S(x^0)$ was denoted by $\theta*$ and a life-style on Π_1 with deviation $\theta*$ by $R(x^0)$. We can regard $\theta*$ as a function which maps intensities of the optimal life-style into the real line.

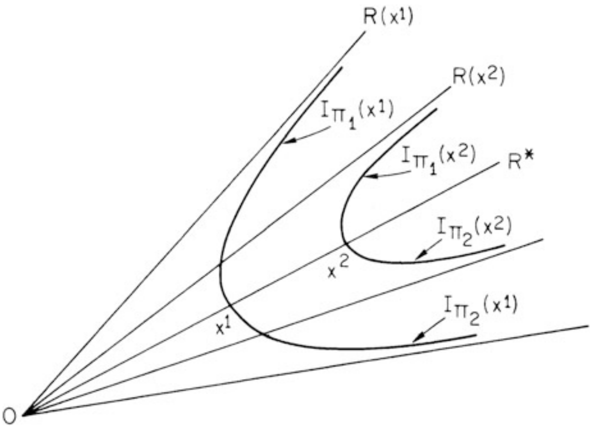


Fig. 3.3 Two two-dimensional slices of indifference hypersurfaces through x^1 and x^2 through $R*$

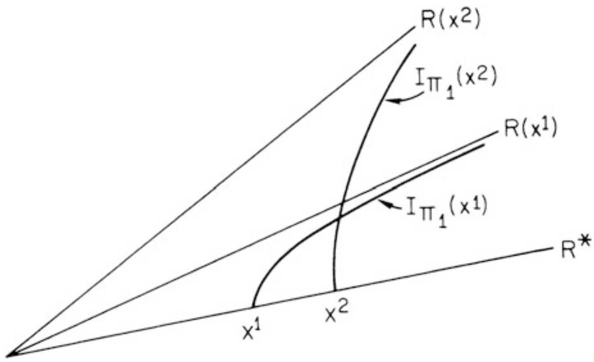


Fig. 3.4 A contradiction that would result if $\theta*(\ )$ were an increasing function

It is now defined by $\theta^*(x) = \sup S(x) = \theta_{R(x)}$, where x is an intensity of the optimal life-style $R*$. Our next theorem states that this function is a nonincreasing function. Figure 3.3 demonstrates two indifference curves when the function is a decreasing kind.

Theorem 4 *On* Π_1, *the function* $\theta*(\ )$ *defined above is a nonincreasing function of intensities of the optimal life-style* $R*$. *That is, for two arbitrary intensities,* x^1 *and* x^2, *of the optimal life-style* $R*$ *with* x^2 *more intensive than* x^1, $\theta*(x^1)$ *is greater than or equal to* $\theta*(x^2)$ (see Fig. 3.3).

This follows directly from transitivity and Assumption 2. If the function were an increasing kind, then $\theta_{R(x^2)}$ must be greater than $\theta_{R(x^1)}$. Being asymptotic to $R(x^1)$ and $R(x^2)$, the indifference curves $I_{\Pi_1}\left(x^1\right)$ and $I_{\Pi_1}\left(x^2\right)$ would intersect as in Fig. 3.4. But, via transitivity, this is a direct violation of Assumption 2. Therefore, the function must be a nonincreasing kind (see Fig. 3.4).

x^1. Furthermore, we can show that the function $\theta*(\ )$ is continuous.

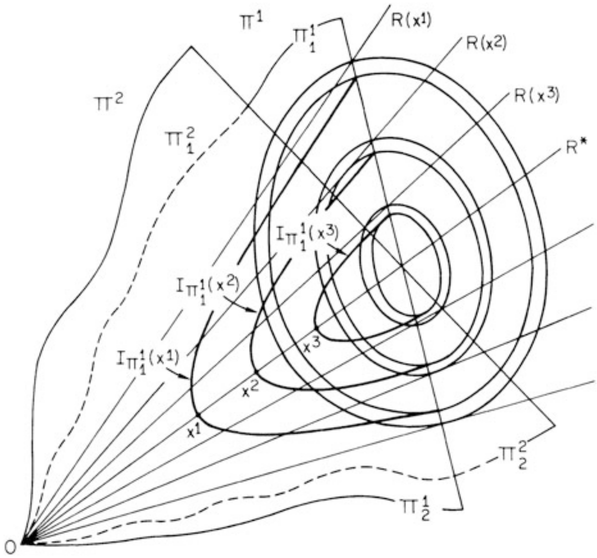


Fig. 3.5 A three-dimensional analog of the *n*-dimensional indifference map

Theorem 5 *The function $\theta^*(\ )$ is a continuous real valued function of intensities of the optimal life-style.*

An underlying reason for Assumption 4 was that group pressure, as a negative argument of the utility function, rises as the deviation from group norms increases. If we further assume that this group pressure is an increasing function of the distance between a bundle of characteristics x and its projection x' on the optimal life-style [where $x \in l_1(x')$], then it is evident that $\theta^*(\ )$ becomes a continuous decreasing function. This is stated separately as a proposition.

Proposition A *If group pressure is an increasing function of the distance of a characteristics bundle from its projection on the optimal life-style R^*, then $\theta^*(\ )$ is a "decreasing" function of intensities of the optimal life-style.*

Thus, we have shown that, under the assumption of Proposition A, the function $\theta^*(\ )$ is a continuous decreasing function of an intensity of the optimal life-style, and that $I_{\Pi_1}(x)$ is asymptotic to $R(x)$. By applying the same analysis to Π_2, we obtain the whole two-dimensional slice of the indifference map which lies on Π. Through repeating the same analysis for all possible Π, the entire, n-dimensional indifference map can be obtained. Because of the directionality and the continuity assumption, such an indifference map is convex. Figure 3.5 is a three-dimensional analog of such an n-dimensional indifference map.

4 Implications and Extensions

We have examined the behavior of a consumer who identifies with and emulates a chosen reference group. To the extent that this hypothesized behavior applies to a substantial number of economic actors, several implications can be drawn for consumer theory. These conclusions, along with suggestions for further research in this area, are presented below.

First, consider the static model in which the consumer is assumed to have chosen a reference group and its associated consumption technology. The resultant consumer preference map, drawn in the characteristics space, will exhibit smooth indifference curves which are convex to the origin. Although the relation between the characteristics space and the commodity space is not a direct one, to the extent that the former can be mapped into the latter without destroying convexity this property is consistent with and indicates the robustness of neoclassical theory. Furthermore, the preference map contains a "relevant range" over which marginal rates of substitution are positive and diminishing and, as prices change, both income and substitution effects operate to affect the consumer's equilibrium response.

However, in its static form our hypothesis does add some information with regard to the general shape of a "typical" preference map. More specifically, the assumption concerning the existence of a completely unacceptable life-style limits the extent of substitutability among characteristics. Moreover, Proposition A implies that the relevant range shrinks as the consumer moves out along the ray R^*.[10] In addition, the income effect of a price change becomes progressively more dominant (vis-a-vis the substitution effect) as income rises.[11] This result implies that stability of general equilibrium within the economic system will decrease (Negishi 1962). This loss must be weighed against the importance of the efficiency substitution effect as commodity space expands along marginally differentiated products to satisfy consumer's desire for variety. This is so, because the life-style hypothesis implies

[10]Note, however, that this characteristic of the indifference map does not necessarily imply an income elasticity of demand close to unity, since even if R^* ray is linear in the characteristic space the corresponding ray in the commodity space need not be. Further- more, as an individual's income rises sufficiently, his R^* ray will rotate to a new position as he abandons his earlier reference group to associate with a higher status group.

[11]In making this observation, it should be noted that we are viewing commodities in a rather broad sense. For any given set of characteristics, one could undoubtedly find a large number of narrowly defined goods which would be highly substitutable; but for broadly defined commodity groups different market baskets will generally entail distinct characteristic mixes. Thus, for broadly defined commodities, the correspondence between characteristic and commodity spaces is more precise and our conclusions are more plausible.

an increased efficiency substitution effect (as opposed to the private substitution effect), and the efficiency choice satisfies the weak axiom of revealed preference (Lancaster 1966, pp. 156–57).[12]

Second, welfare economics is based on the premise that a consumer's preference system as defined on commodity space satisfies the monotonicity assumption. However, according to our life-style hypothesis, a consumer's preference system as defined on characteristics space satisfies monotonicity only in the vicinity of the optimal life-style as an indifference hypersurface eventually approaches asymptotically to a cone with vertex at the origin as in Fig. 3.5. If monotonicity does not hold in the entire characteristics space, then a translation of such a preference relation into commodity space does not satisfy the same property. Thus, the life-style hypothesis renders the monotonicity assumption for consumer preferences a dubious proposition. We can no longer simply assume that "more" necessarily implies "at least as good as".

Third, complementarity in the theory of consumer behavior is seldom mentioned except in the case of clear-cut technical complementarity (such as between bodies of cars and tires). This is partly due to the axiom of substitutability and partly due to the assumption that consumer tastes are personal. But once the notion of a reference group and that of life- style are introduced into consumer choice calculus, we need another type of complementarity, psychological complementarity.

To be sure, whereas classical technical complementarity is treated as an opposite case of substitutability, psychological complementarity is not completely separate from substitutability. To the extent that psychological complementarity emerges from consumption technologies associated with particular life-styles, it is indeed technical. But it would be wise to keep separate from the classical definition. Indeed, psychological complementarity is no less important than its technical counterpart.

So far our discussion of the life-style hypothesis was confined within static grounds. The social group that a consumer emulates was given. It was suggested, however, that as social status changes the reference group will also change in some orderly fashion. The effect of such a change on one's preference system will then be to change the position of his optimal life-style ray as he seeks a new complementarity among characteristics and goods defined by group norms of his new reference group. Consider, for example, the case of a consumer whose income rises by such an amount that he now claims better housing facilities in a more prestigious area. His higher income gives him sufficient means for desiring a new identification, but once he identifies with a new group and chooses to live in a new house, his preference system goes through larger changes. This is partly because of

---

[12]To be precise, this statement has to be modified. For global stability of general equilibrium it is sufficient for Samuelson's weak axiom of revealed preference to be satisfied for the entire economy. But, this condition may not follow even if the axiom is satisfied by individuals.

his new aspirations and partly because of social pressure for conformity to a new identity group. He now demands other attributes in appropriate proportions to his residence. In turn, these proportions will be translated into adequate quantities of goods through his life-style activities. In fact, nothing seems to be more rational for him than to rely on the preferences and consumption technology of his new reference group to solve newly challenged problems in the face of an imperfect environment.

Although we have not formally developed the potential dynamic implication of our hypothesis, it nevertheless appears that our treatment departs substantially from those based on the traditional assumption of "interdependence via individuals" (Gaertner 1973; Krelle 1973). It also suggests that further developments in this area are possible. If a consumer is forced to depart from his optimal ray because of a change in relative prices, will his R* ray eventually rotate toward his equilibrium characteristics mix by altering his choice of a reference group? If so, this would add another dimension to the neoclassical theory of price response and would formally introduce a mechanism through which a change in relative prices institutes a short- and a long-run effect.

Second, we may ask how the exhibited characteristic mix of a particular reference group is determined. If it is affected by the same economic, sociological, and demographic variables which determine the individual's choice of a reference group, then we might expect to observe yet another element of the dynamic path of adjustment to changes in market conditions. As relative prices (or income, or age, etc.) change, and members of the reference group are forced away from their traditional characteristic mix, will they interact in a way that redefines the associated consumption patterns?

Third, we may ask whether the consumer's R^* ray (and thus his full preference map) is susceptible to overt manipulation through information flows. Industrial organization literature contains a running debate over the welfare implications of large advertising outlays. Defenders of current advertising practices typically point to the value of market information in achieving greater welfare. But even a casual glimpse of advertising strategies seems to reveal an effort to appeal to the attractiveness of a reference group rather than to the inherent qualities of the product in question. It is not difficult to recognize that this advertising strategy is followed by various firms in the cigarette and beer industry. If advertising affects consumers' R^* rays, the change in welfare will partially depend upon whether these rays are rotated toward or away from individuals' current consumption equilibrium (abstracting from the problem of externalities). Or, to the extent that public policymakers desire to increase welfare by changing the allocation of resources (for example, to decrease external social costs), it may be possible to affect individuals' consumption behavior by popularizing certain reference groups and therefore R^* rays.

Acknowledgments In writing this paper many friends, colleagues and ex-professors have helped us to articulate our thinking by offering penetrating comments, suggestions, and criticism. In particular, we gratefully acknowledge our debt to Professors W.H. Locke Anderson, M. Manove, P. Ray, S. Winter, and L. Taylor, all of the University of Michigan. We also thank Professors

G. Debreu of the University of California, Berkeley, P. Simpson of the University of Oregon, F. Sebold and D. Stewart of San Diego State University. Finally, we thank Professor H. Johnson and our unknown referee for their penetrating comments. Needless to say, any remaining errors reflect only our limitations.

Addendum: Afterthought[13]

Understanding the way prudential reason works in human decision making has been split between two conceptual poles, i.e., between the pole of *homo economicus* on the one hand, that views individuals as autonomous agents of rational goal orientation dictated by the principle of instrumental rationality, and the pole of *homo* sociologicus on the other, which sees individuals as voluntary agents of socio-cultural orientation, whose dispositions are acquired through socialization and introjection of common normative values. The first relies on the principle of upward causation, starting with the principles that apply universally to all individuals at the bottom and moving upward to the emergence of socio-cultural norms at the top through voluntary exchanges, while the second consists in the principle of downward causation, starting with the presence of socio-cultural norms and institutionalized values at the top and moving downward to the behavior of individuals at the bottom largely determined by such norms and values. While there has been an increasing dialogue between the two approaches in recent years, the gap between them seems to be widening rather than narrowing, despite an obvious fact that human agents are all socialized agents, whose prudence, described by Kant (1785, 4: 416), in the narrowest sense, as the "skill in the choice of means to one's own greatest well-being", dictates the choice of means for the purpose of well-being, or, more generally, whose practical wisdom, understood in the highest sense of Aristotle's *phronesis* in *Nicomachean Ethics*, governs how to live one's life well. The truth must lie somewhere in between, between complete voluntaristic individualism that is blind to cultural symbolism and institutionalized dispositions, and the structural determinism of social and cultural forces that shape individual agents' behavior along social and cultural norms (Hodgson 1986).

Two strains of thought are particularly pertinent to this issue. One strain, known under the rubric of institutional economics, is focused on the evolutionary nature of socio-economic institutions. The thought has been expressed in many different forms, but the core idea dates back to what Adam Smith expounded in *The Theory of Moral Sentiments* (1759). A similar theme was pursued by Thorstein Veblen in *The Theory of the Leisure Class* (1925) and by Pierre Bourdieu in *The Logic of Practice* (1990) and *Distinction* (1984). Smith, Veblen, Bourdieu, and many others along similar veins draw on the essence of socio-cultural evolution, which bifurcates the

[13]This addendum has been newly written by Hiroaki Hayakawa for this book chapter and reflects his afterthought.

social structure into upper and lower classes characterized by different lifestyles and codes of decorum, and through which emulative behavior emerges that constantly seeks upper status identification by relying on the lifestyles of social groups or classes as an effective guide to cultural consumption and profits. Parsons' (1951) functional theory of a social system and institutionalization of common normative values provides a comprehensive framework for placing these evolutional theories in perspective, through a systematic analysis of all essential ingredients necessary for a society to exist as a system.

The other strain, under the name of bounded rationality, focuses on the procedure of problem solving in the face of complexity and uncertainty in the decision making environment. Simon (1955, 1956, 1959) and Cyert and March (1963) argued forcibly that many of the decision making rules adopted by agents whose cognitive and computational capacities are limited are quite different from what the perfect rationality of the mainstream economics requires. Simon's call for procedural rationality (1978) has been echoed by the economics of limited cognition, which has scrutinized the rationality of behavior that looks for economizing modes of choice (e.g., Conlisk 1988; Day and Pingle 1991; Pingle 1992; Pingle and Day 1996). The 84th Dahlem Workshop on Bounded Rationality: The Adaptive Toolbox (1999) was only one of many successful attempts to capture human behavior along these lines. The economics of bounded rationality is now coupled with the experimental economics that owes much of its original insight to Vernon Smith, as well as with psychology and economics that has bloomed after Kahneman and Tversky's contributions (see, e.g., Kahneman and Tversky 1979; Kahneman et al. 1990, 1991; Tversky and Kahneman 1991; Rabin and Thaler 2001; Koszegi and Rabin 2004; Laibson 1997; Ainslie 1991). All of these developments have brought human agents home by asking how they actually behave rather than assuming that they are supposedly rational in the most stringent sense of the term. If a consumer's choice is embedded in a cultural symbolism of emulation and avoidance and has a reference point with respect to where he stands in the social status ladder, this choice must be expressed with respect to this symbolism and status identification. Adam Smith says that the confounding of satisfaction (in the sense of utility), beauty (in the sense of the arrangement), and order (in the sense of the harmony of the economy) is a deception that "rouses and keeps in motion the industry of mankind"; it is the source of all sorts of innovations that embellish our life and push the frontiers of sciences and arts; and this deception is part of Providence that underlies the extensive development of an economic order (1969, pp. 263–264). Smith's insight becomes so much more important when economics as a science of an economic order is combined with how the human mind perceives and expresses under whatever limitations it faces.

At the time this paper was written, all of these developments were yet to come, but the question of how consumers, being bounded in their rationality, make socio-economic choices by referring to the heuristics that can be found in the lifestyles of reference groups was worth pursuing. The paper stipulated a certain set of axioms that capture our lifestyle hypothesis, and showed how an indifference map of a certain shape can be constructed from such axioms. It demonstrated that such norm-oriented consumers can have a convex indifference map, hence are capable of making socially meaningful rational choices. But, this rationality, unlike the instrumental rationality of *homo economicus*, is a socio-economic principle that takes into account life-styles and other socio-cultural norms, not only as low-cost heuristics to otherwise complex problems but also as an effective means of expressing one's identity with a social reference group. The idea was later extended to a more general case in which a consumer is oriented to multiple reference groups and their life-styles (Hayakawa 2000). Simon (1978), referring to the concept of the rational man in economics as a perfect utility maximizer and writing on the trade between economics and other sister social sciences, had this to say:

> It is this concept of rationality that is economics' main export commodity in its trade with the other social sciences. It is no novelty in those sciences to propose that people behave rationally – if that term is taken in its broader dictionary sense. Assumptions of rationality are essential components of virtually all the sociological, psychological, political, and anthropological theories with which I am familiar. What economics has to export, then, is not rationality, but a very particular and special form of it – the rationality of the utility maximizer, and a pretty smart one at that. But international flows have to be balanced. If the program of this meeting aims at more active intercourse between economics and her sister social sciences, then we must ask not only what economics will export, but also what she will receive in payment. An economist might well be tempted to murmur the lines of the tentmaker: "I wonder often what the Vintners buy – Only half as precious as the stuff they sell."
>
> Simon (1978, p. 2)

What this paper demonstrated was that the socio-economic behavior is indeed rational, but not in the sense that a consumer is endowed with a utility function *ex-ante* which is maximized in order to achieve his end, but rather in the sense that an indifference map can be constructed from the way a consumer emulates the lifestyle of a particular reference group in the site of his social space. We might, therefore, be allowed to say that the notion of substantive rationality, when exported to sociology, received its payment in the form of socio-economic rationality, which confirms that the behavior guided by the lifestyles of reference groups is reasonable, agreeable, and rational in this modified sense.

References

Ainslie G (1991) Derivation of rational economic behavior from hyperbolic discount curves. Am Econ Rev 81(2):334–340

Allingham MG, Morishima M (1973) Veblen effects and portfolio selection. In: Morishima M et al (eds) Theory of demand: real and monetary. Clarendon, Oxford, pp 242–270

Aristotle (1984) Nicomachean ethics (trans: Ross WD), revised: Urmson JO. In: Barnes J The complete works of Aristotle (revised Oxford trans), vol 2. Princeton University Press, Princeton

Becker GSA (1965) Theory of the allocation of time. Econ J 75(299):493–517

Bourdieu P (1984) Distinction: a social critique of the judgement of taste (trans: Nice R). Harvard University Press, Cambridge, MA

Bourdieu P (1990) The logic of practice (trans: Nice R). Stanford University Press, Stanford

Charters WW Jr, Newcomb TM (1968) Some attitudinal effects of experimentally increased salience of a membership group. In: Hyman HH, Singer E (eds) Readings in reference group theory and research. Free Press, New York

Clower RM (1951–52) Professor Duesenberry and traditional theory. Rev Econ Stud 19(3):165–178

Conlisk J (1988) Optimization cost. J Econ Behav Organ 9(3):213–228

Cyert RM, March JG (1963) A behavioral theory of the firm. Prentice Hall, Englewood Cliffs

Day RH, Pingle M (1991) Economizing economizing. In: Frantz R, Singh H, Garber J (eds) Handbook of behavioral economics, vol 2. JAI Press, Greenwich, pp 509–522

Duesenberry JS (1949) Income, saving and the theory of consumer behavior. Harvard University Press, Cambridge, MA

Festinger L (1957) A theory of cognitive dissonance. Row Peterson, Evanston

Festinger L (1968) Experiments in group belongingness. In: Hyman HH, Singer E (eds) Readings in reference group theory and research. Free Press, New York

Friedman M (1962) Price theory: a provisional text. University Chicago Press, Chicago

Gaertner W (1973) A dynamic model of interdependent consumer behavior. Dissertation, University of Bonn

Gigerenzer G, Selten R (eds) (2001) The bounded rationality: the adaptive toolbox: Report of the 84th Dahlem workshop, March 1999, The MIT Press, Cambridge, MA

Gintis H (2009) The bounds of reason: game theory and the unification of the behavioral sciences. Princeton University Press, Princeton

Gorman WM (1967) Tastes, habits and choice. Int Econ Rev 8(2):218–222

Hayakawa H (2000) Bounded rationality, social and cultural norms, and interdependence via reference groups. J Econ Behav Organ 43(1):1–34

Hayakawa H, Venieris YP (1974) Life style: a hypothesis on consumer behavior. Manuscript, San Diego State University

Hayakawa H, Venieris Y (1977) Consumer interdependence via reference groups. J Polit Econ 85(3):599–616

Hodgson G (1986) Behind methodological individualism. Camb J Econ 10(3):211–224

Hodgson G (2004) Institutional economics: from Menger and Veblen and North. In: Davis JB, Marciano A, Runde J (eds) The Elgar companion to economics and philosophy. Edward Elgar, Cheltenham

Hollingshead AB (1949) Elmtown's youth. Wiley, New York

Houthakker HS, Taylor LD (1970) Consumer demand in the United States, 2nd edn. Harvard University Press, Cambridge, MA

Hyman HH, Wright CR, Hopkins TK (1968) Reference groups and the maintenance of changes in attitudes and behavior. In: Hyman HH, Singer E (eds) Readings in reference group theory and research. Free Press, New York

Johnson HG (1952) Effects of income-redistribution of aggregate consumption with interdependence of consumers' preferences. Economica 19(74):131–147

Kahl JA (1957) The American class structure. Rinehart, New York

Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica 47(2):263–291

Kahneman D, Knetsch JL, Thaler RH (1990) Experimental tests of the endowment effect and the Coase theorem. J Pol Econ 98(6):1325–1348

Kalman PJ (1968) Theory of consumer behavior when prices enter the utility function. Econometrica 136(3/4):497–510

Kant I (1998). Groundwork of the metaphysics of morals (trans and ed: Gregor M). Cambridge University Press, Cambridge. First published in 1785

Kelly HH (1968) Two functions of reference groups. In: Hymann HH, Singer E (eds) Readings in reference group theory and research. Free Press, New York

Knight FH (1923) The ethic of competition. Q J Econ 37:379–624

Koszegi B, Rabin M (2004) A model of reference-dependent preferences, working paper E04'337, Department of Economics, University of California, Berkeley

Krelle W (1973) Dynamics of the utility function. In: Hicks JR, Weber W (eds) Carl Menger and the Austrian school of economics. Oxford University Press, New York
Laibson D (1997) Golden eggs and hyperbolic discounting. Q J Econ 112(2):443–477
Lancaster KJ (1966) A new approach to consumer theory. J Pol Econ 74(2):132–157
Leibenstein H (1950) Bandwagon, snob and Veblen effects in the theory of consumers' demand. Q J Econ 64(2):183–207
Linder S (1970) The harried leisure class. Harvard University Press, Cambridge, MA
Marshall A (1920) Principles of economics, 8th edn. Macmillan, London
Mayer KB (1955) Class and society. Doubleday, New York
Miller DR (1963) The study of social relationships: situations, identity and social interaction. In: Kock S (ed) Psychology: a study of a science, vol 5. McGraw-Hill, New York
Morgenstern O (1948) Demand theory reconsidered. Q J Econ 62:165–201
Negishi T (1962) The stability of a competitive economy: a survey article. Econometrica 30(4):635–669
Newcomb TM (1968) Attitude envelopment as a function of reference group: the Bennington study. In: Hyman HH, Singer E (eds) Readings in reference group theory and research. Free Press, New York
Newcomb TM, Turner RH, Converse PE (1965) Social psychology. Holt, Rinehart & Winston, New York
Parsons T (1951) The social system. Free Press, New York
Peston MH (1967) Changing utility functions. In: Shubik M (ed) Essays in mathematical economics in honor of Oskar Morgenstern. Princeton University Press, Princeton
Phlips L (1971) Substitution, complementarity, and the residual variation around dynamic demand equations. Am Econ Rev 61(4):586–597
Phlips L (1972) A dynamic version of the linear expenditure model. Rev Econ Stat 54(4):450–458
Pigou AC (1913) The interdependence of different sources of demand and supply in the market. Econ J 23:19–24
Pingle M (1992) Costly optimization: an experiment. J Econ Behav Organ 17(1):3–30
Pingle M, Day RH (1996) Models of economizing behaivior: experimental evidence. J Econ Behav Organ 29(2):191–209
Pollak RA (1970) Habit formation and dynamic demand functions. J Pol Econ 4:745–769
Rabin M, Thaler RH (2001) Risk aversion. J Econ Perspect 15(1):219–232
Roach JL, Gross L, Gursslin O (1969) Social stratification in the United States. Prentice-Hall, Englewood Cliffs
Simon HA (1955) A behavioral model of rational choice. Q J Econ 69(1):99–118
Simon HA (1956) Rational choice and the structure of the environment. Psychol Rev 63(2):129–138
Simon HA (1959) Theories of decision-making in economics and behavioral sciences. Am Econ Rev 49(3):253–283
Simon HA (1978) Rationality as process and as product of thought. Am Econ Rev 68(2):1–16
Smith A (1969) The theory of moral sentiments. Arlington House, New York. First published in 1759
Stigler G (1961) The economics of information. J Polit Econ 59(3):213–225
Taylor LD (1973) On the distinction between habit formation and stock adjustment in demand analysis. Unpublished paper, University of Michigan
Tversky A, Kahneman D (1991) Loss aversion in riskless choice: a reference-dependent model. Q J Econ 106(4):1039–1061
Veblen T (1899) The theory of the leisure class. Macmillan, New York
Venieris YP, Sebold FD, Harper RD (1973) The impact of economic, technological, and demographic factors on aggregate births. Rev Econ Stat 55(4):493–497
von Weizsäcker CC (1971) Notes on endogenous hange of tastes. J Econ Theory 3(4):345–372

Chapter 4
Bounded Rationality, Social and Cultural Norms, and Interdependence via Reference Groups

Hiroaki Hayakawa

Abstract This chapter presents an axiomatic theory of preference orderings similar in abstraction to the standard preference model, but developed for the analysis of bounded rationality where social norms and reference groups serve as sources of low-cost heuristics that can absorb costly deliberation and other limits to rationality. A two-step choice process is proposed, in which physical wants are satisfied sequentially with social want reducing choice indeterminacy. Social want is captured by emulation–avoidance of reference groups, and the serviceabilities of a commodity bundle to it are imputed. The resulting norm-guided behavior is rationalizable in the traditional sense.

Keywords Bounded rationality • Reference groups • Social norms • Social want • Life styles

1 An Overview

Whether individual decision makers are *homo economicus* guided by independent preferences or *homo sociologicus* guided by social norms has been a matter of defining concerns to both economists and sociologists. In broad terms, most economists have maintained that human behavior, whether in isolation or embedded in social situations, can be explained, with proper modifications if necessary, within the rational choice framework where preferences are treated as given data. Sociologists, on the other hand, have largely maintained that it is social norms and

The original article first appeared in the Journal of Economic Behavior and Organization 43; 1–34, 2000. A newly written addendum has been added to this book chapter.

H. Hayakawa (✉)
University of Brunei Darussalam, Bandar Seri Begawan, Negara Brunei Darussalam
e-mail: hiroaki.hayakawa@ubd.edu.bn; hayakawah2@r4.dion.ne.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_4

order that shape preferences of individuals and give meaning to their actions. In recent years, the crossing between the two fields has intensified (Coleman 1990; Baron and Hannan 1994; Smelser and Swedberg 1994; Bowles 1998). Among many questions that have been addressed by both camps is the fundamental one: why do individuals allow their behaviors to be molded by social and cultural norms if they are (assumed to be) capable of making decisions without the help of such norms? In retrospect, the Lucas critique that individuals' decision rules are constantly altered to reflect the structure of the decision-making environment cannot have been aloof to this question, since human behavior in a culturally and socially structured environment is of no exception to this critique (Lucas 1976).

In dealing with social interactions, economists normally start with preferences that are given and represented by a utility function. Adding social factors to the list of the arguments of this function produces a variety of externalities that have been explored in various models. Thus, economists have been answering the fundamental question essentially by saying, rather circularly, that individuals have preferences over social interactions. Most sociologists would defy such an approach for the very reason that the existence of a fixed mapping from social and other factors into utility would imply that social norms are, after an infinite regress, reduced to preferences of individuals (Hodgson 1986). To many of them, social order is something that cannot be reduced to self-interested or voluntary actions of individuals (e.g., Smelser 1990; Tilly 1991), much less to fixed individualistic preferences, although such a position has met serious challenges from the rational choice approach to sociology (Coleman 1990, 1994).[1]

A very promising approach to the fundamental question has been suggested by the economics of bounded rationality *à la* Simon (1955, 1959) and by the recent development of the economics of limited cognition that dates back to, e.g., Cyert and March (1963) (see also Conlisk 1988; Day and Pingle 1991). The insight that decision makers economize on economizing models is not new (e.g., Schumpeter 1934; Alchian 1950), and the idea of habit was an important concept among Western social theorists (Camic 1986; see also Day 1987 for the notion of *habere*). Now the profession recognizes: (1) the decision-making environment including the internal psychology and the cognitive capacity of a decision maker may be significantly short of being perfect; (2) because the time endowment is fixed, every activity including cognition competes for the use of time; (3) information on which decisions are based is almost always incomplete, but information gathering and processing is costly; and (4) many decision making situations are imbued with elements of risk and uncertainty.

If these limitations are serious enough, a decision maker's motivations would apply not only to choice objects but also to how to handle the limitations without

[1]One movement which has been gaining momentum in the field of sociology is what is now called the rational choice sociology, which has been led by Coleman. This approach seeks to explain the nature, the origin, and the reproduction of the social order as consequences of voluntary actions of rational individuals (see Coleman 1990, 1994).

sacrificing too much of his resources. Therefore, what kind of a problem-solving procedure (or mode) is actually adopted to save cognitive efforts and to absorb bounded rationality in general becomes a matter of no trivial importance. In Simon's words, it is the *procedural rationality* that matters in understanding human behavior, not the *substantive rationality* that has pervaded the orthodox economics (Simon 1978).

Thus, it has been suggested that bounded rationality may account for the adoption of simple modes of behavior that take advantage of social and cultural norms and readily available heuristic solutions to complex problems. The implications of such possibilities are far reaching, and we shall explore some of them in this study, particularly ones that relate to the formation of social norm-guided endogenous preferences. If decision makers resort to such modes, there will be non-negligible rigidities in human behavior. Such rigidities often produce adaptive economizing behavior, which may become important sources of dynamic forces that drive the economic system as a whole (Day 1984, 1986; Conlisk 1996).

In this regard, Day (1984, 1986) calls attention to the fact that adaptive economizing by way of simple computable steps (analogous to optimizing algorithms in mathematics) activates two types of mechanisms: (1) information, planning and control mechanisms in disequilibrium, and (2) inventory-order-price adjustment mechanisms that mediate transactions in disequilibrium. In such modeling of behavior, economic units are seen to exercise only local rationality or local search within *zones of flexible responses* in a sequence of recursively connected programs. In his view, agents' behavior is governed by an information-planning-control system under mental-physiological structures of specialized functions. It is not difficult to understand why agents in such a system might turn to economization of economizing itself. According to Day (1984, 1987), seven basic modes are capable of describing economizing choices: (1) obedience to an authority, (2) imitation of others' modes, (3) habit (unconscious repetition of past behavior), (4) unmotivated search, (5) hunch, (6) experimentation (trial and error), and (7) procedural optimizing (see also Pingle and Day 1996 for an explicit statement). The effectiveness of these modes has been examined experimentally by Pingle (1992) and Pingle and Day (1996) (see also Day 1967; Day and Tinney 1968; Shipley 1974; Naish 1993 on the near optimality and the convergence of adaptive decisions; Conlisk 1980 on the coexistence of costly optimizers and cheap imitators; Pingle 1992 for an experiment of costly optimization). There is no reason for some of these economizing modes not to draw on social and cultural norms as sources of low-cost heuristics to complex problems and as instruments by which to endow human action with social meaning. The mode of imitating what *significant others* (Miller 1963) are doing is a good example. Such a mode is expedient because it saves cognitive efforts; at the same time it is socially meaningful because the act of emulating and avoiding certain reference groups by watching what *significant others* are doing is driven by social motives to seek upper status identification.

Bounded rationality and the ensuing adaptive behavior holds promises to answer Hodgson's call for a new framework of thinking beyond the *methodological individualism* (Hodgson 1986). Examining the validity of this methodology and Hayek's

(1952) *composite method* that runs through much of contemporary economics, he argues that in understanding human behavior there is no need to fall into the trap of complete voluntaristic individualism nor into the trap of the structural determinism. What is needed is a theoretical framework that allows human behavior to be molded by social and cultural norms while retaining the autonomy of individual decision makers. If bounded-rationality motivates individuals to orient themselves to social and cultural norms, their goals and preferences will be molded or guided by such norms, but the purposefulness of their actions will remain intact.

Simple modes of behavior suggested by Day above are similar to Beckert's (1996) devices of uncertainty reduction. Expanding on Granovetter's (1985) notion of the *embeddedness* of economic action in situational structures and Simon's (1976) *intended and bounded rationality* that underlies administrative behavior, Beckert has scrutinized the rationality principle in economics and its limitations. He has asked: if the dominant feature of situational structures in which economic actions are embedded is *uncertainty* so that the means–end relations on which economic calculations are based are lost, is it not the case that some mechanisms external to the decision maker (be they social and cultural norms, institutions, and social relations) are called in to reduce the choice set of decision makers and to restore *certainty* in the means–end relations? Because uncertainty, as distinguished from risk by Knight (1921), is not reducible to calculable probabilities, it imposes formidable limitations to the rationality of actions in situations characterized as such. If the means–end relations are lost, agents simply cannot decide what is best to do. It is, therefore, argued that the rational choice theory of economics cannot be a reasonable theory about human behavior in circumstances that are conditioned by uncertainty and that this rationality should be replaced with a more practical one that transcends the dichotomy between rational and irrational behavior. Beckert's argument is reminiscent of Radner's (1975) point over two decades ago that *rationality* does not allow an easy definition once cognitive capacities are recognized as limited.

As an alternative to the objective rationality, Beckert introduces the notion of *intentional rationality* as one that relies on simple devices as instruments of uncertainty reduction when all means–end relationships break down. Such devices are comprised of (1) tradition, habit, and routines, (2) norms and institutions, (3) structural predispositions of decisions such as social networks, organizational structures, and past decisions, and (4) power relations (Beckert 1996, pp. 827–829). The social devices build up rigidities in human behavior, thereby causing it to be adaptive, and more importantly, to be *predictable*. Beckert's argument that making use of these devices narrows the choice set of decision makers and make actions predictable shares much with Heiner's (1983, 1989) insight that the boundedness of the decision-making environment is an important source of predictable behavior as decision makers adopt more inflexible decision rules. It is analogous to Simon's notion of *intended and bounded rationality* (that takes the form of satisficing behavior), which, according to him, forms the theoretical basis of administrative behavior (Simon 1976).

Thus, the recognition that human behavior is boundedly rational and that it is embedded in social norms and order has brought to light possible linkages between the adoption of simple modes of behavior and persistent orientation to such norms. Such linkages have made it an important agendum to look into the possibility that human behavior may be better modeled by considering not only the need for simple modes of behavior that can absorb bounded rationality but also the non-functional benefits from orientation to social and cultural norms.

Many views have been put forth on such linkages. For our purposes, the following views are particularly important: (1) The society takes on a bigger meaning than a mere aggregation of the parts, and it should serve as reliable sources of meaningful simplification devices or heuristics to otherwise complex problems. (2) Decision makers will resort to simple modes of behavior and look for heuristic solutions in order to save their cognitive efforts and to absorb other limiting elements of the decision-making environment. (3) If preferences are composite of various needs which are not necessarily commensurate, they may be prioritized and satisfied sequentially with switching from one need to the next being effectuated as soon as the aspiration level is reached. (4) There are social and cultural norms (social institutions, customs, sanctions, cultural values, etc.) which motivate individuals to behave in ways that endow their actions with social and cultural meaning. To the extent that preferences beyond physical needs are closely related to social and cultural norms, such norms may account for the origin of norm-influenced preferences for social interdependence. (5) More strongly, the formation of ends or preferences itself reflects the desire to act in a socially meaningful fashion when there are serious limits to the objective rationality. Social and cultural norms are, therefore, not just sources of external influences on human behavior, but rather they endow the decision-making environment with a social and cultural structure under which socially meaningful preferences and actions are actively formed and planned to reproduce the structure itself. (6) In acting under a socially structured environment, individuals will be able to exercise only local rationality within their social zones of flexible responses (in Day's terms), which are determined to a large extent by the history of their past emulation and avoidance efforts and on which social, psychological, economic, communicational, and other principles are working.

In this chapter, we propose a theory of choice behavior that takes into account these views, answering by so doing Simon's call for procedural rationality as well as Hodgson's call for norm-oriented purposeful behavior. If individual decision makers are in need of simple devices to solve complex problems, such heuristics must be readily available in the society, and there must be, at the same time, a process by which they are continually sustained and renewed. The view we expound is that low-cost heuristics to complex choice problems are sought and found in life styles of social groups. These life styles constitute a form of social capital that is accumulated collectively by members of social groups through their error-learning processes. If the cost of problem solving is too excessive for individuals to bear, it makes sense to invest in this capital collectively with all parties sharing the cost of the required investment. The benefits of the accumulated consumption know-how are then shared by members of social groups. Thus, with accumulation of such capital, the task of selecting the best from the set of feasible alternatives is reduced to an expedient act of referring to what has been tested and approved by members of social groups.

What is so distinctly important about the life styles of social groups is that they are, in Granovetter's terms, embedded in a cultural-value system so that the act of orienting to them becomes a socially meaningful cultural behavior. It is doubtful that these life styles would ever be developed if identifying with them had no social and cultural value to begin with. The society, therefore, can be viewed as a *culturally directed social field* (analogous to a magnetic field in physics) in which the life styles of social groups exist as norms of consumption behavior and in which individuals sense the direction for higher status identification. This social field gives rise to *social want* as a culturally directed social predisposition to orient oneself to relevant social groups (and their life styles) in the process of seeking higher social status levels. Viewed as such, social want is crucially dependent on one's location in the social field and on his zone of flexible responses, however this zone may be determined.

The rest of the chapter is organized as follows: in Sect. 2, we discuss the nature of bounded rationality and argue that there are genuine needs for simple devices or low-cost heuristics to complex problems. In Sect. 3, we identify those social and cultural arrangements that shape decision makers' orientation to social groups. These arrangements will be abstracted by way of *interdependence* via *reference groups*, in which social status ranking of such groups gives direction to the emulation–avoidance motives. In Sect. 4, holding that physical and social wants are not commensurate generally, we consider a sequential satisficing decision rule over physical wants as an alternative to the usual utility theory. In Sect. 5, we discuss properties of a preference relational system that follows from this sequential rule and propose a two-step procedural model of consumer choice where social want appears in the second step as an instrument of indeterminacy reduction. In Sect. 6, we present a formal model of interdependence via reference groups and define what social want constitutes. We show that a reaction function over relevant social groups and information on the whereabouts of social norms together make it possible to quantify the social want-satisfying property; this quantification eliminates or substantially reduces the indeterminacy of choice that remains in the first step. We show that a norm-oriented choice behavior can be rationalized by a norm-guided ordering of choice objects. In Sect. 7, we relate our model of interdependence to the work of Duesenberry (1949) and Leibenstein (1950), and Sect. 8 concludes the chapter.

2 Bounded Rationality and the Need for Low-Cost Heuristics

It is now widely recognized that the decision-making environment including the internal psychology of decision makers is short of being perfect. Various elements can account for such imperfection. (1) Decision makers' cognitive and computational capacities are significantly bounded (Simon 1955, 1959). (2) The severity of this limitation is compounded by the fact that the time endowment is fixed, so that all activities (including cognitive ones) compete for the use of this endowment (Becker 1965; Linder 1970). (3) Decision makers seldom have perfect information about choice alternatives, but information gathering and processing, like any other activity, is costly in time and other resources (Stigler 1961). (4) Many decision-making situations involve elements of risks, so that the anticipated consequences of decisions can only be assessed in probability terms, subjective or objective. (5) If decision-making situations are imbued with uncertainty so that the means–end relationships that are necessary for economic calculations break down, the consequences of an action cannot be assessed even in probabilistic terms (Keynes 1921; Knight 1921). The total failure of the objective rationality in the face of uncertainty leaves a gap that is beyond the cognitive capacity of decision makers (Beckert 1996). While various models or apparatuses have been developed to cope with some of these limitations, we focus here on decision makers' motives to save the resources that would otherwise be required to solve complex choice problems.

In our view, the key to the success of a decision maker in coping with the limitations of the decision-making environment is the availability of cost-absorbing choice mechanisms (or devices) that are rooted in social capital and order. For sure, such mechanisms have to be simple enough to afford a painless adaptation to the decision-making environment; at the same time they should reflect the social structure of this environment. Specifically, what is needed is a socially meaningful choice mechanism that absorbs risk and uncertainty, reflects costs of acquiring and processing information, reduces the pressure of time constraints, and eases the computational pain of problem solving. If one's cognitive capacity is limited and if there are limits to decision making costs that can be borne, a resulting decision mechanism will be of a simple kind.

Assume, for a moment, that there are neuron locations in human brains to process information and that two kinds of information occupy such locations: (a) decision rules and transformation, and (b) data about choice objects and internal states of a decision maker. Naturally, an economizing problem arises over such locations. The more complex are the problem-solving algorithms, the less room is available for data, and the more complete are the data, the less room is available for complex algorithms. The total computational capacity has an upper limit that is short of global optimization. Given costs of obtaining data and given time constraints on how long is permitted to solve a problem, further limits are placed on the complexity of problem solving. After all, the costs of decision making should not exceed the net value of choices made.

Once costs of reaching a decision are taken into account, the need for a simple-choice mechanism that relies on low-cost heuristics cannot be ignored. Given social capital that has been accumulated in the form of life styles and given social and cultural order that has transformed the environment into a well-directed social field (so that socially desirable ends and means can be identified), it is only natural for individuals to search for low-cost heuristics in the life styles of their relevant social groups.

Viewing the consumer choice process as one of utilizing low-cost heuristics is not an escape from the conventional rationality hypothesis. Rather, it is best viewed as an extension of this hypothesis when the decision-making environment is imperfect. The central issue is still one of cutting costs of problem solving as we focus on the use of socially desirable, cost-saving means (see Vriend 1996 for a much broader interpretation of rationality as pursuance of self-interest). What is novel of this view is that it interprets the life styles of social groups as social capital that has been accumulated through error-learning processes and relates this capital to human behavior of bounded rationality.

3 Social Capital as Sources of Low-Cost Heuristics

Choice decisions made by an individual depend crucially on the organization of his perceptions of choice objects. To the extent that such perceptions are affected by social and cultural elements, they cannot be independent of a particular social and cultural environment in which decisions are made. An individual participates in the economy not simply as an economic abstract with idiosyncratic tastes but as a whole person with a variety of legitimate social and cultural concerns and motivations that are very much part of his economic choices. For this reason, it is important to deal with the full complexity of choice behavior.

An individual occupies certain positions and plays certain roles in a society in which he lives. He also takes part in the life-worlds of social groups through many activities. This entails that his behavior reflects a web of reciprocal role expectations and obligations as well as social sanctions (positive or negative) of various kinds and strengths (see Parsons and Shils 1951 for the notion of *double contingency* of social action or order). It, therefore, follows that human behavior is, to a large extent, a social behavior. Culture then provides a framework of value orientation shared by the majority of the society, without which reciprocal expectations on the responses of different individuals would break down, and coherent perceptions of choice objects in the context of a social environment would not emerge. Among many functions that society performs, we pay attention to its function as the provider of socially desirable heuristics to complex choice problems through clusters of activities that mediate life styles of social groups and through commonly shared cultural values that make adoption of such heuristics a socially desirable act to follow. Thus, the totality of life styles developed by social groups can be viewed as social capital, and a commonly shared cultural value orientation that guides the use of this capital can be viewed as social order.

There are at least four aspects to this capital: economic, psychological, social, and cultural. If a decision maker in a social setting makes use of this capital, we need to know, on top of his usual budget constraint, his psychological motives, the social sanctions (pressures for conformity and sanctions against deviations) that are at work on his choices, and the cultural values shared by members of his society. These aspects are not independent of one another as, e.g., the success-oriented cultural values will give rise to the psychological motives to gain upper status identification as a moral claim to success.

Economists often say that choice theory is or should be independent (in explanatory power) of choice objects. It is the same regardless of what individuals want. It is, therefore, often concluded that a decision maker's motivations are irrelevant to choice theory. This would be the case if it were true that his motivations applied only to the goods bundle's constituent list of commodities. However, the matter of motivation may apply to the safety of a choice as of neighborhood or brand of car, to the status gains, to the costs of handling uncertainty, to the appropriateness of a chosen list of goods to life style activities, or to the handing of high information costs. To the extent that these motivations may play important roles in decision making, they should not be swept aside and branded as irrelevant.

To summarize, social and cultural arrangements of the decision-making environment affect an individual's organization of his perceptions of choice objects, therefore, his choice behavior. The society is not simply a collection of isolated individuals acting on their own idiosyncratic preferences with little interdependence among themselves. It is best viewed as a coherent whole organized around social norms and cultural values that constitute social capital and order of powerful economic significance. An individual's behavior is then part of his entire living in this totality of socio-economic realities.

With this general view, we note that the modern society places high moral values on achievement and success. Moreover, the structure of differentiated social roles and positions is integrated around a system of pecuniary rewards and social prestige based on these values. Therefore, it makes up a very significant part of an individual's motivation to be regarded as a winner and to be respected as worthwhile in his society. Such motivation normally manifests itself in status seeking and emulation of higher status groups. It seizes upon consumption behavior for an obvious reason that it is the best, impersonal way of demonstrating to the society the extent of one's success as a moral claim to the social prestige.

Such an idea is not new in economics. Veblen (1899), writing on the theme of human proclivity to emulation, expounded the notion of pecuniary emulation and conspicuous consumption as a means to an invidious comparison. He writes:

> The accepted standard of expenditure in the community or in the class to which a person belongs largely determines what his standard of living will be. It does this directly by commending itself to his common sense as right and good, through his habitually contemplating it and assimilating the scheme of life in which it belongs; but it does so also indirectly as a matter of propriety, under pain of disesteem and ostracism. To accept and praise the standard of living which is vogue is both agreeable and expedient, commonly to the point of being indispensable to personal comfort and to success in life. The standard of living of any class, so far as concerns the element of conspicuous waste, is commonly as high as the earning capacity of the class will permit — with a constant tendency to go higher. . . . . . (Veblen 1899, pp. 111–112)

Pigou was also among the first to point out the importance of this achievement-oriented moral value as reflected in consumers' constant quest for reputation and distinction bearing goods. He writes:

> …… the essential matter is that people do, in fact, desire many things, not merely for their own sake, but, in the main, on account of the reputation or distinction which the possession of them confers ……. The quantity of a distinction-bearing article that anyone demands at a given price depends, not merely on the price, but also on the extent to which it is "the thing" to buy that article, and thus, indirectly upon the quantity that people in general are buying ……
>
> …… In fact, however, distinction is usually to be found, not in being in the swim in general, nor yet in being out of the swim in general, but in a combination of resemblance to certain persons and of difference from certain other persons. …… Furthermore, both among the persons whom a man wishes to resemble, and among those from whom he wishes to separate himself, some are usually much more important to him than others. …… (Pigou 1913, pp. 20–24)

Thus, given the human proclivity to emulation and avoidance and given a social status scale that has evolved over time with success-oriented moral values, people of higher or equal social statuses tend to have positive effects and those of lower statuses negative effects on one's consumption behavior. Since an individual emulates or avoids behavior of other individuals to the extent they are representative of the social classes or groups of their orientation, Veblen and Pigou's insight can be best abstracted in terms of *interdependence* via *reference groups* (Hayakawa and Venieris 1977).

More precisely, an individual orients his behavior to a number of his reference groups. People in them constitute what social psychologist Miller (1963) calls *significant others*; they are usually friends, neighbors, or associates. Some of them exert positive effects and others negative effects. Both among those groups to be emulated and among those to be avoided, some have stronger positive or negative effects than others. Moreover, under a commonly shared cultural value orientation, it would be possible to rank these groups in terms of their social statuses. It is conceivable that there may be more than one criterion for this ranking. If the ranking of social groups is comprehensive enough, the essential feature of the interdependence via emulation and avoidance of reference groups may be captured by the notion of a reaction function defined on a set of well-ordered social reference groups.

This formalization of social interdependence is to be distinguished from various models of social interdependence or consumption externalities that have been presented in the literature to this day.[2] Most of such models have remained within

[2] There have been many attempts to capture consumption externalities. Some old examples include: Leibenstein's bandwagon, snob, and Veblen effects (Leibenstein 1950); Duesenberry's relative income hypothesis (Duesenberry 1949; Clower 1951–52; Johnson 1952); Veblenian conspicuous consumption through prices or real income as separate arguments in the utility function (Kalman 1968; Allingham and Morishima 1973; Hayakawa 1976); interdependence via reference groups based on emulation–avoidance motives (Hayakawa and Venieris 1977); dynamic modeling of consumer interdependence by way of a system of interdependent linear difference demand equations (Gaertner 1973; Krelle 1973); Becker's (1974) theory of social interactions in which characteristics of other persons enter the production functions of the basic wants or commodities; Pollak's (1976) model of interdependent preferences via consumption of all individuals in the utility function; Frank's (1984) model of wage differentials based on income hierarchies yielding

the confine of *interdependence* via *individuals* or at the level of Millers' *significant others* without considering social groups in the background. This is understandable in the light of the fact that the *hedonistic creature* of the traditional utilitarian individualism has dominated economic theorizing. But, it is useful to go beyond the individualistic level of social interdependence and to address the fundamental reason for it. In our view, an individual is affected by *significant others* because they are representative of his reference groups targeted for emulation and avoidance along the social-status scale. Furthermore, if the social and cultural proclivities of human behavior are to be linked to bounded rationality of attempting to reduce the cost of complex problem solving in an imperfect decision-making environment, social interdependence ought to be abstracted by a scheme of *interdependence* via *reference groups*. For, after all, it is social groups, not individuals, that possess cost-saving heuristics for complex problems.

Some observations are in order: first, which social groups an individual targets in an invidious system is not random, but is very much guided by other groups' social statuses. Therefore, two individuals of different social statuses will orient their behaviors to different social groups. Yet, to the extent that the same principles (psychological, economic, sociological, and communicational) are likely to be operating within a given society, the general characteristics of emulation and avoidance patterns, hence, the general shape of a reaction function will be similar among most individuals. In fact, it is this fact that supports the notion of social order (an orderly pattern of social orientation) in an invidious system. Moreover, a reaction function is best viewed as a (net) summary of (a) psychological motives to seek higher social statuses, (b) social pressures for conformity to group norms and sanctions against deviations from such norms, (c) the amount of useful factual information about consumption patterns of different groups across society, and (d) the economic cost of gathering information and experimenting something new (at the risk of disturbing the complementarities among activities and goods that mediate particular life styles) (Ray 1973, pp. 284–288).

---

within-group status, and his model (Frank 1985) of the effect of Hirsch's (1976) *positional goods* on the demand for non-positional goods; Granovetter and Soong's (1986) threshold models of interpersonal effects in consumer demand. An interdependence model via reference groups by Hayakawa and Venieris (1977) differs from these other models as it tries to derive indifference curves axiomatically from a life style hypothesis. The notion of social reference groups consisting of *significant others* is also applied in Kapteyn (1980).

More recently, the literature has witnessed a resurgence of interest in social interactions as sources of promising answers to some intriguing questions. Among them are the equity premium puzzle (see Mehra and Prescott 1985; Abel 1990; Gali 1994; Kocherlakota 1996), the convexity of the utility function in the controversy between the New Keynesian and the New Classical economics (Mankiw et al. 1985; Hall 1988; Ikeda 1994), and the concave–convex–concave utility function (Robson 1992; Coelho and McClure 1998). Furthermore, with growing interest in new growth theory, some attempts have been made to relate status seeking and consumption externalities to capital and wealth accumulation, economic growth and dynamics, and economic performance in general (Cole et al. 1992; Fershtman and Weiss 1993; Futagami and Shibata 1995; Hof 1997).

Second, wants are not distributed randomly or evenly throughout the entire society. Most likely they exist in clusters associated with different social groups. To the extent that individuals select certain social groups as their reference groups, social wants of individuals will reflect such clusters. Heuristically, there are islands of clustered wants of different social groups, which may be disjoint or overlapped. These clusters are then learned and acquired by individuals through interactions with *significant others*. In this process, individuals acquire a collection of wants at a time lest their complementarities be lost.

One further observation: the notion of a life style introduces a parallelism between complementarities among goods and particular life styles. Some clusters of goods are better than others in serving a particular life style. In this context, a social group can be thought of as having borne the cost of testing various consumption technologies and associated clusters of goods for their serviceabilities to its life style. Such testing may be understood as follows: suppose that members of a social group are in continual contact with one another, sharing common ideas, values, standards, information, etc., and learning from experiences of one another. Then, *simple learning* will set off an elementary *hill-climbing process* of incremental adjustments of a life style and a cluster of goods serving that life style, and of similar adjustments of consumption technologies that mediate properties, activities, and goods (Day 1967; Day and Tinney 1968). What comes out over time, therefore, is a well-defined life style and an efficient cluster of goods to live that life style. The latter is called *complementary* because of the interconnections among goods that are established by virtue of life style activities and related consumption technologies.

Thus, the society is best viewed as a *social space* in which various styles of living and consumption technologies have been tested and accumulated with their benefits being shared by members of social groups and in which there is movement from one style of living to another as one seeks higher social prestige and status. The accumulated life styles constitute *social capital*, and the cultural value orientation that drives upper social status seeking defines *social order*. Our scheme of social interdependence via reference groups is based crucially on the existence of such capital and order. Faced with an imperfect decision-making environment, it is natural for individuals to turn to social capital as sources of low-cost heuristics that not only save the cost of problem-solving but also meet their social needs, and to turn to social order for orientation of their emulation–avoidance behavior.

The question is, how do the social capital and order help reduce problem-solving costs borne by individuals? In our view, they do so by sending to individuals signals or directions as to which reference groups ought to be emulated and which reference groups ought to be avoided, all as a function of their social statuses (i.e., as a function of their positions in the social field). Once their reference groups are identified, their attention turns to such non-functional attributes of choice objects as how popular they are among members of such groups, hence how the acquisition of them helps achieve a desired social status identification. Each commodity bundle, therefore, comes with a set of non-functional attributes, some of which are based on its popularity across the reference groups. If an individual seeks to emulate a particular group, those goods that are in vogue with that group have, in their vectors of non-functional attributes, a component which gives him a sense of belonging to that group and of having attained the social status of that group. Hence, given a set of socially ranked reference groups and a vector of non-functional attributes of goods, one may be able to contrive a measure that indicates to what extent any given commodity bundle satisfies one's social want by taking the convolution of a reaction function and vectors of non-functional attributes. The social want-satisfying property measured this way makes up an important part of the total serviceabilities of a commodity bundle. Thus, the social capital and order we have identified reduce the burden of complex problem solving by substituting, in place of utilitarian psychologizing, learning of social norms called life styles and by directing individuals to make use of these norms in the quest for higher social status identification.

4 A Sequential Satisficing of Wants

The traditional approach to consumer behavior assumes that rationality is global and that wants are commensurable (i.e., reducible to utility) regardless of their origins. It ignores the possibility that the internal physiology of a decision maker may not allow a common measurement of all different physiological needs. Moreover, as we argue, if the limited cognitive capacity and other limiting conditions are the dominating features of the decision-making environment, individuals may be significantly motivated to economize on economizing by searching for low-cost heuristics to otherwise complex choice problems and to turn to social and cultural norms (i.e., life styles of social groups and a commonly shared system of cultural value orientation) as real sources of these heuristics and direction to guide emulation and avoidance motives. Embedded in a system of cultural value orientation, such motives give rise to some well-structured social want, which, by its very nature, is distinctly different from any of the physical needs. If human wants, physical or social, differ so much in nature to make them practically non-commensurate, it makes sense to model consumer choice as one of assigning priorities to differing wants and satisficing with respect to their relative satiation levels.

For this reason, we take an approach based on Georgescu-Roegen's (1954) hierarchical nature of human wants, Simon's (1955, 1959) principle of satisficing (see also Radner 1975; Radner and Rothschild 1975), and Day's (1987) prioritization of multiple ends. We assume (1) that wants are specific and not commensurable, (2) that they are prioritized and satisfied sequentially under the principle of satisficing with respect to their aspiration levels (relative satiation levels), and (3) that wants of physical nature are bounded by their aspiration levels whereas no such restriction is placed a priori on want of social nature because this want is inherently relative to social norms. The primary physiological needs are assigned higher priorities, to be followed by less basic ones and eventually by want of social nature.

Reviewing the literature, we find that the view that the decision maker's actual choice process is sequential is at least as old as Menger and Jevons. They write:

> An isolated farmer, after a rich harvest, has more than 200 hundred bushels of wheat at his disposal. A portion of this secures him the maintenance of his own and his family's lives until the next harvest, and another portion the preservation of health; a third portion assures him seed-grain for the next seeding; a fourth portion may be employed for the production of beer, whiskey, and other luxuries; a fifth portion may be used for the fattening of his cattle. Several remaining bushels, which he cannot use further for these more important satisfactions, he allots to the feeding of pets in order to make the balance of his grain in some way useful.
>
> We have seen that the efforts of men are directed toward fully satisfying their needs, and where this is impossible, toward satisfying them as completely as possible. If a quantity of goods stands opposite needs of varying importance to men, they will first satisfy, or provide for, those needs whose satisfaction has the greatest importance to them. If there are any goods remaining, they will direct them to the satisfaction of needs that are next in degree of importance to those already satisfied. Any further remainder will be applied consecutively to the satisfaction of needs that come next in degree of importance. (Menger 1950, pp. 129–131)
>
> . . . . . . Nor, when we consider the matter closely, can we say that all portions of the same commodity possess equal utility. Water, for instance, may be roughly described as the most useful of all substances. A quart of water per day has the high utility of saving a person from dying in a most distressing manner. Several gallons a day may possess much utility for such purposes as cooking and washing; but after an adequate supply is secured for these uses, any additional quantity is a matter of comparative indifference. All that we can say, then, is that water, up to a certain quantity, is indispensable; that further quantities will have various degrees of utility; but that beyond a certain quantity the utility sinks gradually to zero; it may even become negative, that is to say, further supplies of the same substance may become inconvenient and hurtful. (Jevons 1957, pp. 43–44)

As seen in these quotations, three principles run through Menger's and Jevons's views on human wants: (1) wants are specific and qualitatively different; (2) in the process of satisfying them, they are prioritized; and (3) they are bounded by their relative satiation levels. In the subsequent development of economists' choice theory, however, it has been assumed that wants of all origins are reducible to a common measurement called utility. This utility theory then has expounded that it is only ordering of choice objects that matters in consumer choice. Thus, the fact that wants are prioritized has become replaced with the principle of the diminishing marginal utility, or, more generally, with that of the diminishing marginal rate of substitution. As we shall show below, a sensible ordering of choice objects is equally possible with non-commensurable and prioritized wants. It is this possibility that legitimizes a two-step choice process we propose below.

A couple of observations are in order. First, it has often been pointed out that human wants are *dynamic* in nature in the sense that the satisfaction of lower, more primary wants awakens higher wants so that wants themselves are destined to grow in number over time. This aspect of wants has been dubbed as *the principle of the subordination of wants* or *the principle of the growth of wants* (see Menger 1950, pp. 82–83; Marshall 1920, pp. 86–91). While wants grow dynamically over time, they may be taken to be finite in number at a given point in time.

Second, one's priority ordering of wants and setting of their aspiration levels are likely to be socially and culturally influenced. Hence, it would not be surprising if individuals who belong to the same society and share an identical cultural value orientation exhibited similar priorities and aspiration levels of wants. Different societies and cultures attach different values to human activities, and individuals' orderings of wants are likely to reflect such social and cultural differences.

Third, the fact that social want is placed at the end of an ordering does not imply that social considerations are of the least importance when one's budgetary resources are too limited to satisfy many of the basic needs. Even in such situations, multiple means are likely to be available, and this multiplicity often calls for social considerations for further guidance. This may account that even in a less affluent society, where individuals are still struggling for the basic needs, social pressures for conformity can coerce individuals to choose, among many alternatives, those that invite less social sanctions. In an affluent society on which our attention is focused, most of the primary needs are satisfied and social considerations occupy the mind of most individuals as they seek constantly higher social status identification. In such a society, social considerations do play a crucial role as an instrument of indeterminacy reduction on the choice set. It is this fact that supports our assumption that the satisficing feasibility set (i.e., the set of those choice objects that are feasible and meet all of the physical needs to their aspiration levels) is non-empty.

5 Sequential Satisficing of Wants

Sequential satisficing of wants reviewed above is based on three premises: (1) All wants, physical or social, are prioritized, and social want appears at the end of this ordering. (2) Each physical want is bounded by its aspiration level whereas no such bound is imposed a priori on social want because this want is determined by social norms. (3) It is social interdependence via reference groups (embedded in a cultural value orientation) that structures social want such that it serves as an instrument of indeterminacy reduction in choice decision making.

In the sequel, we derive an ordering (of choice objects) from the prioritization of differing wants and show that behavior based thereon is perfectly rational. As a matter of our strategy, we first deal with prioritization of physical wants and get an ensuing ordering of choice objects, which is incomplete to the extent that the satisficing feasibility set is non-empty. We then combine this ordering with another ordering based solely on social want considerations. We show that this composite ordering is well defined and that the resulting choice behavior is rationalizable.

Suppose, we have a fixed number of physical wants. Let want i be denoted by ω_i and the property that satisfies want i by x_i. A greater value of ω_i and x_i indicates, respectively, a greater degree of satisfaction of want i and a greater capacity to satisfy want i. We start with the following postulates:

Postulate 1 An individual has a finite number, say m, of physical (functional) wants, denoted $\omega_1, \omega_2, \ldots, \omega_m$. Let $W \equiv \{\omega_1, \omega_2, \ldots, \omega_m\}$. W is individual-specific.

Postulate 2 An individual has a strong ordering R_ω on the set W, i.e., R_ω satisfies:

1. *Transitivity*: $\omega_i R_\omega \omega_j$ and $\omega_j R_\omega \omega_k$ imply $\omega_i R_\omega \omega_k$ for any ω_i, ω_j, and ω_k in W.
2. *Asymmetry*: $\omega_i R_\omega \omega_j$ implies not $\omega_j R_\omega \omega_i$ for any ω_i and ω_j in W.
3. *Completeness*: Either $\omega_i R_\omega \omega_j$ or $\omega_j R_\omega \omega_i$ holds for any ω_i and ω_j in W.

Postulate 3 Each want, ω_i, has its aspiration (or relative satiation) level, denoted ω_i^*. This is equivalent to assuming that the corresponding want-satisfying property (or quality) x_i has its aspiration level x_i^*.

Postulate 2 implies that any subset of W has a unique maximal element. Postulate 3 implies satisficing, i.e., that physical wants are only satisfied to their (relative) satiation levels. These postulates give a sequential satisficing decision rule: an individual first prioritizes physical wants, and satisfies them sequentially, each to its aspiration level. Without loss of generality, let the elements of set W be arranged according to the priority ordering, so that $\omega_i R_\omega \omega_j$ holds if and only if $i < j$.

The space of goods is the non-negative orthant of an n-dimensional Euclidean space, denoted

$$G = \{y : y \in R^n \text{ and } y \geq 0\},$$

and the space of the want-satisfying properties (qualities) is the non-negative orthant of an m-dimensional Euclidean space, denoted

$$X = \{x : x \in R^m \text{ and } x \geq 0\}.$$

Assume that there is a (perceived, hence basically subjective) transformation function $\Phi : G \to X$, which reflects the amount of information that an individual possesses, his cognitive limitations, and the way he perceives the merits of choice objects. Since information and cognition are not free, the perceived transformation reflects the cost of both information gathering and cognition. Given such a transformation function, a commodity bundle y_i is transformed to a bundle of physical want-satisfying properties:

$x^i = \left(x_1^i, x_2^i, \ldots, x_m^i\right) = \Phi\left(y^i\right) \equiv \left[\Phi_1\left(y^i\right), \Phi_2\left(y^i\right), \ldots, \Phi_m\left(y^i\right)\right]$, where $\Phi_j(y^i)$ is the jth component of $\Phi(y^i)$. Let the vector of the aspiration levels of m want-satisfying properties be $x* \equiv \left(x_1^*, x_2^*, \ldots x_m^*\right)$.

The transformation function $\Phi : G \to X$ induces an ordering of objects in G, which we call a preference relation. This relation is more general than the usual one that underlies the traditional utility theory for the reason that it is affected, among other things, by the aspiration levels, experiences, and information in possession, which are not independent of social and psychological predispositions of the decision maker, the cost of cognition, and the cost of information gathering and processing.

Postulate 4 An individual has a preference relation $\succ$ on the space of goods G, which is defined as follows: for any two commodity bundles, y^i and y^j, in G, $y^i \succ y^j$ if and only if any one of the following conditions holds:

1. $x_1^j < x_1^*$ and $x_1^i > x_1^j$.
2. $x_1^i = x_1^j < x_1^*$, $x_2^j < x_2^*$, and $x_2^i > x_2^j$.
3. $x_1^i > x_1^*$, $x_1^j > x_1^*$, $x_2^j < x_2^*$, and $x_2^i > x_2^j$ and so on up to the *m*th want-satisfying property, where

$$x^i \equiv \left(x_1^i, x_2^i, \ldots, x_m^i\right) = \Phi\left(y^i\right) \text{ and } x^j \equiv \left(x_1^j, x_2^j, \ldots, x_m^j\right) = \Phi\left(y^j\right).$$

In determining preferences between any given two commodity bundles, y^i and y^j, an individual first examines the extent to which they satisfy the want of the first priority. If the two bundles satisfy this want either to the same extent or in excess of its aspiration level, then his attention shifts to the want of the second priority. This will be repeated sequentially. A preference relation of this kind is in general not complete for the following reason: to be able to state that either $y^i \succ y^j$ or $y^j \succ y^i$ holds for any two alternatives, one needs some decisive relation at the margin. But, after all preceding ones have failed to make preferences determinate one way or the other, even the last (the *m*th) property may still fail to do so because this property is again satisfied either to the same extent or in excess of its aspiration level. It is this incompleteness or the indeterminacy that, in our view, motivates individuals to seek further guidance in social and cultural norms, which will be discussed in detail in the next section.

The space of goods, G, together with this preference relation, constitutes a relational system denoted $(G, \succ)$. Because $\succ$ is asymmetric and negatively transitive (i.e., not $y^i \succ y^j$ and not $y^j \succ y^k$ imply not $y^i \succ y^k$), this is a weak order system. Also, given $\succ$ on G, an indifference relation, denoted $\sim$, can be defined by the absence of preferences one way or the other. That is, for any two bundles, y^i and y^j, in G, $y^i \sim y^j$ (y^i is indifferent to y^j) if and only if neither one is preferred to the other. Then, combining $\succ$ with $\sim$, we may form a composite relation R_G, on G, defined by $y^i R_G y^j$ if and only if either $y^i \succ y^j$ or $y^i \sim y^j$. For any two bundles in G, it is always the case that either one is preferred to the other or there are no definite preferences between the two. Therefore, the composite relation R_G is reflexive and complete. It is also transitive. Thus, an induced relational system (G, R_G) is a preference-ordering system. Let these results be summarized as follows.

Proposition 1 *A preference relational system* (G, $\succ$), *where* $\succ$ *is defined as in Postulate* 4, *is a weak-order system (*i.e., $\succ$ *is asymmetric and negatively transitive).*

Proposition 2 *An induced relational system (G, R_G), where* $\sim$ *is defined by the absence of definite preferences, is a preference-ordering system (*i.e., R_G *is reflexive, transitive, and complete).*

Moreover, if the relation $\succ$ is weakly complete (i.e., for any two bundles, y^i and y^j $\left(y^i \neq y^j\right)$, in G, either $\mathrm{y}^i \succ \mathrm{y}^j$ or $\mathrm{y}^j \succ \mathrm{y}^i$), the composite relation R_G, defined by $y^i R_G y^j$ if and only if either $\mathrm{y}^i \succ \mathrm{y}^j$ or $\mathrm{y}^i \sim \mathrm{y}^j$, becomes a chain. The only way that $\succ$ becomes weakly complete in the context of a sequentially satisficing decision rule is by being able to come up, at the margin, with some want-satisfying property that makes preferences determinate one way or the other. Notice that by the asymmetry of $\succ$ and by the definition of $\sim$, $y^i R_G y^j$ and $y^j R_G y^i$ must imply that y^i is indifferent to y^j; but if the weak completeness is met, y^i is indifferent to y^j if and only if y^i is identical to y^j. Thus, if the axiom of weak completeness is satisfied, our composite relation R_{G} satisfies the property of antisymmetry (i.e., for any two bundles, y^i and y^j, in G, $y^i R_G y^j$ and $y^j R_G y^i$ imply $y^i = y^j$). This is the case taken up by Georgescu-Roegen (1954) to demonstrate that lexicographic preferences are not measurable.

Proposition 3 *An induced relational system, (G, R_G), of Proposition 2 becomes a chain system (*i.e., *R_G is reflexive, transitive, complete, and antisymmetric) if $\succ$ is weakly complete.*

Now, let $B(P, M)$ be a consumer's budget set corresponding to a price vector P and income M; i.e., $B(P, M) \equiv \{y : y \in G \text{ and } P \cdot y \leq M\}$. Also, let $A(x^*)$ be his satisficing set; i.e., $A(x^*) \equiv \{y : y \in G \text{ and } \Phi(y) \geq x^*\}$ ($\Phi(y) \geq x^*$ means $\Phi_i(y) \geq x_i^*$ for all $i = 1, 2, \ldots, m$).

We next postulate that the intersection of the budget set and the satisficing set is non-empty. Call this intersection *the satisficing feasibility set*. The idea is that basic physical needs are within the feasibility of the budget set.

Postulate 5 A consumer's satisficing feasibility set is non-empty, i.e.,

$$A\left(x^*\right) \cap B\left(P, M\right) \neq \varnothing.$$

Whether the satisficing feasibility set is non-empty or not, or, more importantly, how large this set is, depends on to what extent the aspiration levels of wants are adjusted dynamically when choice decisions are repeated. For instance, depending on the type of want, the aspiration level may be adjusted upward or downward, all according to the degree of easiness or difficulty experienced in day-to-day choices. Overall, however, to the extent that many of the functional wants arise from physical needs, it would not be too unrealistic to assume that they are more or less satisfied. In an affluent society, a typical middle-class individual's total expenditure most likely exceeds what his basic needs require (see Baxter and Moosa 1996 for a basic need hypothesis on consumption behavior).

If the satisficing feasibility set is large, it begs a question as to how to reduce the size of this set and where to turn for effective guidance. We argue that it is social want that serves as an instrument of indeterminacy reduction through a well-directed orientation to social and cultural norms. The fact that social want is distinctly different from physical ones and the fact that the satisficing feasibility set is most likely to be non-empty in an affluent society suggests that a typical individual may be solving his choice problems in two steps. In the first step, the satisficing feasibility set (i.e., the set of all R_G-maximal elements in the budget set) is identified. That is, physical wants are arranged by their priorities so as to have them satisfied sequentially to their aspiration levels. In the second step, his attention shifts to social want, whose structure, combined with the whereabouts of social norms, leads to determinate choices by identifying those objects (in the satisficing feasibility set) that yield the highest social gratification.

A Two-Step Procedural Choice Process In the first step, identify the satisficing feasibility set $A(x^{*}) \cap B(P, M)$; i.e., select all R_G-maximal elements from the budget set $B(P, M)$. In the second step, select those elements of this set that yield the highest satisfaction of social want.

If income is too low to yield a non-empty satisficing feasibility set, the first step will suffice to make determinate decisions. With the rise of income, the attention shifts to higher (less primary) functional wants and eventually to social considerations. As the satisficing feasibility set becomes non-empty, how best to meet social want acquires the status of an important criterion for selection of desirable objects. In an affluent society, such non-functional aspects of decision making cannot be ignored. Decisions are therefore eventually guided by such considerations as how fit choice objects are to present life styles, how effective they are for status identification and seeking, how fashionable they are to current modes of tastes, etc. Thus, social want, as it is embedded in social capital and order, motivates individuals to select their most socially desirable objects with resources that are typically not enough to catch up with ever increasing social needs and expectations.

This two-step procedure should be contrasted with one suggested by Kornai (1971), in which a single element is chosen from the set of eligible alternatives at the final stage of an elementary process with no deterministic decision rules for this selection. Final choices are randomly made with a decision distribution being defined on the set of eligible alternatives. In our model, the criterion of how best to meet social want narrows the choices from the satisficing feasibility set.

6 Interdependence via Reference Groups and Social Want

With the role of social want made explicit, we next turn to the modeling of interdependence via reference groups that endows social want with a useful structure to decision making at the final stage. Since an individual guides his behavior by perceiving, in a bundle of goods, certain properties (or qualities) that contribute to the satisfaction of his social want, we need to measure such properties somehow. Recall that our scheme of social interdependence had three features (Sect. 3): (1) an individual belongs to a social group and takes a number of social groups as his *reference groups* in the process of seeking upper status identification; (2) some of these groups are to be emulated (positive orientation) and others are to be avoided (negative orientation); and (3) these reference groups are ordered in terms of their social statuses. These features were then consolidated into the notion of a reaction function defined on a set of well-ordered reference groups. This scheme associates each bundle of goods with a vector of non-functional attributes derived from its popularity across social groups. Therefore, to what extent a given bundle of goods meets social want can be measured, to a first approximation, by convoluting a reaction function with vectors of such non-functional attributes. To formalize this measurement, we start with the following postulates.

Postulate 6 There are a finite number of social groups in the society,

$$\widehat{S} \equiv \{g_1, g_2, \ldots, g_n\}.$$

An individual is oriented to a subset of these groups,

$$S \equiv \{g_1, g_2, \ldots, g_k\} \subset \widehat{S}.$$

Groups in S are referred to as the individual's *relevant social groups* or *reference groups.* S is individual-specific and depends largely on his position in the social field. The individual belongs to at least one group in S.

Postulate 7 The set of relevant social groups, S, is divided into two disjoint subsets, S_1 and S_2 ($S = S_1 \cup S_2, S_1 \cap S_2 = \varnothing$):

$$S_1 = \{g_i : g_i \in S \text{ and } g_i \text{ is a group of positive orientation}\},$$
$$S_2 = \left\{g_j : g_j \in S \text{ and } g_j \text{ is a group of negative orientation}\right\}.$$

That is, $g_i \in S_1$ is a group to be emulated, and $g_j \in S_2$ is a group to be avoided.

Postulate 8 An individual has orderings, R_1 on S_1 and R_2 on S_2, defined as follows:

1. For any two groups, g_i and g_j, in S_1, $g_i R_1 g_j$ if and only if g_i has emulation effects stronger than or equal to g_j.
2. For any two groups, g_i and g_j, in S_2, $g_i R_2 g_j$ if and only if g_i has avoidance effects stronger than or equal to g_j.

We assume that there is a continuum of social status levels over the range [0, h] (alternatively, one may assume that there are only a finite number of such levels). It is important that there exists a social-status ranking function that is accepted by most members of the society, so that the status disparity between any two groups can be measured with little personal biases.

Postulate 9 There exists a social-status ranking function $r : \widehat{S} \to [0, h]$ which satisfies

1. $r(g_i) > r\left(g_j\right)$ if and only if g_i is higher than g_j in social status, and
2. $r(g_i) = r\left(g_j\right)$ if and only if g_i and g_j are identical in social status, where $\hat{S}$ is the set of all social groups.

This social-status ranking function makes it possible to measure the social status disparity between any two groups.

Definition 1 A social-status disparity function is a real-valued function $d : \widehat{S} \times \widehat{S} \to R$ that associates each pair $\left(g_i, g_j\right) \in \widehat{S} \times \widehat{S}$ with a real number

$$d\left(g_i, g_j\right) = r\left(g_i\right) - r\left(g_j\right).$$

The absolute value $|d(g_i, g_j)|$ is referred to as the *social distance*.

A social-status disparity function is analogous to a distance function (a metric) in mathematics, but it differs in that it is allowed to take both positive and negative values. If the function takes a positive (negative) value for a pair of groups, the first is ranked higher (lower) than the second in social status. If there is little risk of confusion, we shall use *social status disparity* and *social distance* interchangeably.

An individual's orientation to social groups (for emulation or avoidance) is in general limited to a subset $S \subset \widehat{S}$. Without loss of generality, let it be assumed that the elements of S are ordered by the social-status ranking function with the larger subscript indicating the higher social status, so that $r\left(g_i\right) > r\left(g_j\right)$ if and only if $i > j$. We assume that no two groups in S have an identical social status rank. With this convention, g_1 takes the lowest value $r(g_1)$ and g_k the highest value $r(g_k)$ among all $g_i \in S$.

Suppose that group $g_m \in S$ is a group of an individual's current belonging; the case of multiple group belonging is excluded. From the vantage point of g_m, the above disparity function, when restricted to his relevant set S, takes its maximum value at some social group and its minimum value at some other group with all other groups taking their values in between.

Definition 2 Let an individual belong to a group $g_m \in S$. For his relevant set of social groups, S, define

$$\begin{aligned}
\delta_{\min} &\equiv \min\left\{d\left(g_i, g_m\right) : g_i \in S, i = 1, \ldots, k;\ g_m \text{ is fixed}\right\},\\
\delta_{\max} &\equiv \max\left\{d\left(g_i, g_m\right) : g_i \in S, i = 1, \ldots, k;\ g_m \text{ is fixed}\right\}, \text{and}\\
\Omega &\equiv \left\{d\left(g_i, g_m\right) : g_i \in S, i = 1, \ldots, k;\ g_m \text{ is fixed}\right\}.
\end{aligned}$$

With our convention, it holds that $d\left(g_1, g_m\right) = \delta_{\min}$ and $d\left(g_k, g_m\right) = \delta_{\max}$.

To formalize the notion that an individual is oriented positively to some groups and negatively to others, we introduce the notion of a *reaction function* defined on set Ω.[3]

Definition 3 The *reaction function* of an individual who belongs to group g_m is a real-valued function $V : \Omega \to R$ such that

1. $V\left(d\left(g_i, g_m\right)\right) > 0$ if g_i is in S_1,
2. $V\left(d\left(g_i, g_m\right)\right) < 0$ if g_i is in S_2,

[3]We define a reaction function on the social status disparities of relevant social groups rather than directly on the set of such groups, although a reaction function defined as such, together with a social-status disparity function, induces an indirect reaction function on the set of relevant social groups. The reason for this definition is that the extent to which one emulates or avoids a reference group depends largely on how disparate the group is in social status from his own. That is, if a relevant reference group is replaced with another one of a similar status, the group should be emulated or avoided to a similar degree.

3. for any two groups, g_i and g_j, in S_1, $V(d(g_i, g_m)) \geq V\left(d\left(g_j, g_m\right)\right)$ if and only if $g_i R_1 g_j$, and
4. for any two groups, g_i and g_j, in S_2, $V(d(g_i, g_m)) \leq V\left(d\left(g_j, g_m\right)\right)$ if and only if $g_i R_2 g_j$, where S_1 and S_2 are the sets of groups for emulation and avoidance in Postulate 7.

The last two conditions require that the reaction function $V : \Omega \to R$ agree with orderings R_1 and R_2 defined on S_1 and S_2, respectively. Listing the images V(d(g_i, g_m)), $g_i \in S$, gives a reaction vector, whose first and last components are $V(\delta_{\min})$ and $V(\delta_{\max})$, respectively.

Definition 4 A reaction vector, denoted $[V(g_i, g_m)]$, is a vector defined by

$$[V(g_i, g_m)] \equiv [V(d(g_1, g_m)), V(d(g_2, g_m)), \ldots, V(d(g_k, g_m))],$$

where $V(d(g_1, g_m)) = V(\delta_{\min})$ and $V(d(g_k, g_m)) = V(\delta_{\max})$.

An example is useful to illustrate graphically what has been captured by Postulates 6–8 and Definitions 1–4. Consider a case in which an individual belongs to group g_4 and has seven relevant social groups including g_4. In Fig. 4.1, the social-status disparity of each group from g_4 is plotted on the x-axis. Since the disparity of g_4 from itself is zero (i.e., $d(g_4, g_4) = r(g_4) - r(g_4) = 0$), the origin corresponds to the social-status rank of g_4. The individual emulates groups g_3, g_4, g_5, and g_6 and avoids groups g_1, g_2, and g_7. The reaction function $V : \Omega \to R$, therefore, takes positive values at $d(g_3, g_m)$, $d(g_4, g_m)$, $d(g_5, g_m)$, and $d(g_6, g_m)$, and negative values at $d(g_1, g_m)$, $d(g_2, g_m)$, and $d(g_7, g_m)$. The absolute value of $V(d(g_i, g_m))$ measures the intensity of the individual's emulation and avoidance. Capturing the nature of the orientation to social and cultural norms, $V(d(g_i, g_m))$, $i \in S$ (or a reaction vector) characterizes the individual's social want.

The shape of reaction functions is not a matter of personal tastes. Because orientation to social and cultural norms grows out of the need for reciprocal expectations and workable heuristic solutions to otherwise complex problems, it is contrary to the notion of social want itself to assume that such functions are idiosyncratically formed. But, if they are to have a common structure, some principles must be identified that contribute to its formation. At least four such principles can be identified along the psychological, social, communicational, and economic dimensions (Ray 1973, pp. 284–288).

The Psychological Principle Under interdependence via reference groups that is embedded in success-oriented moral values, psychological motives that drive emulation and avoidance should be closely related to the rewards of upward status identification as well as to the threat of lower status identification. The psychic satisfaction from such motives will, most likely, be an increasing function of the gains in social status. Written as a function of the social status disparity defined above, this function will be positive in its positive range and negative in its negative range, but increasing over the entire range. This satisfaction will be subjected to

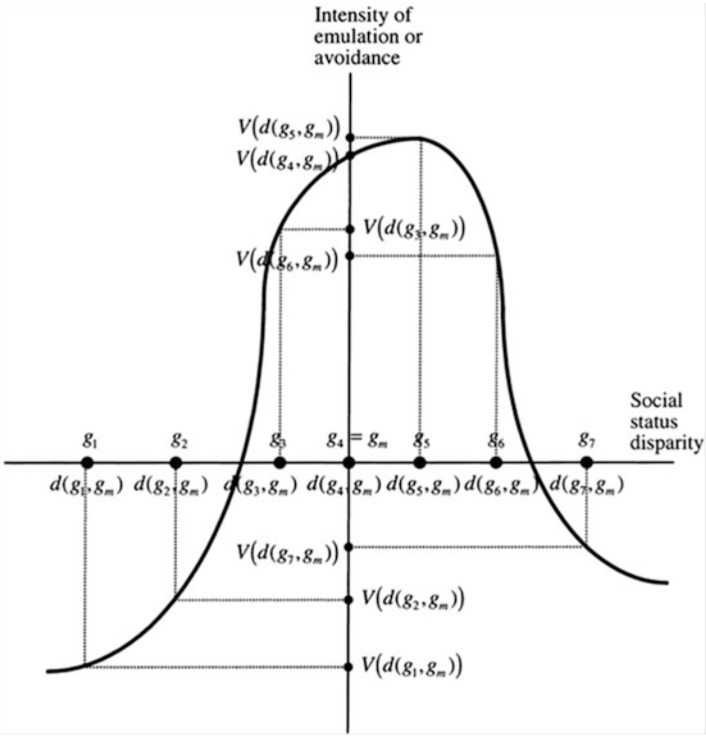


Fig. 4.1 A reaction function: an example

the law of diminishing marginal satisfaction (analogous to the law of diminishing marginal utility).

The Social Principle Group norms tend to sanction positively (favor) those activities that are in vogue among members of own and similar status groups and to sanction negatively (penalize) those that deviate from these activities. Peer group pressures are often very strong and cannot be ignored (e.g., Evans et al. 1992). The fear of a loss of reputation by acting differently from norms (e.g., Akerlof 1980) or the fear of inviting envy from others by doing too well (Mui 1995) can be strong enough to coerce individuals to behave in conformity with group norms. Viewed as a function of the social status disparity, these sanctions will therefore be positive around zero disparity but will fall to negative values in a bell-shaped fashion as the social distance increases in either direction.

The Communicational Principle The more distant (in social status) are the groups that are being emulated or avoided, the more difficult it becomes to obtain useful factual information about their life styles. People who belong to identical or similar social groups tend to communicate more often through socializing activities, and knowing more about what *significant others* are doing makes emulation and avoidance easier and more effective. Therefore, the quantification of this principle as a function of the social status disparity will be analogous to a density function of a normal distribution in probability.

The Economic Principle A life style is mediated by a certain set of complementarity relationships among goods and services. Hence, trying something new always runs the risk of threatening to break some of these relationships. To the extent that the disparity in life style becomes more pronounced with the social distance, such costs will rise as a function of this distance. Furthermore, successful emulation and avoidance requires factual information about life styles to be emulated and avoided, but information gathering is costly. Because sources of such information diminish in availability with the social distance, this cost will also increase with the same distance.

An individual's reaction function is the net outcome of these four and possibly more principles working underneath. The function will take (1) negative values in the extreme negative range of the social-status disparity, (2) positive values in the neighborhood of zero social status disparity, and again (3) negative values in the extreme positive range of the same disparity. That it takes negative values at the extreme positive and negative ranges of the social status disparity can be attributed to the dominating influences of the social and economic principles. Moreover, because the psychological principle is not symmetric in its effect (i.e., the psychic satisfaction from higher status identification is positive while that from lower status identification is negative), a reaction function will generally take a skewed bell shape, tilted toward the upper status identification, as in Fig. 4.1.

A word of caution is in order: It is by no means easy to construct a reaction function without running into a normalization problem. To anchor the position of a reaction function, it would be necessary to introduce some additional hypotheses that help fix the position of its underlying component functions (including those representing the four principles above). But, such positioning cannot be left to the whim of personal tastes because one's social orientation has to be reciprocated by similar orientation by others if social norms are to be sustained. The empirical testing of a reaction function, therefore, would require that some additional restrictions be imposed on its component functions in order that these functions may yield a normalized reaction function that can be tested empirically in actual situations (see Ray 1973, for a specific example).

We have now demonstrated that social want can be represented by a reaction function. The importance of this function is twofold. First, expressed as a function of the social-status disparity, it structures social want itself (i.e., the pattern and the intensity of emulation and avoidance against relevant social groups). Second, combined with another piece of information, it becomes possible to quantify the *social want-satisfying property* of choice objects. We now turn to this quantification.

If orientation to social norms and seeking of upper status identification is what constitutes social want, the quantification of the social want-satisfying property requires information (or perception) regarding the whereabouts of these norms, in particular, information about how vogue choice objects are (perceived to be) among members of various social groups. This information is essential in determining whether consumption of a particular choice object enhances the image of being part of those groups that are being emulated or puts distance to those that are being avoided. We, therefore, need to characterize choice objects in terms of their popularity across relevant social groups.

Definition 5 Take an individual with set S of his relevant social groups, and consider commodity y_j . Let $z(y_j, g_i)$ be a popularity indicator (index) of commodity y_j for group $g_i \in S$.

For certain commodities, such as most household durables and semidurables, the popularity in group g_i can be measured by the proportion of the group's population actually using them. Because such commodities are normally not purchased in bundles, the quantities purchased are close to the number of the purchasing households. For others, however, measuring their popularity in this fashion can only be a rough approximation. In general, individuals do not possess accurate information on such proportions, nor do they have exact information on any other popularity measure. To that extent, the content of popularity indicators will be short of being objective; it may also be affected by channels of information diffusion such as advertising and word-of-mouth. Again, what is important here is that an individual has a certain perception on how popular any choice object is in each of his relevant social groups. Given the cost of information gathering and cognition, this perception is bound to be subjective and biased. But, by actually observing or hearing what significant others are purchasing, individuals will develop some idea on the social desirability of their choice alternatives, without which norm-orientation becomes an empty endeavor.

Given a reaction vector and popularity indicators of choice objects, the social want-satisfying property of a given commodity bundle $y = [y_1, y_2, \ldots, y_n]$ can be measured, as a first approximation, by multiplying each component of the reaction vector with a corresponding popularity indicator and summing this product over all relevant social groups and all components of the commodity bundle; this is a mapping from the commodity space G to the set of real numbers, $F : G \to R$; i.e.,

$$x_s = F(y) \equiv \sum_{j=1}^{n} \sum_{i \in S} z\left(y_j, g_i\right) V\left(d\left(g_i, g_m\right)\right),$$

where $z(y_j, g_i)$ is a popularity indicator of commodity y_j for relevant social group g_i, $V(d(g_i, g_m))$ is the ith component of the individual's reaction vector, subscript m denotes the group of the individual's belonging, and x_s denotes the social want-satisfying property. Such imputation summarizes the total serviceabilities of a commodity bundle to the goal of getting better social status identification through emulation and avoidance of the life styles of relevant social groups. In this vein, Hirsch's (1976) *positional goods* can be interpreted as those that take relatively high values in this imputation.

The imputation of the social want-satisfying property completes the two-step choice process of the preceding section. According to this process, an individual first identifies his satisficing feasibility set through a sequential satisficing decision rule applied to his physical wants, and then selects from this set those objects that yield the highest satisfaction of his social want. Since each commodity bundle is associated with its social want-satisfying property through mapping $x_s = F(y)$, this process yields the following choice set:

$$C(B) \equiv \left\{ y : y \in A\,(x^*) \cap B(P, M) \text{ and } F(y) \geq F\,(y')\ \forall y' \in A\,(x^*) \cap B\,(P, M) \right\},$$

where $A(x^*) \equiv \{y\colon y \in G \text{ and } \Phi(y) \geq x^*\}$ and $B(P, M) \equiv \{y\colon y \in G \text{ and } P \cdot y \leq M\}$. It is possible to rationalize this choice set by an ordering defined on the goods space. To demonstrate this point, we first define on G a relation that is induced by function F. Call this relation *a social want relation* and denote it by R_S.

Definition 6 For any two commodity bundles, y^i and y^j, in G, $y^i R_S y^j$ if and only if $F\left(y^i\right) \geq F\left(y^j\right)$. This relation can be combined with our composite relation R_G above to form a new relation R_G^* on G.

Definition 7 For any two commodity bundles, y^i and y^j, in G, $y^i R_G^* y^j$ if and only if any one of the following conditions holds:

1. $y^i \in A\,(x^*)$, $y^j \in A\,(x^*)$, and $y^i R_S y^j$,
2. $y^i \in G - A\,(x^*)$, $y^j \in G - A\,(x^*)$, and $y^i R_G y^j$,
3. $y^i \in A\,(x^*)$ and $y^j \in G - A\,(x^*)$, where $G - A(x^*) \equiv \{y\colon y \in G \text{ and } y \notin A\,(x^*)\}$.

Under this definition, the choice set $C(B)$ consists of the R_G^*-greatest elements of $B : C(B) = \left\{ y : y \in B \text{ and } yR_G y' \forall y' \in B \right\}$.

Consider now the goods space G and the set β of all possible budget sets of the form $B(P, M) \equiv \{y\colon y \in G \text{ and } P \cdot y \leq M\}$, where P is a price vector and M is income. The two constitute a so-called budget space (G, β). The question is whether an individual h is *rational* in the sense of having a preference relation R defined on G such that the choice set $h(B)$ is the set of the R-greatest elements of B for every $B(P, M) \in \beta$ (Richter 1971). We have seen that with two relations, R_G and R_S, combined, there essentially is a (preference) relation R_G^* on space G such that for each budget set $B(P, M) \in \beta$ the choice set $C(B)$ is the set of the R_G^*-greatest elements of $B(P, M)$. Thus, the individual is *rational*; in fact, he is *regular rational* since R_G^* is reflexive, transitive, and complete.

Proposition 4 *A system* (G, R_G^*) *is a preference ordering relational system* (i.e., R_G^* *is reflexive, transitive, and complete*).

This result confirms that orientation to social and cultural norms is no less rational than behavior based on idiosyncratic tastes. Faced with the bounded rationality, an individual is motivated to cut the cost of problem-solving. In our modeling, this motivation takes the form of positive and negative orientation to social groups as an individual allows his choice decisions to be guided by the life styles of his relevant social groups. In this sense, the life styles as social norms are serving as real sources of low-cost heuristics to complex problems. An individual acting in this fashion is perfectly *rational* in the formal sense of Proposition 4, because the act of referring to social norms facilitates comparison of choice objects in a socially meaningful way. But, the resulting behavior differs from what the traditional utilitarian individualism envisions in that social norms are now internalized into an individual's preferences. Among many possible modes of behavior, it is sensible and meaningful to adopt this particular mode when an individual looks for heuristic solutions that not only save the problem-solving cost but also are effective to the goal of upper social status identification.

Proposition 4 implies that it is misleading to distinguish tastes-driven behavior in economics and norm-oriented behavior in sociology by the litmus test of the existence of a well-defined preference relation on the goods space. Whether behavior is guided by personal tastes or social norms, individuals would have to have some idea as to which alternatives are more desirable than others. Orientation to social norms provides a guide no less effective on such comparison than personal tastes. The essence of Proposition 4 is that it is entirely possible for an individual to form a consistent ordering of his choice objects when his preferences are guided by social and cultural norms. It can be viewed as an ordering that prevails over a zone of socially meaningful flexible responses in the social field in Day's terms or as an ordering that justifies the practice of intentional rationality in Beckert's terms.

The existence of such a norm-guided ordering is an answer to Hodgson's call that human behavior be characterized as a purposeful norm-oriented behavior. With this ordering, an individual's decision-making autonomy is exercised at two levels, first at the level of assigning priorities to different physical wants and setting their aspiration levels, and second at the level of identifying relevant social groups for emulation–avoidance purposes and evaluating the social merits of choice objects by way of a reaction function and the (perceived) popularity indicators of choice objects. A norm-guided ordering reflects one's position in the social field, the multiple principles that underlie reaction functions, and the perception on the whereabouts of social norms. This implies that one's choice behavior changes whenever any one of the underlying forces of his reaction function shifts and whenever his perception of social norms is influenced by advertising, word-of-mouth, and other diffusion processes of information. On this point, it is useful to recall that Corneo and Jeanne (1997) have examined the possibility that norms may be intentionally created by investment in social norms by producers with market power and that if social pressures for conformity transmitted via the feeling of *envy* by others become excessive, healthy efforts aimed at upper status identification may be seriously hampered (Kolm 1995; Mui 1995).

Thus, the social disposition and cultural value orientation, jointly, absorb bounded rationality by suggesting how to integrate preferences around social and cultural norms. If the deliberation costs and other limits to rationality are substantial, norm-guided behavior will be an effective way of adapting to the environment in a socially meaningful way.[4]

What has been confirmed here, therefore, is that human behavior reflects (endogenizes) and thereby reinforces the social structure of the decision-making environment. As individuals are oriented to social and cultural norms, such norms become internalized in their preferences. Norm-guided preferences, in turn, serve as an instrument to the reproduction and evolution of these norms. Moreover, to the extent that social orientation is a consequence of the imperfect decision-making environment, the social structure reflects the nature of bounded rationality. Human behavior in society, therefore, rests on a triad of relations involving bounded rationality, the existence of social and cultural norms, and the formation of norm-guided preferences.

We close this section with an observation that the above norm-oriented choice behavior (of an individual who belongs to social group g_m and has k-relevant social groups, $S = \{g_1, \ldots, g_m, \ldots, g_k\}$) can be viewed formally as a demand correspondence $\mathrm{D} : R_+^{n+1} \to G$ such that

$$D\Big(P, M; x^*; z\left(y_j, g_i\right), V\left(d\left(g_i, g_m\right)\right),\ i = 1, \ldots, k,\ j = 1, \ldots, n\Big)$$
$$= \Big\{y : y \in B\left(P, M\right) \cap A\left(x^*\right) \text{ and } F(y) \geq F\left(y'\right)\ \forall y' \in A\left(x^*\right) \cap B\left(P, M\right)\Big\}.$$

This correspondence does not presume that relevant social groups and reaction functions are exogenous to decision makers. To the extent that reference group taking is very much influenced by one's economic and social status, which groups to emulate or avoid for social status identification is as much part of one's choices as commodity bundles. For this reason, the reference group taking itself should be viewed as a choice correspondence $\Psi\left(g_m; Z\right) : \widehat{S} \to \widehat{S}$ such that

$$\Psi\left(g_m; Z\right) = \mathrm{S} \subset \widehat{S},$$

[4]Sometimes a model of bounded rationality reduces to a model of unbounded rationality as a special case. No simple relationship of this kind exists for our model, because deliberation costs are not explicitly considered in individuals' optimization problems as in Conlisk (1988), Day and Pingle (1991), and Pingle (1992). Rather than presuming that social orientation or status seeking is an end in itself dictated by invariant (utility-measurable) preferences regardless of the limits to rationality, we have maintained that because rationality is seriously bounded, individuals see the need to develop social capital of life styles collectively as real sources of low-cost heuristics to otherwise complex choice problems and that in so doing they simultaneously develop the desire to act in socially meaningful ways by internalizing social norms and cultural values into their preference formation. We leave undefined what the world would be like if individuals were completely unbounded in their rationality, hence were not in need of any instrument, social or otherwise, by which to absorb the boundedness of rationality. Therefore, our model of norm-guided preferences and socially-oriented behavior under bounded rationality cannot be related in simple terms to a model of unbounded rationality in the limit as the limits to rationality are lifted.

where Z is an indicator of one's economic resources such as income and wealth. A reaction function is then formed *vis-à-vis* the selected reference groups. Such static characterizations of norm-oriented behavior and the reference-group taking are incomplete and should be supplemented by their dynamic features. But, taken as snapshots, these correspondences capture choice behavior and reference group taking within a zone of flexible responses that is tied to a position in the social field.

7 Leibenstein and Duesenberry Revisited

Leibenstein (1950) and Duesenberry (1949) were among the first to address issues pertaining to externalities in consumption. Leibenstein's classifications of bandwagon, snob, and Veblen effects still serve as useful characterizations of such externalities. Similarly, Duesenberry's relative income hypothesis has by no means become a dead idea. With the resurgence of strong interest in sociological perspectives to consumer behavior, we now relate our model of interdependence via reference groups to the work of Leibenstein and Duesenberry.

7.1 Leibenstein's Bandwagon, Snob, and Veblen Effects

Leibenstein (1950) argued that a market demand schedule cannot be obtained merely by adding individuals' demand schedules when externalities, such as bandwagon, snob, and Veblen effects, are present in consumption. Two points of difficulty were noted in incorporating such externalities into the neoclassical demand theory. One pertained to the assumption to be made about the nature and the amount of information that individual agents are likely to possess, and the other to the assumption to be made on their behavior. Our model of choice decision making is relevant to both of these points. The amount of information that a decision maker possesses is reflected in the shape of his reaction function through the communication principle, and this reaction function, when convoluted with the information about the whereabouts of social norms, in particular, with the popularity indicators of choice objects, can quantify the social want-satisfying property.

We find it useful to characterize Leibenstein's bandwagon, snob, and Veblen effects in terms of the characteristics of the demand correspondence above. Suppose we take an individual who belongs to group g_{m} with social status $r(g_m) = r_m$, and examine how this individual responds to an increase in the popularity of a choice object y among members of a higher status reference group at least some distance away. That is, we study $\Delta y / \Delta z(y, g_i)$ for group $g_i \in S$ whose social status rank is non-trivially higher than his; $\triangle y$ denotes change in y and $z(y,\ g_i)$ is the popularity indicator above. If this quantity turns out to be positive, the individual's demand for y may be said to exhibit Veblen effects (or emulation effects). But, with our measurement of social want-satisfying property, this response is positive only if the individual's reaction function takes a positive value for the reference group in question. This implies that the fact that a reference group is higher in social status alone is not enough to generate Veblen effects. If social sanctions against deviant behavior (the sociological principle) and the limitation of useful factual information (the communication principle) dominate the positive psychic satisfaction from upper status identification (the psychological principle), the individual's reaction function may take a negative value even for a higher status group. If this is the case, for a higher status group, an increase in the popularity of a choice object among its members does not give rise to Veblen effects.

Veblen Effects Veblen effects (emulation effects) from group $g_i \in S$ are present in an individual's demand for y if $\Delta y/\Delta z\,(y, g_i) > 0$ and if the social status rank of g_i is non-trivially higher than his. Such effects arise only if his reaction function takes a positive value for the group, i.e., $V\,(d\,(g_i, g_m)) > 0$.

If the response to an increase in the popularity of object y in a lower status reference group is to reduce consumption of y, the individual must be avoiding the image of being associated with this group. Such responses constitute snob effects. But, again, snob effects are not synonymous to avoiding whatever is popular in lower status groups. For those groups that are lower but not so distant in social status, positive sanctions for conformity and the amount of useful factual information may outweigh the psychic dissatisfaction from lower status identification, in which case the reaction function takes positive values. As long as this is the case, an increase in the popularity of object y in a somewhat lower status group may actually encourage the individual to consume more of it. Thus, the snob effects should be reserved for those groups that are low enough in social status for the reaction function to take negative values.

Snob Effects Snob effects (avoidance effects) from group $g_i \in S$ of lower social status are present in the individual's demand for y if $\Delta y/\Delta z\,(y, gi) < 0$. Such effects are present only if his reaction function takes a negative value for the group, i.e., $V\,(d\,(g_i, g_m)) < 0$.

The reaction function will, in general, take negative values for groups that are substantially higher in social status. For such groups, negative sanctions against deviant behavior and the lack of factual information for emulation will dominate even the psychic satisfaction so that the reaction function takes negative values. An increase in the popularity of object y among members of such groups will reduce its social want-satisfying property. Such effects are distinctly different from snob effects, which only apply to groups with sufficiently low social statuses.

For the group to which the individual belongs to and for those groups that are within a small social distance (from the one of his belonging), the reaction function takes positive values, so that an increase in the popularity of a choice object among such groups invites favorable responses. These responses can be identified as bandwagon effects to the extent that the individual finds it socially meaningful to consume those goods that are in vogue in groups that are clustered in social distance around his, although such effects may be more safely reserved for externalities that cut across groups of many social statuses.

Bandwagon Effects Bandwagon effects from group $g_i \in S$ are present in the individual's demand for y if $\Delta y/\Delta z\,(y, g_i) > 0$ and if g_i is within a small social distance from his. Such effects arise only if his reaction function takes a positive value for the group, i.e., $V\,(d\,(g_i, g_m)) > 0$.

Note that while Leibenstein applied such classifications to aggregate demand behavior, it is useful to apply them to demand behavior at the level of individual consumers. The aggregate demand behavior, being a mixture of all three effects, may exhibit, in net, Veblen, snob, or bandwagon effects, but such classifications on a disaggregate level serve to characterize the origins of externalities that arise from reference group taking and social status seeking, and by so doing, bring to light some of the underlying forces of social dynamics.

7.2 Duesenberry's Model of Social Interdependence

Duesenberry (1949) attempted to reformulate traditional consumer theory to resolve an apparent contradiction between the long-run constancy of the saving–income ratio and its cyclical fluctuations. In doing so, he introduced two notions: the *interdependence of preferences* among consumers and the *non-reversibility* of consumption over time.

Consider a situation (analyzed by Duesenberry), in which there are a finite number of socially interdependent individuals whose utility functions are written as a function of their own consumption and assets over some finite planning horizon, each *deflated* by a weighted average of all individuals' current consumption. If these utility functions are maximized subject to initial holdings of assets and to current and expected incomes and interest rates, each individual's *deflated* current consumption becomes a function of his own current and expected incomes, his *deflated* initial holdings of assets, and current and expected interest rates.

Thus, consumption behaviors of all individuals taken as a group are described by a system of these functions. Under the continuity and the convexity of preferences, a unique solution to the system can be found as long as the system is consistent. The crucial feature of this system is that it is homogeneous in that if certain values of consumption of all individuals are a solution under a given set of current and expected incomes and initial asset holdings of all individuals (given also current and expected interest rates), then these consumption values multiplied by a common positive factor are also a solution if all the income and asset quantities in the set are multiplied by the same factor. This property follows directly from the fact that the individuals' utility functions are homogeneous of degree zero in own consumption and assets over the planning horizon and current consumption of all interdependent individuals (see Clower 1951–52 for more general conditions).

While this homogeneity property describes consumer behavior in steadily growing situations, consumers in recessions do not easily retreat from the standards of living that the preceding booms made possible. This reluctance causes consumption to be irreversible in time with the saving–income ratio declining during recessions. Hence, over time one observes an upward drift of living standards, which gives rise to *ratchet effects* in the aggregate consumption function.

The way our reaction function shifts over different phases of business cycles offers a new perspective to such ratchet effects. Recall the economic principle underlying a reaction function, which captures both the cost (risk) of trying something new and the cost of gathering information about life styles of other social groups. These costs, as argued then, increase symmetrically as functions of social distance. Call their composite the economic cost function.

Changes in the position of the economic cost function affects the shape of a reaction function, hence the degrees to which different statuses are sought or avoided. If the economy is steadily growing with rising current and expected incomes and asset positions, the economic cost pressures will ease to allow more vigorous emulation of higher status groups. Accordingly, an individual's reaction function will have a wider spread and tilt more toward higher status groups.

When a recession sets in, there will be contractions in current and expected incomes and asset positions. Deviations from a current life style become costlier, and so does to gather information on the life styles of various social groups. Such increased cost pressures cause an individual's reaction function to have a narrower spread with its hump centered around the zero social status disparity. In a recession, therefore, an individual becomes less aggressive at emulation and clings to what has already been achieved. As the economy regains its normal growth, the cost pressures ease again to restore the reaction function to its original position favoring a more vigorous emulation of higher status groups.

To explain Duesenberry's ratchet effects, we combine this shifting of reaction functions with the fact that an affluent society has an important pair of hierarchies ordered by social prestige, status, or even cost. One is on goods, commanding that socially more prestigious goods be costlier than less prestigious ones. The other is on income with higher income commanding higher social prestige and status. Given these hierarchies, the income level of an individual places him at a certain position in the income hierarchy, and he will choose to live a life style fit to its social status. Where he is in the income hierarchy then dictates the average prestige or costliness of goods to be consumed and the average committed expenditure to live the chosen life-style.

As an individual climbs up the income ladder in booms, his reaction function shifts in favor of more vigorous emulation of higher status groups, therefore, in favor of the life styles that demand more prestigious goods. This raises both the average costliness of goods to be consumed and the average committed expenditure required of new life styles. As income declines in a recession, an individual's reaction function slides back to its conservative position, and there will be a tendency to adhere to the life style that has already been achieved. This life style has a certain social prestige, requiring accordingly a certain level of committed expenditures to sustain it. Thus, such shifting of an individual's reaction function over different phases of a business cycle, combined with the pair of hierarchies on goods and income, can explain Duesenberry's *ratchet effects* in the aggregate consumption function.

Other principles underlying the reaction function also contribute to such effects. For instance, in a steadily growing full employment economy, one's neighbors, friends, and associates will be trying new commodities more frequently than in recessions. Therefore, the amount of useful factual information passed on to individuals will increase. Moreover, with more frequent attempts at something new, social sanctions against deviant behavior will ease, which creates a more conducive environment to upper status seeking. Thus, the communicational and sociological principles tend to reinforce the economic principle in the shifting of the reaction function over business cycles. The ratchet effects, therefore, become all the more plausible when the four principles are jointly considered.

8 Conclusion

We have attempted to characterize human behavior from three perspectives: bounded rationality, the presence of social and cultural norms, and the formation of norm-guided preferences. If decision makers are seriously bounded in their rationality, they will be motivated to economize on economizing by turning to low-cost heuristics to otherwise complex choice problems. In a society that features reference-group taking and success-oriented cultural values, individuals are motivated to climb the social status ladder by making use of the low-cost heuristics that are found in the life styles of their reference groups. We have consolidated such social motives into a reaction function over relevant social groups and discussed its underlying principles in terms of economic costs, social pressures, psychic satisfactions, and communicational advantages that are involved in emulating and avoiding of social groups. Convoluting this function with information about the whereabouts of social norms makes it possible to order choice objects by their overall serviceabilities to the satisfaction of social want. When this ordering is combined with one based on the priorities of physical wants, a comprehensive ordering of choice objects emerges that rationalizes human behavior under bounded rationality.

In our view, preferences that are invariant to social and cultural structures of the decision- making environment and which are independent of the boundedness of the decision maker's rationality are too restrictive to account for human behavior that features social orientation and economization of economizing. If preferences were entirely idiosyncratic in nature, social norms would not reproduce themselves. If individuals were assumed to possess utility functions that favored certain social interactions, social norms might be explained, after an infinite regress, as resulting from voluntary choices based on such functions, but in that case the preference formation itself would remain independent of social and cultural norms with the causation running from fixed preferences to social norms. The utility theory of social norm formation, therefore, cannot account for the interaction between social norms and the formation of norm-guided preferences, nor can it account for possible mechanisms by which social and cultural norms are internalized into preferences of individuals.

We argue that the shaping of preferences that favor social orientation and the reproduction of social norms reinforce each other through the formation of norm-guided preferences, and that bounded rationality is instrumental to the formation of such endogenous preferences. The traditional utilitarian individualism provides a micro-foundation of human behavior that abstracts basically from social and cultural elements of decision making. Our view that individuals internalize social and cultural norms into their preferences provides both a macro- and a micro-foundation of human behavior in a socially and culturally structured decision-making environment. The autonomy in decision making as well as in formation of consistent preferences is retained at a micro-level while macro-phenomena of social and cultural norms feed back into the formation of preferences that can account for the presence and evolution of social norms. With such internalization, individual decision makers behave in the middle ground between the complete voluntaristic individualism on the one extreme and the structural determinism on the other – a view that answers Hodgson's (1986) call for a new theory of human behavior.

Acknowledgements Many ex-professors, friends, and colleagues have helped me shape my thought on this topic. I am particularly thankful to Paul H. Ray, with whom I exchanged views on many occasions during period 1973–1974 at the University of Michigan. This exchange contributed greatly to my first conceptualization of possible connections between social interdependence and bounded rationality (Hayakawa 1973), which has evolved to take its present form with the help of the present enormous literature on related subjects. I am also grateful to Lester D. Taylor, W.H. Locke Anderson, Michael Manove, Suezo Ishizawa, Vai-Lum Mui, Allan G. Sleeman, Hajime Miyazaki, John T. Sjöström, Koichi Futagami and many others for their insightful comments and suggestions at various stages of the development. I would also like to thank an anonymous referee of this journal for his suggestions. Needless to say, all remaining shortcomings reflect only my own limitations.

Addendum: Afterthought and Possible Extensions[5]

Adam Smith's theory of moral sentiments (1759), Veblen's theory of the leisure class (1899), Bourdieu's logic of practice and distinction (1984, 1990), and Parsons's theory of social systems (1951), all support the view that human wants or dispositions are rooted in institutionalized values, and expressed in symbolic forms. If the social value of a good consists in symbolic profits it yields, this value cannot be known a priori before knowing how meaningful it is to acquire or consume it in the light of a culture in which such expression takes place. This fact implies

[5]This addendum has been newly written by the author for this book chapter.

that at least four factors need to be considered in understanding expressive behavior guided by common normative values: (1) a social space of life-styles of different social groups, that are characterized by proper codes of decorum and consumption knowhow, and which mediate consumption for its symbolic expression, (2) a measure of social distance that separates social groups on a commonly recognized social status ladder, (3) an effective reaction pattern, not in a game-theoretical sense of an optimal response function derived from a payoff function given a priori, but rather as an effective composite pattern of orientation based on economic, sociological, psychological, communicational, and other hidden factors, and (4) a measure of the serviceabilities of goods or a collection of goods as the status symbols of life-styles, which informs which goods, if consumed or owned, can yield symbolic profits in terms of higher status identification. If the social value of choice objects is determined through these factors, consumer preferences can no longer be isolated from this social space, nor from the motivational structure of individual agents in which normative or cultural values are introjected. In the final analysis, individual agents' preferences must be convoluted, as a product of socially acquired dispositions on the one hand and the symbolic values of choice objects that serve the life-styles to be emulated or avoided on the other.

In this chapter, based on Hayakawa (2000), I related such symbolic expression and cultural consumption to the economics of bounded rationality *à la* Simon and to the economics of limited cognition, by drawing on the following points: (1) The decision-making environment including the internal psychology and the cognitive capacity of a decision maker is not perfect; (2) the time endowment is fixed so that all activities including cognition must compete for the use of time; (3) information is insufficient, but information gathering and processing is costly; (4) typical circumstances in which decisions are made are imbued with risk and uncertainty that cannot be reduced to probabilistic terms. Facing these limitations, decision makers may turn to the procedurally rational way of handling them in order to economize on the cost of problem solving, by making use of those heuristic modes of behavior that emerged through an error-learning process of cultural evolution, particularly when the logic of expressive behavior requires not only a shared code of interpretation but also a motivational structure that evaluates goods in accordance with it. This is quite consistent with Simon's insight on the procedural rationality (1978).

In the light of the fact that the society takes on a bigger meaning than a mere aggregation of its parts, I proposed a theory of choice behavior that answers not only Simon's call for procedural rationality but also Hodgson's call for norm-oriented purposive behavior, by positing that low-cost heuristics to otherwise complex problems can be found in the life-styles of social groups that have been accumulated and refined collectively. Since the life styles of social groups are expressions of cultural consumption mediated by a symbolism that permeates the social space, the act of referring to them as a real source of guide for emulation and avoidance is very much in accord with the need-dispositions that are disposed to interpret consumption as part of cultural expression. Such norm-guided behavior, combined with the boundedly rational decision making procedure, suggests that there is an ample ground to be cultivated for productive research on the institutional nature of decision making. This agenda resonates with Hodgson's recapitulation, of reconstitutive effects of institutions on the preferences of individuals, of the habit formation through institutional channels and constraints as the key to the mechanism of such reconstitution, and of the degree on which institutional evolution may depend on this habit formation (Hodgson 2004). It also shares the stand that Gintis (2009) takes, namely, that humans have a normative predisposition that allows common beliefs and social norms to choreograph a correlated equilibrium, which points to a new direction in research on how the bounds of reason and forms of sociality can be integrated by a higher principle that correlates conflicting interests of social actors. It may be regarded as an outgrowth of Parsons's theory of institutionalization, which says that social systems are constituted of the need-dispositions acquired by individual actors through internalization of common normative values in their motivational structures. If we are to understand human behavior and socio-economic institutions in which this behavior is embedded, social ontology with its emphasis on downward causation, should be integrated with economic ontology with its emphasis on upward causation, for the actual expressive behavior takes place in the space that opens up at the crossing point of the two. In such integration, there remains many questions that need to be addressed: how endogenous preference formation, resulting from norm-orientation, can be explained by an evolutionary game theory of norm formation on the one hand and by an epistemological game theory dealing with knowledge of social norms on the other, how the legal framework and the psychological make-up of individuals contribute to the formation of social norms and prosocial preferences, how the spirit of independence, innovation, and defiance against the status-quo is retained while orientation to social norms and sanctions does not lose its force. To make headway with these questions, we need an extensive interdisciplinary research across many fields, with game theory providing one possible unifying principle. I have made an extensive inquiry into the institutional nature of decision making in my multidisciplinary decision science symposium paper, with its emphasis on how the voluntary nature of decision making is sustained when the need-dispositions, through internalization of the institutionalized values, shape the motivational structure of individual agents (Hayakawa 2010). This is one small step into the research on the 'institutionalized rationality' of socio-economic agents as opposed to the 'independent rationality' of autonomous agents that has prevailed in mainstream economics for so many years.

References

Abel AB (1990) Asset prices under habit formation and catching up with the Joneses. Am Econ Rev 80(2):38–42

Akerlof G (1980) A theory of social custom, of which unemployment may be a consequence. Q J Econ 94(4):749–775

Alchian AA (1950) Uncertainty, evolution, and economic theory. J Pol Econ 58(3):211–221

Allingham MG, Morishima M (1973) Veblen effects and portfolio Selection. In: Morishima M et al (eds) Theory of demand: real and monetary. Clarendon, Oxford, pp 242–270

Baron JN, Hannan MT (1994) The impact of economics on contemporary sociology. J Econ Lit 32(3):1111–1146

Baxter JL, Moosa IA (1996) The consumption function: a basic needs hypothesis. J Econ Behav Organ 31(1):85–100

Becker GS (1965) A theory of the allocation of time. Econ J 75(299):493–517

Becker GS (1974) A theory of social interactions. J Pol Econ 82(6):1063–1093

Beckert J (1996) What is sociological about economic sociology? Uncertainty and the embeddedness of economic action. Theory Soc 25(6):803–840

Bourdieu P (1984) Distinction: a social critique of the judgement of taste (trans: Nice R). Harvard University Press, Cambridge, MA

Bourdieu P (1990) The logic of practice (trans: by Nice R). Stanford University Press, Stanford

Bowles S (1998) Endogenous preferences: the cultural consequences of markets and other economic institutions. J Econ Lit 36(1):75–111

Camic C (1986) The matter of habit. Am J Sociol 91(5):1039–1087

Clower RM (1951–52) Professor Duesenberry and traditional theory. Rev Econ Stud 19(3):165–78

Coelho PRP, McClure JE (1998) Social context and the utility of wealth: addressing the Markowitz challenge. J Econ Behav Organ 37(3):305–314

Cole HL, Mailath GJ, Postlewaite A (1992) Social norms, saving behavior, and growth. J Pol Econ 100(6):1092–1125

Coleman JS (1990) Foundations of social theory. Harvard University Press, Cambridge, MA

Coleman JS (1994) A rational choice perspective on economic sociology. In: Smelser NJ, Swedberg R (eds) The handbook of economic sociology. Princeton University Press, Princeton, pp 166–180

Conlisk J (1980) Costly optimizers versus cheap imitators. J Econ Behav Organ 1(3):275–293

Conlisk J (1988) Optimization cost. J Econ Behav Organ 9(3):213–228

Conlisk J (1996) Bounded rationality and market fluctuations. J Econ Behav Organ 29(2):233–250

Corneo G, Jeanne O (1997) Snobs, bandwagons, and the origin of social customs in consumer behavior. J Econ Behav Organ 32(3):333–347

Cyert RM, March JG (1963) A behavioral theory of the firm. Prentice Hall, Englewood Cliffs

Day RH (1967) Profits, learning and the convergence of satisficing to marginalism. Q J Econ 81(2):302–311

Day RH (1984) Disequilibrium economic dynamics: a post-Schumpeterian contribution. J Econ Behav Organ 5:57–76

Day RH (1986) Disequilibrium economic dynamics: a post-Schumpeterian contribution. In: Day RH, Eliasson G (eds) The dynamics of market economies. North-Holland, Amsterdam, pp 51–70

Day RH (1987) The general theory of disequilibrium economics and of economic evolution. In: Batten D, Johansson B, Casti J (eds) Economic evolution and structural adjustment. Springer, Berlin, pp 46–63

Day RH, Pingle M (1991) Economizing economizing. In: Frantz R, Singh H, Gerber J (eds) Handbook of behavioral economics, vol 2B. JAI Press, Greenwich, pp 509–522

Day RH, Tinney EH (1968) How to co-operate in business without really trying: a learning model of decentralized decision making. J Pol Econ 76:583–600

Duesenberry JS (1949) Income, saving and the theory of consumer behavior. Harvard University Press, Cambridge, MA

Evans WN, Oates WE, Schwab RM (1992) Measuring peer group effects: a study of teenage behavior. J Pol Econ 100(5):966–991

Fershtman C, Weiss Y (1993) Social status, culture and economic performance. Econ J 103(419):946–959

Frank RH (1984) Are workers paid their marginal products? Am Econ Rev 74(4):549–571

Frank RH (1985) The demand for unobservable and other non-positional goods. Am Econ Rev 75(1):101–116

Futagami K, Shibata A (1995) Keeping one step ahead of the Joneses: status, the distribution of wealth, and long run growth. Discussion paper no 38. Department of Economics, Ritsumeikan University
Gaertner W (1973) A dynamic model of interdependent consumer behavior. Dissertation, University of Bonn
Gali J (1994) Keeping up with the Joneses: consumption externalities, portfolio choice, and asset prices. J Money Credit Bank 26(1):1–8
Georgescu-Roegen N (1954) Choice, expectations, and measurability. Q J Econ 68:503–534
Gintis H (2009) The bounds of reason: game theory and the unification of the behavioral sciences. Princeton University Press, Princeton
Granovetter M (1985) Economic action and social structure: the problem of embeddedness. Am J Sociol 91(3):481–510
Granovetter M, Soong R (1986) Threshold models of interpersonal effects in consumer demand. J Econ Behav Organ 7(1):83–99
Hall R (1988) Intertemporal substitution in consumption. J Pol Econ 96(2):339–357
Hayakawa H (1973) Mr. Q, a resident of an imperfect world and a social satisficer. Manuscript. University of Michigan
Hayakawa H (1976) Consumer theory when prices and real income affect preferences. South Econ J 42(1):840–845
Hayakawa H (2000) Bounded rationality, social and cultural norms, and interdependence via reference groups. J Econ Behav Organ 43(1):1–34
Hayakawa H (2010) Institutionalization of common normative values and economics of limited cognition in search for a middle ground between economics and sociology: theories of Veblen, Adam Smith, Bourdieu, and Parsons. Paper presented at the multidisciplinary decision science symposium, Singapore, Aug 2010
Hayakawa H, Venieris Y (1977) Consumer interdependence via reference groups. J Pol Econ 85(3):599–616
Hayek FA (1952) The counter-revolution of science. Free Press, Glencoe
Heiner RA (1983) The origin of predictable behavior. Am Econ Rev 73(4):560–595
Heiner RA (1989) The origin of predictable dynamic behavior. J Econ Behav Organ 12(2):233–257
Hirsch F (1976) Social limits to growth. Harvard University Press, Cambridge, MA
Hodgson G (1986) Behind methodological individualism. Cambridge J Econ 10(3):211–224
Hodgson G (2004) Institutional economics: from Menger and Veblen and North. In: Davis JB, Marciano A, Runde J (eds) The Elgar companion to economics and philosophy. Edward Elgar, Cheltenham
Hof FX (1997) Relative consumption in dynamic macroeconomic models of closed economies. Manuscript. University of Technology Vienna, Vienna
Ikeda S (1994) Consumption externalities, optimal saving, and dynamics. Manuscript, Osaka University
Jevons WS (1957) The theory of political economy, 5th edn. Kelley and Millman, New York
Johnson HG (1952) Effects of income-redistribution of aggregate consumption with interdependence of consumers' preferences. Economica 19(74):131–147
Kalman PJ (1968) Theory of consumer behavior when prices enter the utility function. Econometrica 136(3/4):497–510
Kapteyn A (1980) The dynamics of preference formation. J Econ Behav Organ 1(2):123–157
Keynes JM (1921) The treatise on probability. In: The collected writings of John Maynard Keynes, vol 8. Macmillan, London, 1973
Knight F (1921) Risk, uncertainty and profit. Kelly, Boston
Kocherlakota NR (1996) The equity premium: it's still a puzzle. J Econ Lit 34(1):42–71
Kolm SC (1995) The economics of social sentiments: the case of envy. Jpn Econ Rev 46(1):63–87
Kornai J (1971) Anti-equilibrium. North-Holland, Amsterdam
Krelle W (1973) Dynamics of the utility function. In: Hicks JR, Weber W (eds) Carl Menger and the Austrian school of economics. Oxford University Press, New York

Leibenstein H (1950) Bandwagon, snob and Veblen effects in the theory of consumers' demand. Q J Econ 64(2):183–207

Linder S (1970) The harried leisure class. Harvard University Press, Cambridge, MA

Lucas RE Jr (1976) Econometric policy evaluation: a critique. In: Brunner K, Meltzer AH (eds) The Phillips curve and labor markets. Carnegie–Rochester conference series on public policy 1. North-Holland, Amsterdam, pp 19–46

Mankiw N, Rotemberg J, Summers L (1985) Intertemporal substitution in macroeconomics. Q J Econ 100(1):225–251

Marshall A (1920) Principles of economics, 8th edn. Macmillan, London

Mehra R, Prescott EC (1985) The equity puzzle. J Monet Econ 15(2):145–161

Menger C (1950) Principles of economics (trans: Dingwall J, Hoselitz BF). Free Press, Glencoe. Original work published 1871

Miller DR (1963) The study of social relationships: situations, identity and social interaction. In: Kock S (ed) Psychology: a study of a science, vol 5. McGraw-Hill, New York

Mui VL (1995) The economics of envy. J Econ Behav Organ 26(3):311–336

Naish HF (1993) The near optimality of adaptive expectations. J Econ Behav Organ 20(1):3–22

Parsons T (1951) The social system. Free Press, New York

Parsons T, Shils EA (1951) Some fundamental categories of the theory of action: a general statement. In: Parsons T, Shils EA (eds) Toward a general theory of action. Harvard University Press, Cambridge, MA

Pigou AC (1913) The interdependence of different sources of demand and supply in the market. Econ J 23:19–24

Pingle M (1992) Costly optimization: an experiment. J Econ Behav Organ 17(1):3–30

Pingle M, Day RH (1996) Models of economizing behavior: experimental evidence. J Econ Behav Organ 29(2):191–209

Pollak RA (1976) Interdependent preferences. Am Econ Rev 66(3):309–320

Radner R (1975) Satisficing. J Math Econ 2(2):253–262

Radner R, Rothschild M (1975) On the allocation of effort. J Econ Theory 10(3):358–376

Ray PH (1973) Residential location, consumer demand and urban structure. Dissertation, University of Michigan

Richter MK (1971) Rational choice. In: Chipman JS, Hurwicz L, Richter MK, Sonnenschein HF (eds) Preferences, utility, and demand. Harcourt Brace and Jovanovich, New York, pp 29–58

Robson AJ (1992) Status, the distribution of wealth, private and social attitudes to risk. Econometrica 60(4):837–857

Schumpeter JA (1934) The theory of economic development. Harvard University Press, Cambridge, MA

Shipley FB II (1974) Convergence of adaptive decisions. In: Day RH, Groves T (eds) Adaptive economic models. Academic, New York, pp 247–272

Simon HA (1955) A behavioral model of rational choice. Q J Econ 69(1):99–118

Simon HA (1959) Theories of decision-making in economics and behavioral sciences. Am Econ Rev 49(3):253–283

Simon HA (1976) Administrative behavior, 3rd edn. Free Press, New York

Simon HA (1978) Rationality as process and as product of thought. Am Econ Rev 68(2):1–16

Smelser NJ (1990) Can individualism yield a sociology? Contemp Sociol 19(6):778–783

Smelser NJ, Swedberg R (1994) The sociological perspective on the economy. In: Smelser NJ, Swedberg R (eds) The handbook of economic sociology. Princeton University Press, Princeton, pp 3–26

Smith A (1969) The theory of moral sentiments. Arlington House, New York. First published in 1759

Stigler G (1961) The economics of information. J Pol Econ 59(3):213–225

Tilly C (1991) Individualism askew. Am J Sociol 96:1007–1011

Veblen T (1899) The theory of the leisure class. Macmillan, New York

Vriend NJ (1996) Rational behavior and economic theory. J Econ Behav Organ 29(2):263–285

Chapter 5
Keeping One Step Ahead of the Joneses: Status, the Distribution of Wealth, and Long Run Growth

Koichi Futagami and Akihisa Shibata

Abstract Assuming that the utility of each agent depends on its relative wealth position in the society, this chapter constructs an endogenous growth model. It is shown that even if the subjective discount rates differ across agents, there exists a unique balanced growth equilibrium in which each agent owns a positive share of the world wealth. It is also shown that if the agents are identical then an increase in savings incentives always raises the long run growth rate but if they are heterogeneous then an increase in savings incentives may lower the long run growth rate.

Keywords Wealth preference • Interpersonal dependency of preference • Social status • Endogenous growth • Wealth distribution

1 Introduction

To be rich or to be poor, it is one of the most important issues for human beings. If you are rich, then people treat you differently from the ordinary man. Thus, your utility depends on whether you are rich or not, that is, your utility would depend on your wealth holdings as well as your consumption. We call this dependency of utility

---

The original article first appeared in the Journal of Economic Behavior and Organization 36: 109–126, 1998. A newly written addendum has been added to this book chapter.

K. Futagami (✉)
Graduate School of Economics, Osaka University, Toyonaka, Osaka 560-0043, Japan
e-mail: futagami@econ.osaka-u.ac.jp

A. Shibata
Kyoto Institute of Economic Research, Kyoto University, Yoshida, Kyoto 606-8501, Japan
e-mail: shibata@kier.kyoto-u.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_5

on wealth holdings *wealth preference*. The existence of wealth preference has been emphasized by David Hume, Adam Smith, J. S. Mill, Karl Marx, among others.[1]

In addition to the fact that agents' utility would depend on wealth holdings, it should be stressed that the notions of richness and poverty are also relative ones. An individual is interested not only in the absolute level of his or her position in the society but also in the relative position in that society. This point has been also pointed out by many authors including Marx, Hume and Thorstein Veblen.[2] For example, Hume (1978) states: "One of the most considerable of these passions is that of love or esteem in others, which therefore proceeds from a sympathy with the pleasure of the possessor. But the possessor has also a secondary satisfaction in riches arising from love and esteem he acquires by them, and this satisfaction is nothing but a second reflection of that original pleasure, which proceeded from himself. This secondary satisfaction or vanity becomes one of the principal recommendations of riches, and is the chief reason, why we either desire them for ourselves, or esteem them in others".

In other words, an agent's utility depends on its relative position in the society rather than the absolute level of its own wealth. We call this dependency of utility on other people's behavior *interpersonal dependency of preference*. This chapter aims to construct a formal dynamic model in the spirit of Hume, Marx, Veblen and others.

The two concepts, wealth preference and interpersonal dependency of preference, have been modeled by many studies mainly in a static framework. See, for example, Duesenberry (1949), Frank (1985), Robson (1992) and Fershtman and Weiss (1993). Recently, however, a few important *dynamic* studies on the subjects have appeared.[3] Ono (1994) emphasizes the importance of wealth preference and constructs a dynamic general equilibrium model with Keynesian features.[4] In his pioneering work, Ikeda (1993, 1995) construct a dynamic general equilibrium model of interpersonal dependency of preference (in which an agent's utility depends on other agents' consumption) and analyzes the effects of the interpersonal dependency of preference on wealth distributions.[5] Along the same lines of Ono and Ikeda, this chapter combines these two concepts; wealth preference and the interpersonal dependency of preference. In this chapter, the utility of each agent is assumed to depend on its relative wealth position in the society to whichit belongs as well as on

[1] A typical example is found in Marx (1964, p. 167). See also Chap. 1 of Ono (1994) for this point.

[2] See Cole et al. (1995) for this point.

[3] An early exception is Kurz (1968). He develops an optimal growth model in which the representative agent gets utility directly from holding capital.

[4] Gylfason (1993) and Zou (1995) also construct a growth model with wealth preference.

[5] Abel (1990) and Gali (1994) also develop dynamic models of consumption externalities and analyze the effects of the existence of consumption externalities on asset price determination. However, their attention is restricted to a pure exchange economy in which all agents are symmetric.

its absolute consumption level. We call the person's relative wealth position in the society *status*. This can be considered a natural formulation of the ideas of Hume, Marx, Veblen and others

Cole et al. (1992, 1995), Bakshi and Chen (1996) and Fershtman et al. (1996) are also closely related to our study. The two papers by Cole, Mailath and Postlewaite stress the importance of incorporating concern for relative wealth into economic models. Cole et al. (1992) examines a similar issue as that of this chapter, that is, how the existence of status preference affects capital accumulation, but they choose not to assume that status is directly included in the utility function and, instead, regard status as a ranking device determined by the allocation of nonmarket goods.[6] Although we basically agree with their assertion that the direct incorporation of status into the utility function is not desirable, we nevertheless assume it to be an argument of the utility function. The reasoning behind this is that their two papers can be regarded as providing the microfoundations of our formulation and that it is very difficult to analyze the dynamic implications of status preference (which is the main concern of this chapter) without this simplifying assumption.

Using three type of status preference, Bakshi and Chen examine the effects of status preference on consumption, portfolio choice and stock prices. Although one of their specifications of status preference (model 2 in their paper) is very close to ours, only the effect of status preference on the first order conditions of consumers is analyzed. Characterizing the general equilibrium dynamics in the setting remains unsolved in their paper.

Fershtman, Murphy and Weiss construct an overlapping generations model with heterogeneous agents whose nonwage income and learning ability differ across them. The distribution of nonwage income and the learning ability is exogenously given, and is assumed to be not time varying. Each agent is assumed to get utility directly from relative educational position or occupational status. Under this setting, they analyze how the long run growth rate is affected by changes in rewards on social status and by exogenous changes in distribution of nonwage income.

In their model, status is determined by the relative level of human capital. Although we agree with their view that in evaluating social status human capital is an important factor, we think that financial assets are also important, as stated in the classics. Thus, in this chapter, we define social status as a function of the relative level of financial assets rather than human capital.[7]

Our major results in this chapter are summarized below. If all consumers are identical, the stronger the degree of status preference is, the higher the long run rate of balanced growth path is. However, if agents are heterogeneous then the results obtained in the identical agents case can be overturned, and three interesting results are obtained. First, even if the subjective discount rates differ across agents, there

[6] Corneo and Jeanne (1997a) construct a related model and show that social segmentation stimulates economic growth through status-seeking motives.

[7] Corneo and Jeanne (1997b) provide a microfoundations of this formulation. They show that both too much equality and too much inequality may be detrimental to economic growth.

exists a unique steady growth path on which each agent owns a positive share of the world wealth. This result is contrast with the well known work by Ramsey (1928) and Becker (1980), whose results show that the most patient agent will eventually hold all of the non-human wealth existing in the economy. Rather, our result somewhat resembles those of Epstein and Hynes (1993) and Ikeda. Examining a model in which time preference is endogenous and dependent on the level of consumption, Epstein and Hynes show that, even if the shapes of time preference functions differ across agents, each agent possesses a positive share of the world wealth in the steady state. Ikeda develops a model in which each agent's utility depends on other agents' consumption and shows that, even if the discount rates differ across agents, each agent would own a positive amount of wealth.

Second, the most myopic agent, that is, the agent with the largest subjective discount rate, may hold a larger share of non-human wealth than more patient agents, if the strength of each individual's status preference differs. This result reconfirms Ikeda's analysis of the snob effect (one type of consumption externality) in a different setting.[8]

Third, when the status preference of an agent is strengthened, the long run rate of growth could be lowered. Despite the fact that each agent has perfect foresight in this chapter's model and that agents strongly desire to accumulate wealth, the long run growth rate decreases contrary to their desire. This result is close to Fershtman, Murphy and Weiss. They show that the long run growth rate can be decreased through a misallocation of talent when social rewards on status are strengthened. That is, to obtain a higher position in the society wealthy but low ability persons acquire schooling and crowd out high ability but poor persons. Although their result bears a close resemblance to ours, there is a significant difference between their and our model. In their model, the distribution of wealth (nonwage income) is given exogenously. In contrast, the long run distribution of wealth in our model is determined endogenously dependent on the strength of status preference and the subjective discount rates. As stressed above, the determination of long run wealth distribution itself is an important topic in the economic growth literature. Our model allows analyses on the simultaneous determination of wealth distribution and the long run growth rate.

The organization of this chapter is as follows. In the next section, as a benchmark case, a symmetric agents model is presented, and it is shown that the existence of status preference enhances the long run growth rate. Section 3 extends the benchmark model to an asymmetric agents model, and characterizes the general equilibrium. In Sect. 4 conclusions are presented.

[8] Ikeda (1993, 1995) shows that, when there is a bandwagon effect, the opposite case occurs.

2 A Model with Symmetric Agents

2.1 The Basic Model

Let us first present the basic structure of the model. There is one commodity, which is used for both consumption and investment. For simplicity assume the production function

$$Y = AK, \tag{5.1}$$

where K is capital.

Assume that the total size of population is constant and normalized to unity. Agent i has the following utility function:

$$U_i = \int_{t=0}^{\infty} u_i\left(C_i, \frac{W_i}{\overline{W}}\right) e^{-\rho t}\, dt,$$

where C_i and ρ_i are consumption and the constant subjective discount rate of agent i. The emphasis of this chapter's formulation is the dependence of utility on the ratio of agent i's wealth (W_i) to the average level of wealth in the economy ($\overline{W}$). This term represents the agent's preference regarding status.[9] As mentioned in the introduction, this formulation follows the spirit of Hume, Marx, Veblen and others.[10] For simplicity, in the following analysis we specify the utility function as

$$U_i = \int_{t=0}^{\infty} \frac{\left[C_i^{\alpha} V_i \left(\frac{W_i}{\overline{W}}\right)^{\beta}\right]^{1-\gamma} - 1}{1-\gamma} e^{-\rho t}\, dt, \tag{5.2}$$

where γ is the inverse of the rate of intertemporal substitution. V is assumed to be a monotonically increasing function of $W_i/\overline{W}$. In order to assure the concavity of the utility function, we impose the following restriction[11]:

$$1 - \alpha\,(1-\gamma) > 0. \tag{5.3}$$

[9] See Corneo and Jeanne (1995) for a different specification of status preference.

[10] A similar utility function is employed by Bakshi and Chen (1996). Using three stochastic models, they show that the existence of status preference can be a driving force behind stock market volatility.

[11] This condition implies that the intertemporal substitution of consumption without wealth preference ($\beta = 0$) is positive. See discussion below (5.22) for the empirical validity of the assumption.

The curvature of the V function represents the strength of status preference. The higher the degree of the curvature is, the stronger the degree of the status preference. Each agent is endowed with a positive initial non-human wealth $W_i(0)$, which is assumed to be identical in this section. Furthermore, assume, in this section, that preferences of all agents are identical, so that there are many symmetric agents in this economy. Under this assumption, the structure of our model seems to be similar to Gylfason (1993) and Zou (1995). However, the basic idea behind our formulation differs from theirs, and this difference plays an important role in the next section.

The flow budget equation of agent i is:

$$\dot{W}_i = AW_i - C_i. \tag{5.4}$$

Regarding the average level of wealth ($\overline{W}$) as exogenously given, as Romer (1986) assumes in the case of production externalities, agent i maximizes (5.2) subject to the flow budget equation (5.4) under perfect foresight. Denoting the co-state variable of W_i as q_i, we obtain the optimal conditions to this maximization problem:

$$\alpha C_i^{\alpha(1-\gamma)-1} V\left(\frac{W_i}{\overline{W}}\right)^{\beta(1-\gamma)} = q_i, \tag{5.5}$$

$$-\dot{q}_i + \rho q_i = \frac{\beta V\left(\frac{W_i}{\overline{W}}\right)^{(1-\gamma)\beta)-1} V'\left(\frac{W_i}{\overline{W}}\right) C_i^{\alpha(1-\gamma)}}{\overline{W}} + Aq_i, \tag{5.6}$$

$$\lim_{t\to\infty} q_i W_i e^{-\rho t} = 0. \tag{5.7}$$

Equation (5.5) is the optimal condition with respect to C_i, (5.6) the Euler equation and (5.7) the transversality condition (TVC).

2.2 Market Equilibrium

Next consider the market equilibrium conditions. The asset and goods markets equilibrium conditions are given by

$$\sum_i W_i = K, \tag{5.8}$$

$$\dot{K} = AK - \sum_i C_i. \tag{5.9}$$

Since the agents are assumed to be symmetric and the size of the population is unity, the following relations hold in equilibrium:

$$W_i = \overline{W} = K_i = K, \tag{5.10}$$

$$\sum_i C_i = C = C_i.$$

2.3 Equilibrium Dynamics

Using (5.8) and (5.10), we can rewrite (5.5) and (5.6) as

$$\alpha C^{\alpha(1-\gamma)-1} V(1)^{\beta(1-\gamma)} = q, \tag{5.11}$$

$$-\dot{q} + \rho q = \frac{\beta V(1)^{\beta(1-\gamma)-1} V'(1) C^{\alpha(1-\gamma)}}{K} + Aq. \tag{5.12}$$

Combining (5.11) and (5.12) gives

$$-\frac{\dot{q}_i}{q_i} = A - \rho + \theta\left(\frac{C}{K}\right), \tag{5.13}$$

where $\theta \equiv \beta V'(1)/\alpha V(1)$. θ reflects the strength of status preference. The greater the value of θ is, the stronger the status preference is. Note that the RHS of (5.13) implies that the subjective discount rate is modified by the term which represents the strength of status preference, $\theta(C/K)$. We call this modified rate, $\rho - \theta\,(C/K)$, the effective discount rate and distinguish it from the subjective discount rate ρ.

Differentiating (5.11) with respect to time, we have

$$\frac{\dot{q}}{q} = \frac{[\alpha\,(1-\gamma) - 1]\,\dot{C}}{C}, \tag{5.14}$$

Thus, it follows from (5.13) to (5.14) that

$$\frac{\dot{C}}{C} = \frac{A - \rho + \theta\,(C/K)}{1 - \alpha\,(1-\gamma)}. \tag{5.15}$$

On the other hand, (5.9) can be rearranged as

$$\frac{\dot{K}}{K} = A - \frac{C}{K}. \tag{5.16}$$

Since (5.15) and (5.16) include only C and K, they constitute the full dynamics of C and K.

2.4 The Balanced Growth Equilibrium: Existence and Stability

Consider a balanced growth path and denote the growth rate by g. Then, by definition we have

$$\mathrm{g} = \frac{\dot{C}}{C} = \frac{\dot{K}}{K}. \tag{5.17}$$

From (5.16) to (5.17) it follows that

$$\frac{C}{K} = A - \mathrm{g}. \tag{5.18}$$

Substituting (5.18) into (5.15) gives

$$\mathrm{g} = \frac{A(1+\theta) - \rho}{1 - \alpha(1-\gamma) + \theta}, \tag{5.19}$$

which shows how the long run growth rate (if it exists) is determined. The long run growth rate depends on the strength of status preference, θ as well as α, γ, A, and ρ.

Next consider the condition for the existence of a balanced growth equilibrium. Define a new variable as

$$x \equiv \frac{C}{K}.$$

Then we have

$$\frac{\dot{x}}{x} = \frac{\dot{C}}{C} - \frac{\dot{K}}{K} = \frac{A - \rho + \theta x}{1 - \alpha(1-\gamma)} - (A - x) = \frac{A\alpha(1-\gamma) - \rho}{1 - \alpha(1-\gamma)} + \frac{\theta + 1 - \alpha(1-\gamma)}{1 - \alpha(1-\gamma)} x \tag{5.20}$$

or

$$\dot{x} = \left[\frac{A\alpha(1-\gamma) - \rho}{1 - \alpha(1-\gamma)} + \frac{\theta + 1 - \alpha(1-\gamma)}{1 - \alpha(1-\gamma)} x \right] x.$$

Denote the steady state value of x as x^*. Then, there always exists a positive x^* for any value of $\theta \in [0, \infty)$ if the following condition is satisfied[12]

$$\rho - A\alpha(1-\gamma) > 0. \tag{5.21}$$

Note that this condition is always satisfied when $\gamma > 1$. Throughout the paper, we assume that γ is greater than one, so that (5.21) always holds. This assumption implies that the degree of intertemporal substitution ($1/\gamma$) is relatively low and thus the sensitivity of saving to changes in the interest rate is low. Since estimates for

[12]Note that the concavity of the utility function requires condition (5.3), that is, $1 - \alpha(1-\gamma) > 0$.

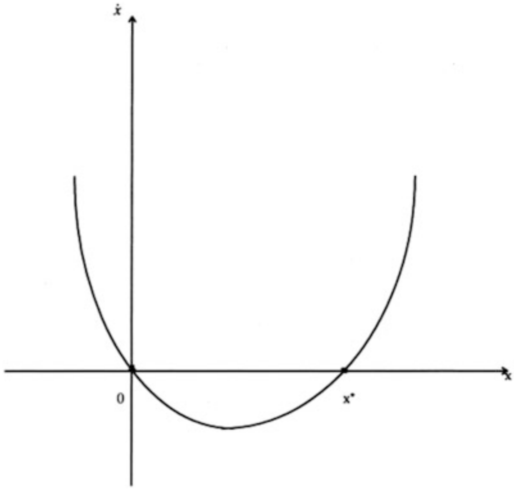


Fig. 5.1 The symmetric agents case

$1/\gamma$ by Hall (1988) and Caballero (1988) concentrate around zero, this assumption would be satisfied empirically.

It is easily verified that differential (5.20) is unstable and thus the economy is always on the balanced growth path. See Fig. 5.1 for a proof. Furthermore, on this balanced growth path, the TVC condition, (5.7), is automatically satisfied under (5.21).[13]

2.5 *The Effect of Status Preference on Long Run Growth*

Let us here examine the relation between the strength of status preference and the long run growth rate. By differentiating (5.19) with respect to θ, we obtain

$$\frac{dg}{d\theta} = \left(\frac{1}{1-\alpha(1-\gamma)+\theta}\right)^2 [\rho - A\alpha(1-\gamma)] > 0. \tag{5.22}$$

Thus the stronger the status preference is, the higher the economic growth rate is. Remembering (5.13), this result can be immediately understood; an increase in θ means that the agents become greedier for wealth accumulation, and it decreases the

[13] This can be verified as follows. From (5.14) the asymptotic growth rate of q is given by $[\alpha(1-\gamma)-1]\,g$. Since the asymptotic growth rate of W is given by g, the asymptotic growth rate of qW is represented as $\alpha(1-\gamma)\,g$. Thus, if $\alpha(1-\gamma)\,g < \rho$, the TVC is satisfied. Since g is given by (5.19), this inequality is reduced to $\alpha(1-\gamma)\,A < \rho$, which is equivalent to (5.21).

effective discount rate of agents.[14] Since the agents are assumed to be symmetric in this section, this decrease in the effective discount rate is common across the agents, and thereby raises the economic growth rate.

3 A Model with Asymmetric Agents

3.1 The Structure of the Extended Model

This section extends the basic model developed in the previous section to an asymmetric agents model. For simplicity, only the case of an economy consisting of a large number of two types of agents. The size of each agent is assumed to be the same. It is also assumed that not only the level of each person's wealth position but also each person's subjective discount rate ρ_i and the strength of status preference differ across the agents.

Let us denote the variables of person j with subscript ($j = 1, 2$). Then optimal conditions of agent j are

$$\alpha C_j^{\alpha(1-\gamma)-1} V_j \left(\frac{W_j}{\overline{W}}\right)^{\beta(1-\gamma)} = q_j, \tag{5.23}$$

$$-\dot{q}_j + \rho_j q_j = \frac{\beta V_j\left(\frac{W_j}{\overline{W}}\right)^{\beta(1-\gamma)-1} V_j'\left(\frac{W_j}{\overline{W}}\right) C_j^{\alpha(1-\gamma)}}{\overline{W}} + A q_j, \tag{5.24}$$

$$\lim_{t\to\infty} q_j W_j e^{-\rho_j t} = 0. \tag{5.25}$$

Taking the logarithms of both sides of (5.23) and differentiating them with respect to time, we have

$$\frac{\dot{q}_j}{q_j} = [\alpha(1-\gamma) - 1]\frac{\dot{C}_j}{C_j} + \beta(1-\gamma)\frac{V_j'}{V_j}\left(\frac{\dot{W}_j}{W_j} - \frac{\dot{\overline{W}}}{\overline{W}}\right). \tag{5.26}$$

Furthermore, combining (5.23) and (5.24) gives

$$-\frac{\dot{q}_j}{q_j} = A - \rho_j + \beta V_j'\left(W_j/\overline{W}\right) C_j/\alpha V_j\left(W_j/\overline{W}\right)\overline{W}, \tag{5.27}$$

which is the modified Euler equation.

[14] More precisely, if the change in the long run value of (C/K) is small relative to the change in θ, this statement is valid.

Next consider the market equilibrium conditions. In market equilibrium, the following two equations hold.

$$\sum_j W_j = \sum_j K_j = \overline{W} = K, \tag{5.28}$$

$$\dot{K} = AK - \sum_j C_j. \tag{5.29}$$

Equations (5.28) and (5.29) are the asset market equilibrium condition and the goods market equilibrium condition, respectively.

3.2 Balanced Growth Equilibrium

In the balanced growth equilibrium the following relations must be satisfied if the equilibrium exits:

$$\frac{\dot{W}_j}{W_j} = \frac{\dot{\overline{W}}}{\overline{W}} = \frac{\dot{K}}{K} = \frac{\dot{C}_j}{C_j} = \mathrm{g}. \tag{5.30}$$

Defining new variables as:

$$\varepsilon_j \equiv \frac{W_j}{\overline{W}}, \text{ and } x_j \equiv \frac{C_j}{K}, \tag{5.31}$$

and combining (5.26), (5.27), (5.30) and (5.31) gives

$$[1 - \alpha(1 - \gamma)]\,\mathrm{g} = A - \rho_j + \beta V_j'\left(\varepsilon_j\right) x_j / \alpha V_j\left(\varepsilon_j\right),\ (j = 1, 2) \tag{5.32}$$

Furthermore, from (5.28), (5.30) and (5.31) we have

$$\sum_j \varepsilon_j = 1, \tag{5.33}$$

$$\mathrm{g}\varepsilon_j = A\varepsilon_j - x_j,\ (j = 1, 2) \tag{5.34}$$

Five (5.32), (5.33) and (5.34) determine five values of endogenous variables on the balanced growth path; x_j^*, ε_j^* ($j = 1, 2$), and g*.

Let us define a function $\mu_j(\varepsilon_j)$ as follows

$$\mu_j\left(\varepsilon_j\right) \equiv \frac{V_j'\left(\varepsilon_j\right)\varepsilon_j}{V_j\left(\varepsilon_j\right)}\,(> 0).$$

That is, μ_j represents the elasticity of utility derived from status for agent j. Substituting (5.32) into (5.34) gives the following two equations

$$g_j = A + \frac{A\alpha(1-\gamma) - \rho_j}{1 - \alpha(1-\gamma) + \beta\mu_j(\epsilon_j)/\alpha}, \quad (j = 1, 2), \tag{5.35}$$

Define the right hand side of (5.35) as $g_j(\epsilon_j)$ ($j = 1, 2$). Differentiating g_j with respect to ε_j yields

$$\frac{dg_j}{d\varepsilon_j} = \left(\frac{\rho_j - A\alpha(1-\gamma)}{1 - \alpha(1-\gamma) + \beta\mu_j(\varepsilon_j)/\alpha}\right)\frac{d\mu_j}{d\varepsilon_j}. \tag{5.36}$$

Hence the sign of $dg/d\varepsilon_j$ is the same as that of $d\mu_j/d\varepsilon_j$ since the sign of the term inside the bracket is positive.[15] Let us assume that the sign of $d\mu_j/d\varepsilon_j$ does not change as ε_j varies.

Without loss of generality, we assume that $\rho_1 < \rho_2$, that is, agent 1 is more patient than agent 2. As is proved in the appendix, when $\frac{dg_1}{d\varepsilon_1} + \frac{dg_2}{d\varepsilon_2} < 0$ is satisfied, the balanced growth equilibrium is stable. The intuition of the result can be interpreted as follows. Assume that $\frac{dg_i}{d\varepsilon_i} > 0$ (i.e., at $\frac{d\mu_i}{d\varepsilon_i} > 0$) and the magnitude is larger than the other's. Since the effective discount rate is given by $\rho_i - \beta\mu_i(\varepsilon_i)/\alpha$, when the share of wealth becomes greater, the agent accumulates more. Then, the agent's relative wealth position becomes higher, and that the effective discount rate of the agent becomes lower. This decrease in the effective discount rate induces more accumulation of wealth, implying explosive movements of wealth accumulation.

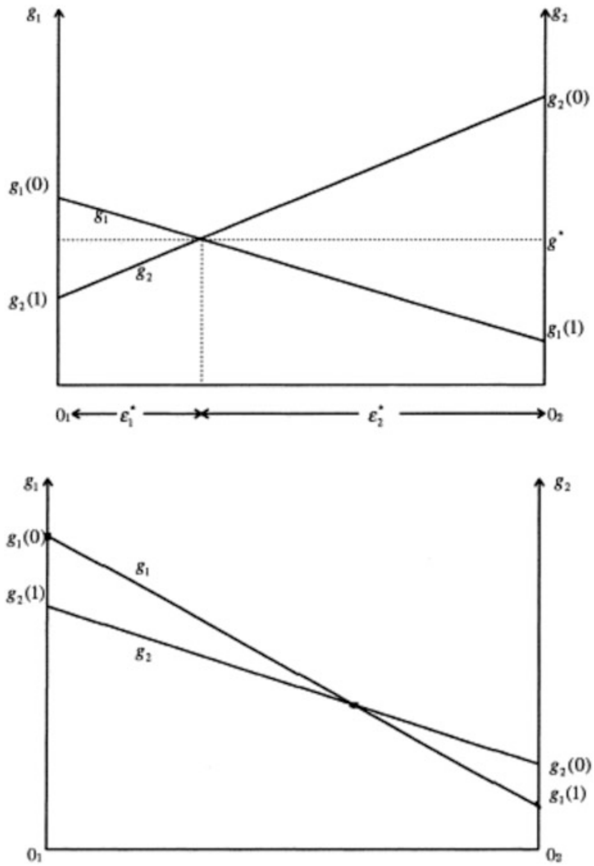
This result is analogous to the endogenous time preference model developed by Epstein and Hynes. In their model the subjective discount rate is assumed to be a function of the level of consumption. They show that the equilibrium is stable only when the function is increasing with respect to consumption. In our model, this corresponds to the case where the effective discount rate is an increasing function of the share of wealth.

3.3 Steady State Wealth Distribution

Let us now examine the wealth distribution in the steady state. In the following analysis, we treat only the cases where the stability condition presented above holds. We can depict the graphs of $g_1(\varepsilon)$ and $g_2(\varepsilon)$ as in Fig. 5.2. Figure 5.2a corresponds to the case where the signs for both values of $\frac{d\mu_j}{d\varepsilon_j}$ ($j = 1, 2$) are negative. On the other hand, Fig. 5.2b corresponds to the case where only one of the signs of $\frac{d\mu_j}{d\varepsilon_j}$ is negative.

[15]Remember that $1 - \alpha(1-\gamma) > 0$, $\rho - A\alpha(1-\gamma) > 0$ and $\mu_j > 0$.

Fig. 5.2 The balanced growth equilibrium



As can be seen from these figures, if the following inequalities hold, then the balanced growth equilibrium exists

$$g_1(0) > g_2(1),\ g_2(0) > g_1(1).$$

Subtracting $g_1(\varepsilon)$ from $g_2(\varepsilon)$, we obtain

$$g_1(\varepsilon) - g_2(\varepsilon) = \frac{A\beta(1-\gamma)(\mu_1-\mu_2) + [1-\alpha(1-\gamma)](\rho_1-\rho_2) + \frac{\beta(\mu_1\mu_2-\rho_2\mu_1)}{\alpha}}{\left[1-\alpha(1-\gamma)+\frac{\beta\mu_1}{\alpha}\right]\left[1-\alpha(1-\gamma)+\frac{\beta\mu_2}{\alpha}\right]}.$$

Therefore, even if ρ_1 is smaller than ρ_2, when μ_2 is larger than μ_1, the case in which inequality $g_2(\varepsilon) > g_1(\varepsilon)$ holds could exist since $\rho_j > A\alpha(1-\gamma)$. Figure 5.2a depicts this situation. In this case, we find

$$\varepsilon_1^* < \varepsilon_2^*.$$

Thus, if the status preference of agent 2 is stronger than that of agent 1, then agent 2, that has the higher subjective discount rate, can hold a larger share of world wealth on the balanced growth path. In other words, even if an agent is less patient than other agents, the agent can posses larger wealth than others when the agent is much greedier.

It is also worthwhile emphasizing that even if the subjective discount rates differ across agents, there exists a unique balanced growth path on which each agent holds a positive share of the world wealth. These results are summarized as:

Proposition 1 *Even if the subjective discount rates differ across agents, there can exist a unique steady state path with constant growth on which each agent holds a positive share of the world wealth.*

Proposition 2 *Even if $\rho_1 < \rho_2$, the steady state allocation of wealth can be $\varepsilon_1 < \varepsilon_2$.*

Using a model without status preference, Ramsey (1928) and Becker (1980) show that if the subjective discount rate differs across agents, the most patient agent eventually holds all the nonhuman wealth. Here, we show that in a model incorporating status preference, all agents can hold nonhuman wealth, and the agent who is less patient can hold larger nonhuman wealth. When agents have status preference, as assumed in this chapter, even agents with a high discount rate could continue to hold a positive share of the world wealth. This is because their utility from their relative wealth position decreases unless they catch up with more patient agents.

Constructing a dynamic general equilibrium model with consumption externalities, Ikeda (1993, 1995) shows the possibility that the less patient agent finally possesses all the nonhuman wealth when consumption of one agent reduces the utility of the other agent (the snob effect case). In his model, consumption externalities modify the subjective discount rate in a way similar to status in our model, so that the two analyses generate similar conclusions.

3.4 Stronger Status Preference May Depress Long Run Growth

Finally, consider the effects of changes in ρ_i and μ_i. These changes represent changes in the strength of saving motives. A decrease in ρ_i implies that agent i cares about his future life more seriously, and thus his savings incentives become higher. A higher value of μ_i. means a higher status preference and thus a strong motive to save more than others. It is easily seen that an increase in μ_i. or a decrease in ρ_i shifts up (5.35). Thus the effects can be depicted as in Figs. 5.3, 5.4, and 5.5. In the cases of Figs. 5.3 and 5.5, when the agent becomes more patient or attaches more weight to his status, the growth rate of the economy increases and his share of nonhuman wealth rises. This result is basically the same as that in the homogenous agents case analyzed in the previous section. In contrast, in the case of Fig. 5.4 the effect of an increase in savings incentives is rather different. In this case, when

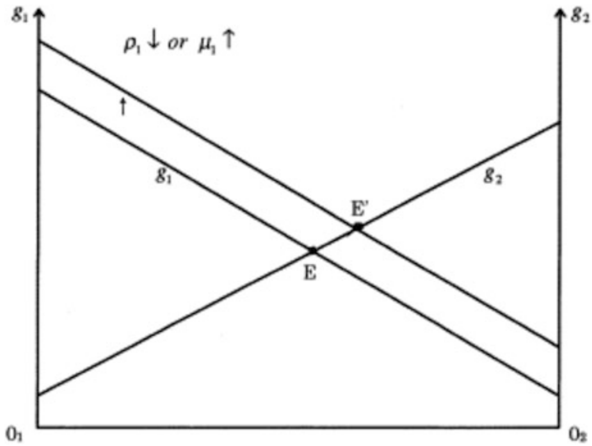


Fig. 5.3 Stronger status preference enhances long run growth

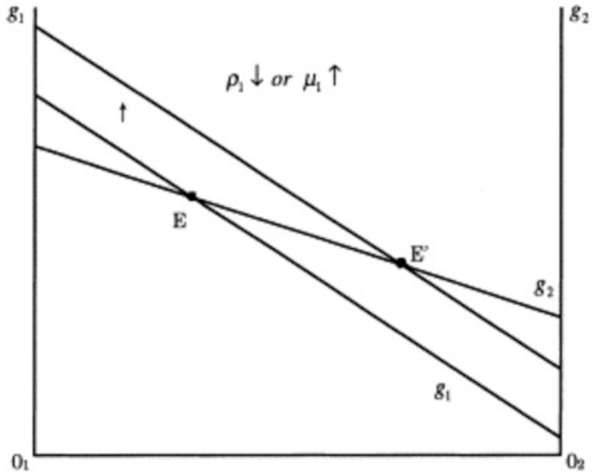


Fig. 5.4 Stronger status preference depresses long run growth

the time horizon of agent 1 becomes longer and agent 1s desire to differentiate his wealth position relative to the other agent becomes stronger, the growth rate of the economy declines. Therefore we have the following proposition:

Proposition 3 *Suppose that* $\mu_1'(\varepsilon_1) < 0$ and $\mu_2'(\varepsilon_2) > 0$. *Then, the equilibrium growth rate is reduced when agent* 1 *increases its savings incentives.*

The intuition behind this result is as follows. An impact effect of an increase in agent 1s saving incentives raises its wealth share ε_1 and worsens agent 2s wealth position. Thus, it lowers both values of $\mu_i(\varepsilon_i)$ $(i = 1{,}2)$. From (5.32), (5.34) and the definition of $\mu_i(\varepsilon_i)$ $(i = 1{,}2)$, it is easily seen that a decline in the value of $\mu_i(\varepsilon_i)$ has the same effect on the growth rate as an increase in the subjective discount rate. Now, since both agents' "modified" discount rates is raised by an increase in agent 1s saving incentives, the rate of capital accumulation as a whole is reduced, resulting

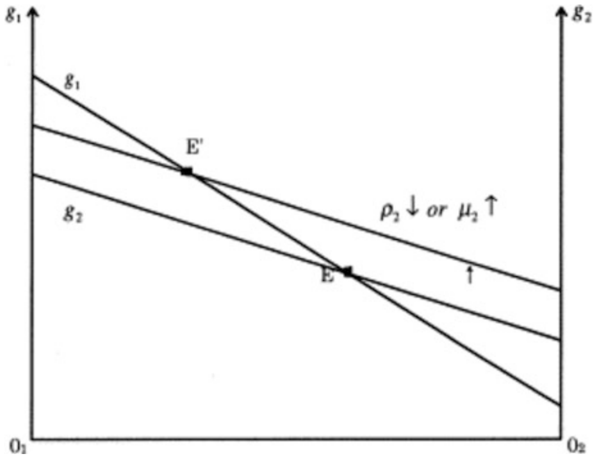


Fig. 5.5 Stronger status preference enhances long run growth

in a reduction in the long run growth rate. This result can be regarded as an example of fallacy of composition in a dynamic general equilibrium model. However, it is worthwhile stressing that Proposition 3 is obtained in the framework of optimizing agents with rational expectations. Propositions 2 and 3 show that, in analyzing the growth process of economies, considering the heterogeneity of agents is important since these results cannot be obtained in the representative agent framework.[16]

4 Concluding Remarks

The results obtained in this chapter can be extended to many directions. First, we can easily apply our model to international economy settings. For example, the two agents could be interpreted as two countries called 1 and 2. Suppose that both countries' governments impose a residence tax on capital income. In this case, a difference in the residence tax rates would have a similar effect to a difference in the subjective discount rates between countries. A residence tax is levied on income of residents regardless of its source. The residents in country ($j = 1, 2$) whose tax rate is τ_j faces the after tax interest rate $\left(1 - \tau_j\right) A$ instead of A.[17] Thus, from (5.35) it is easily understood that the implication of alarger value of τ_j is similar to that of a

[16]Using endogenous growth models with production-side externalities, Tamura (1991), Torvik (1993) and Benabou (1996) also investigate how the heterogeneity of agents affects the rate of long run growth.

[17]See, for example, Iwamoto and Shibata (1991) for a detailed explanation on the functioning of residence tax.

higher subjective discount rate. Proposition 3 implies that when a country raises its residence tax rate, the long run growth rate may increase contrary to the standard prediction.[18]

Furthermore, when each agent is relatively large the agent would recognize that its own and other's wealth accumulation have externalities. In this situation, the interaction among agents should be formulated as a dynamic game, as in Shibata (1996). Under this setting, depending applied equilibrium concepts such as open-loop Nash and closed-loop Nash equilibria, the equilibrium dynamics and the steady state allocation of wealth among agents would exhibit various patterns. These are interesting questions for future research.

Acknowledgments An earlier version of this chapter was presented at the 1994 Southeastern International Economics Meeting held at the University of Virginia. The authors are deeply indebted to a referee of this journal for very constructive comments. They also acknowledge E. K. Choi, E. Dinopoulos, H. Hayakawa, N. Okuno, Y. Ono, J. Raymo, Y. Seoka, T. Yagi, A. Yakita and seminar participants at Osaka City University, Yokohama National University, Nagoya University and Tokyo Institute of Technology for their helpful comments and suggestions.

Appendix: Stability of the Differential Equation System

In this appendix the stability of the model in Sect. 3 is examined. Let us first summarize the dynamics of the model. From the definition of ε_j, (5.31), we obtain the following equations:

$$\frac{\dot{\varepsilon}_j}{\epsilon_j} = \frac{\dot{W}_j}{W_j} - \frac{\dot{\overline{W}}}{\overline{W}} = x_j \left(1 - \varepsilon_j^{-1}\right) + x_i, \ (j \neq i, \ j = 1, 2)$$

Since $\varepsilon_1 + \varepsilon_2 = 1$, one of these two equations is used to obtain the fundamental dynamic equations. On the other hand, from (5.26), (5.27) and the definition of x_j, the following equations are derived:

$$\frac{\dot{x}_j}{x_j} = \frac{\beta}{\delta} v_j' \left[(1-\gamma) \left\{ x_j \left(1 - \varepsilon_j^{-1}\right) + x_i \right\} + \frac{x_j}{\alpha} \right] + x_j + x_i - A + \frac{A - \rho_j}{\delta}, \ j = 1, 2$$

where $\delta = 1 - \alpha(1-\gamma)$ and $v_j' \equiv V_j'/V_j$.

In order to examine the stability of this system, linearizing above equations around the balanced growth equilibrium, we obtain the following system of differential equations:

[18] Uhlig and Yanagawa (1996) also show the possibility that an increase in capital income tax leads to faster growth by using an overlapping generations model with production externalities. However, the underlying mechanism is basically different from ours.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{\varepsilon}_1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 1-\varepsilon_1^{-1} & 1 & x_1/\varepsilon_1^2 \end{bmatrix} \begin{bmatrix} x_1 - x_1^* \\ x_2 - x_2^* \\ \varepsilon_1 - \varepsilon_1^* \end{bmatrix}. \tag{5.37}$$

In the coefficient matrix, each element is defined as:

$$a_{11} = \frac{\beta}{\delta} v_1 \left[(1-\gamma)\left(1-\varepsilon_j^{-1}\right) + \frac{1}{\alpha} \right] + 1, \ \ a_{12} = \frac{\beta}{\delta} v_1 (1-\gamma) + 1,$$

$$a_{13} = \frac{\beta}{\delta} v_1' B_1 + \frac{\beta}{\delta} v_1 (1-\gamma) \frac{x_1}{\varepsilon_1^2}, \ \ a_{12} = \frac{\beta}{\delta} v_2' (1-\gamma) + 1,$$

$$a_{22} = \frac{\beta}{\delta} v_2 \left\{ (1-\gamma)\left(1-\varepsilon_2^{-1}\right) + \frac{1}{\alpha} \right\}, \ \ a_{23} = -\frac{\beta}{\delta} v_2' B_2 - \frac{\beta}{\delta} v_2 (1-\gamma) \frac{x_2}{\varepsilon_2^2},$$

$$B_j \equiv (1-\gamma) \left\{ x_j \left(1-\varepsilon_j^{-1}\right) + x_i \right\} + x_j/\alpha, \ \ v_j \equiv v_j\left(\varepsilon_j\right) \ (j=1,2).$$

We can prove the following lemma concerning the stability of this three-dimensional linear differential equation system.

Lemma *Suppose that* $\frac{dg_1}{d\varepsilon_1} + \frac{dg_2}{d\varepsilon_2} < 0$. *Then the characteristic equation of the coefficient matrix of the above linearized differential equation system has at least one negative eigen value.*

Proof The determinant of the coefficient matrix of differential equations is given by

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 1-\varepsilon_1^{-1} & 1 & x_1/\varepsilon_1^2 \end{vmatrix}.$$

The definitions of a_{ij} are presented below (5.37). Since $A - g = x_1/\varepsilon_1 = x_2/\varepsilon_2$ on the balanced growth equilibrium path, we have

$$\begin{vmatrix} a_{11} & a_{12} & \frac{\beta v_1}{\delta \varepsilon_1} x_1 \left[\frac{\eta_1}{\alpha} + \frac{1-\gamma}{\varepsilon_1} \right] \\ a_{21} & a_{22} & -\frac{\beta v_2}{\delta \varepsilon_2} x_2 \left[\frac{\eta_2}{\alpha} + \frac{1-\gamma}{\varepsilon_2} \right] \\ 1-\varepsilon_1^{-1} & 1 & x_1/\varepsilon_1^2 \end{vmatrix},$$

where $\eta_j \equiv \varepsilon_j v_j'/v_j$. Subtracting the second column from the first column and using the relation $x_1/\varepsilon_1 = x_2/\varepsilon_2$, we can rearrange the determinant as follows:

$$\frac{x_1}{\varepsilon_1} \begin{vmatrix} \frac{\beta}{\delta} v_1 \left\{ \frac{1}{\alpha} - \frac{1-\gamma}{\varepsilon_1} \right\} & \frac{\beta}{\delta} v_1 (1-\gamma) + 1 & \frac{\beta v_1}{\delta} \left[\frac{\eta_1}{\alpha} + \frac{1-\gamma}{\varepsilon_1} \right] \\ -\frac{\beta}{\delta} v_2 \left\{ \frac{1}{\alpha} - \frac{1-\gamma}{\varepsilon_2} \right\} & \frac{\beta}{\delta} v_2 \left\{ (1-\gamma)\left(1-\varepsilon_2^{-1}\right) + \frac{1}{\alpha} \right\} + 1 & -\frac{\beta}{\delta} v_2 \left[\frac{\eta_2}{\alpha} + \frac{1-\gamma}{\varepsilon_2} \right] \\ -\varepsilon_1^{-1} & 1 & -\varepsilon_1^{-1} \end{vmatrix}$$

$$= \frac{\beta v_1}{\delta \varepsilon_1} \left\{ \frac{\beta}{\delta} v_2 \left(-(1-\gamma) \frac{\varepsilon_1}{\varepsilon_2} + \frac{1}{\alpha} \right) + 1 \right\} \frac{1+\eta_1}{\alpha} + \left\{ \frac{\beta}{\delta} v_1 (1-\gamma) + 1 \right\}$$

$$\frac{\beta}{\delta} \frac{v_2 (1+\eta_2)}{\varepsilon_1 \alpha}$$

$$+ \left(\frac{\beta}{\delta} \right)^2 v_1 v_2 \left\{ \left(\frac{1}{\alpha} - \frac{1-\gamma}{\varepsilon_1} \right) \left(\frac{\eta_2}{\alpha} + \frac{1-\gamma}{\varepsilon_2} \right) - \left(\frac{1}{\alpha} - \frac{1-\gamma}{\varepsilon_2} \right) \left(\frac{\eta_1}{\alpha} + \frac{1-\gamma}{\varepsilon_1} \right) \right\}$$

$$= \frac{\beta}{\delta} \frac{1+\eta_1}{\varepsilon_1} v_1 \left(\frac{\beta}{\delta\alpha} v_2 + 1 \right) + \frac{\beta}{\delta\alpha} v_2 (1+\eta_2) + \left(\frac{\beta}{\delta\alpha} \right)^2 v_1 v_2 (\eta_2 - \eta_1)$$

$$= \frac{\beta}{\delta\alpha\varepsilon_1} \left[\frac{\beta}{\delta\alpha} v_1 v_2 \varepsilon_2 (1+\eta_1) + \varepsilon_1 (1+\eta_2) + v_1 (1+\eta_1) + v_2 (1+\eta_2) \right].$$

Since $\frac{d\mu_j}{d\varepsilon_j} = \left(\frac{V_j'}{V_j} \right) \left[1 + \varepsilon_j \left(\frac{V_j''}{V_j'} - \frac{V_j'}{V_j} \right) \right] = v_j \left(1 + \eta_j \right)$ and $\mu_j = v_j \varepsilon_j$, the sign of the determinant is the same as that of the following:

$$\left(\frac{\beta}{\delta\alpha} \mu_2 + 1 \right) \frac{d\mu_1}{d\varepsilon_1} + \left(\frac{\beta}{\delta\alpha} \mu_1 + 1 \right) \frac{d\mu_2}{d\varepsilon_2}.$$

From (5.36), we can rearrange this equation as follows

$$\frac{(\beta\mu_1/\alpha + 1)(\beta\mu_2/\alpha + 1)}{\delta (A - \mathrm{g})} \left(\frac{d\mathrm{g}_1}{d\varepsilon_1} + \frac{d\mathrm{g}_2}{d\varepsilon_2} \right).$$

Because the sign of δ is positive and $A > \mathrm{g}$ on the balanced growth path, the lemma holds. Q.E.D.

Since, in (5.37) there is one predetermined variable, ε_1, the balanced growth equilibrium is asymptotically stable. Under the assumption that $\gamma > 1$, it is easily proved, in the same way of Sect. 2, that on these convergent path the TVC of each agent is always satisfied. However, note that the perfect foresight equilibrium path may be either unique or multiple since we cannot identify the sign of the trace of the coefficient matrix.

Addendum[19]

Before millennium, models of status preference were constructed based on naïve formulations. Researchers constructed the models in simple and intuitive ways.

[19] This addendum has been newly written for this book chapter.

After millennium, Dupor and Liu (2003) defined new notions called *Keeping up with the Joneses* (KUJ) and *Running away from the Joneses* (RAJ).[20] Let us consider a utility function, $U(s, \bar{s})$. s stands for his own wealth and $\bar{s}$ for his reference point, for example, the average wealth of a society he lives. They first define two notions, jealousy and sympathy. If $\frac{\partial U}{\partial \bar{s}} < 0$, preferences exhibit jealousy. On the other hand, if $\frac{\partial U}{\partial \bar{s}} > 0$, preferences exhibit sympathy. These definitions are intuitively reasonable. They further define the two important definitions. If $\frac{\partial}{\partial \bar{s}}\left(\frac{\partial U}{\partial s}\right) > 0$, preferences exhibit *keeping up with the Joneses*. If $\frac{\partial}{\partial \bar{s}}\left(\frac{\partial U}{\partial s}\right) < 0$, preferences exhibit *running away from the Joneses*. Afterwards, researches on the status preference have been conducted based on these two important notions. For example, Kawamoto (2009) showed that changes of income inequality over time depend on whether people have KUJ preferences or RAJ preferences. In the KUJ economy, income inequality shrinks over time, whereas it expands in the RAJ economy.

We reconsider the results of Futagami and Shibata (1998) (hereafter FS) by using the notions, that is, KUJ and RAJ. It is obvious that FS's status function exhibits jealousy. We next consider whether FS's status function exhibits KUJ or RAJ. The cross derivative of the status function is given by

$$\frac{\partial}{\partial \overline{W}}\left[\frac{\partial}{\partial W_i}\left(V_i\left(\frac{W_i}{\overline{W}}\right)\right)\right] = -\frac{\left(1 + \frac{W_i}{\overline{W}}\frac{V_i''}{V_i'}\right)V_i'}{\overline{W}^2}.$$

Therefore, the following holds.

1. When $1 < -\frac{W_i}{\overline{W}}\frac{V_i''}{V_i'}$, the status function exhibits KUJ.
2. When $1 > -\frac{W_i}{\overline{W}}\frac{V_i''}{V_i'}$, the status function exhibits RAJ.

We next examine the sign of $\frac{d\mu_j}{d\epsilon_j}$. The derivative is given by

$$\frac{d\mu_j}{d\epsilon_j} = \frac{d}{d\epsilon_j}\left[\frac{V_j'(\epsilon_j)\epsilon_j}{V_j(\epsilon_j)}\right] = \frac{V_j'V_j\left[\left(1 + \frac{\epsilon_j V_j''}{V_j'}\right) - \frac{\epsilon_j V_j'}{V_j}\right]}{V_j^2}.$$

Thus, when the following inequality holds, the derivative takes a positive value, that is, $\frac{d\mu_j}{d\epsilon_j} > 0$.

$$\left(1 + \frac{\epsilon_j V_j''}{V_j'}\right) - \frac{\epsilon_j V_j'}{V_j} > 0.$$

[20]See Yamada and Sato (2013) for recent empirical analyses on this issue.

In FS's model, the paradoxical case can occur when $\frac{d\mu_1}{d\epsilon_1} > 0$ and $\frac{d\mu_2}{d\epsilon_2} < 0$ hold. Only if $1 + \frac{\epsilon_1 V_1''}{V_1'} > 0$, that is, the preference of agent type 1 exhibits RAJ, $\frac{d\mu_1}{d\epsilon_1} > 0$ holds. On the other hand, when the preference of agent type 2 exhibits KUJ, $\frac{d\mu_2}{d\epsilon_2} < 0$ holds. Even if the preference of agent type 2 exhibits RAJ, the paradoxical case obtains when $\left(1 + \frac{\epsilon_2 V_2''}{V_2'}\right) - \frac{\epsilon_2 V_2'}{V_2} < 0$. Thus, when the paradoxical case obtains, the preference of one type of the agents at least exhibits RAJ. Therefore, the notions, KUJ and RAJ play an important role in FS's model.

References

Abel AB (1990) Asset prices under habit formation and catching up with the Joneses. Am Econ Rev 80:43–47

Bakshi GS, Chen Z (1996) The spirit of capitalism and stock-market prices. Am Econ Rev 86:133–157

Becker RA (1980) On the long-run steady state in a simple dynamic model of equilibrium with heterogeneous households. Q J Econ 95:375–382

Benabou R (1996) Heterogeneity, stratification, and growth: macroeconomic implications of community structure and school finance. Am Econ Rev 86:584–609

Caballero RJ (1988) Intertemporal substitution and consumption fluctuation, Columbia University discussion paper no 409

Cole HL, Mailath GJ, Postlewaite A (1992) Social norms saving behavior, and growth. J Polit Econ 100:1092–1125

Cole HL, Mailath GJ, Postlewaite A (1995) Incorporating concern for relative wealth into economic models, Federal Reserve Bank of Minneapolis Quarterly Review, Summer, 12–21

Corneo G, Jeanne O (1995) Status-seeking can generate long-run growth in the Solow-Cass Model, University of Bonn, discussion paper no A497

Corneo G, Jeanne O (1997a) Social organization in an endogenous growth, mimeo, University of Bonn and ENPC-CERAS

Corneo G, Jeanne O (1997b) Pecuniary emulation inequality, and growth. Eur Econ Rev 43:1665–1678

Duesenberry JS (1949) Income, saving, and the theory of consumer behavior. Harvard University Press, Cambridge, MA

Dupor B, Liu W-F (2003) Jealousy and equilibrium overconsumption. Am Econ Rev 93:423–428

Epstein LG, Hynes AJ (1993) The rate of time preference and dynamic economic analysis. J Polit Econ 91:611–635

Fershtman C, Weiss Y (1993) Social status culture and economic performance. Econ J 103:946–959

Fershtman C, Murphy KM, Weiss Y (1996) Social status education, and growth. J Polit Econ 104:108–132

Frank RH (1985) The demand for unobservable and other nonpositional goods. Am Econ Rev 75:101–116

Futagami K, Shibata A (1998) Keeping one step ahead of the Joneses: status, the distribution of wealth, and long run growth. J Econ Behav Organ 36:109–126

Gali J (1994) Keeping up with the Joneses: consumption externalities portfolio choice, and asset prices. J Money Credit Bank 26:1–8

Gylfason T (1993) Interest rates optimal saving, and endogenous growth. Scand J Econ 95:517–533

Hall RE (1988) Intertemporal substitution in consumption. J Polit Econ 96:339–357

Hume D (1978) In: Selby-Bigge LA, Neddich P (eds) A treatise of human nature, 2nd edn. Clarendon, London
Ikeda S (1993) Interdependent preferences and intertemporal equilibrium, mimeo, Osaka University
Ikeda S (1995) Time preference, intertemporal substitution, and dynamics under consumer interdependence, ISER discussion paper no 386, Osaka University
Iwamoto Y, Shibata A (1991) Capital income taxation and the current account in a small open economy. J Int Money Financ 10:480–496
Kawamoto K (2009) Status-seeking behavior, the evolution of income inequality, and growth. Econ Theory 39: 269–289
Kruz M (1968) Optimal economic growth and wealth effects. Int Econ Rev 9:348–357
Marx K (1964) Economic and philosophic manuscripts of 1844 (trans: Milligan M). International Publishers, New York
Ono Y (1994) Money, interest, and stagnation: dynamic theory and Keynes's economics. Oxford University Press, Oxford
Ramsey F (1928) A mathematical theory of saving. Econ J 38:543–559
Robson AJ (1992) Status, the distribution of wealth private and social attitudes to risk. Econometrica 60:837–857
Romer PM (1986) Increasing returns and long run growth. J Polit Econ 94:1002–1037
Shibata A (1996) Patterns of growth in a dynamic game model with infrastructure capital, mimeo, Kyoto University
Tamura R (1991) Income convergence in an endogenous growth model. J Polit Econ 99:522–540
Torvik R (1993) Talent growth and income distribution. Scand J Econ 95:581–596
Uhlig H, Yanagawa N (1996) Increasing the capital income tax may lead to faster growth. Eur Econ Rev 40:1521–1540
Yamada K, Sato M (2013) Another avenue for anatomy of income comparisons: evidence from hypothetical choice experiments. J Econ Behav Organ 89:35–57
Zou Heng-fu (1995) The spirit of capitalism' and long run growth. Eur J Polit Econ 10:279–293

Chapter 6
Macroeconomic Implications of Conspicuous Consumption: A Sombartian Dynamic Model

Katsunori Yamada

Abstract This chapter presents a dynamic general equilibrium model in which consumers have status preference. I investigate the manner in which capital accumulation is impeded by conspicuous consumption à la Corneo and Jeanne (J Public Econ 66:55–71, 1997). Following the literature, social norms are given as either bandwagon type or snob type. I then show that when the economy is characterized by a bandwagon type social norm, capital accumulation exhibits interesting patterns. Those patterns include, for example, an oscillating convergence path: the rise of the economy feeds its decay through conspicuous consumption and that decay suppresses conspicuous consumption and engenders prosperity, as predicted by Sombart (Liebe, Luxus und Kapitalismus, 1912 (reprinted 1967)).

Keywords Status preference • Conspicuous consumption • Capital accumulation

JEL classifications E10; Z13; E30

1 Introduction

This chapter presents an investigation into how capital accumulation is impeded by conspicuous consumption in a framework of a dynamic general equilibrium model. In the model, motivation of conspicuous consumption is related to signaling of status, which is consistent with the original definition of conspicuous consumption by Veblen (1899). The framework of the signaling game is given by the microeconomic study of Corneo and Jeanne (1997a) (henceforth, CJ) and status utility is generated by *rank utility*. It is shown that the capital accumulation path might exhibit

The original article first appeared in the Journal of Economic Behavior & Organization 67: 322–337, 2008. A newly written addendum has been added to this book chapter.

K. Yamada (✉)
Faculty of Economics, Kindai University, Higashiosaka, Osaka, Japan
e-mail: kyamada@kindai.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_6

oscillating convergence as well as polarization caused by multiple equilibria when the social norm is of bandwagon type.

It has been well recognized that consumer decisions might not be explained adequately merely using utility derived from consumption. Rather, economic agents have been considered to derive additional utility from his social status, as highlighted, i.e., by Smith (1759), Hume (1739), and Veblen (1899). Many empirical analyses support the hypothesis of status preference through tests of Veblen's views of *pecuniary emulation*. Microeconomic evidence includes that presented by Clark and Oswald (1996), McBride (2001), Blanchflower and Oswald (2004), and Saito et al. (2005), among others. Basmann et al. (1988), Hayes et al. (1988, 1992), and Slottje (1992) provide evidence for Veblen's view of *conspicuous consumption* with data of the U.S., North America, European countries, and Japan, respectively. Frijters and Leigh (2008) examine conspicuous leisure in the U.S. and find strong effects of conspicuous leisure on work hours.

Furthermore, quantitative approaches also exist to test the validity of status preference such as Abel (1990), Gali (1994), and Bakshi and Chen (1996), who argue that observed asset price volatility can be explained through inclusion of the motivation of *keeping up with the Joneses*. Then, it seems that status preference has many consequences at the macroeconomic level as well as on individual decision problems.

Along with the empirical studies mentioned above, theoretical studies have examined how external effects of status preference affect the equilibrium, and especially, the economic growth rate. These studies propose that analyses with status preference might engender markedly different outcomes from those obtained through analyses in which utility depends only on consumption.[1]

In contrast to previous macroeconomic studies of the literature such as Cole et al. (1992), Konrad (1992), Zou (1994), Corneo and Jeanne (1997b), and Futagami and Shibata (1998), who investigate effects of status preference among consumers on the economic growth rate, this chapter analyzes disturbing effects of status preference through *conspicuous consumption* on the accumulation of capital. I use a model that has the same property as the well-known Solow model to extract effects of relative concern with clarity. This is the same strategy as that used by Konrad (1992), Zou (1994), and Corneo and Jeanne (1997b), who based their analyses on the well-known Ramsey economy.[2]

The intuition underlying the analysis is as follows. When status utility is obtained by *conspicuous consumption*, this type of preference might cause serious distortion in capital accumulation because conspicuous consumption behavior, by Veblen's definition, works effectively when the resource is devoted to economically meaningless activities. Hence, as argued by Frank (1985), conspicuous consumption

---

[1] See, for comprehensive surveys and discussions of theoretical studies of social status: Hayakawa (2000), Easterlin (2001), and Hollander (2001).

[2] Although the propensity to save in the model is constant, as it is in the Solow model, here is a micro-foundation to explain it.

might cause serious inefficiencies in the form of downward distortions in demand for non-conspicuous goods. In turn, this distortion might badly affect accumulation of an economically valuable good, i.e. productive capital.[3]

A simple and tractable model with conspicuous consumption behavior is constructed to clarify dynamic implications of this distortion. The model has the same properties as the textbook Solow model if there is no status utility. I introduce status preference by adopting the framework of CJ; fundamentally, this chapter can be considered as a first attempt to extend the static work of CJ into a two-sector general equilibrium model with dynamics.[4] I note here that this chapter is a first attempt of dynamic analysis of conspicuous consumption because full analytical solutions require to consider a special case in which the preference function for *substantial expenditure* converges to linearity. By this condition, I will neglect a level effect of changing wealth distribution on economic decisions by consumers.

It will be shown that the conspicuous consumption motivation can cause rich types of capital accumulation paths with the model. I propose that the model has consequences related to a prediction of Sombart (1912) and a convergence controversy summarized, for example, in Galor (1996). The rise of the economy might feed decay through conspicuous consumption and the decay might suppress conspicuous consumption and bring prosperity, as predicted by Sombart. Simultaneously, quests for social status can engender polarization of two economies with same economic fundamentals, initial conditions, and social norms.

This chapter is organized as follows. The next section describes the model. Explanations of CJ economy are included. Section 3 analyzes the equilibrium paths. Section 4 contains discussions of the relevance of the model to the modern economy. The last section concludes the chapter.

2 The Model

Time is discrete and extends to infinity. The economy is one in which people derive utility from social status as well as from consumption of intrinsic goods and granting bequests. The economy is populated by a continuum of agents who belong to

---

[3]It is intuitively plausible that wealth accumulation is impeded by human vanity; history shows that the rise and fall of aristocratic lineages might be explained by rat races in a quest for ever greater social status. For a comprehensive survey of human vanity and desire for social status, especially among the aristocratic class and wealthy merchants, see the third chapter of Sombart (1912).

[4]One feature of the model is that I illustrate the status good as a marketable good, whereas Cole et al. (1992, 1995) stress that the analyses of status utility are meaningful when utility from some non-market action has consequences related to market decisions. Nevertheless, I adopt the present strategy following work by Bagwell and Bernheim (1996), Corneo and Jeanne (1997a,c), and Becker (1991), in which the demand for conspicuous goods is determined in the market in accordance with some social norms.

lineages; each agent lives for one period and has one offspring.[5] Each individual of the lineage has index $i \in [0, 1]$. Within the lineage, generations are connected by bequest motives. All lineages in this economy are homogeneous except for the amount of bequests inherited from ancestors, although it is assumed that no lineage is endowed with the same level of wealth in the initial period. Aside from inherited wealth, agents are endowed with one unit of labor, which is in-elastically supplied to production sectors, and earns the wage income.

2.1 Consumers and Social Status

Each agent's utility is given as

$$U_{i,t} = u(q_{i,t}) + \nu_{i,t},$$

where q is called substantial expenditure, which comprises consumption $c_{i,t}$ and bequest $\omega_{i,t+1}$. Also, ν reflects a social reward from conspicuous consumption and is determined endogenously. As is apparent in the equation, total utility consists of $u(q)$ and ν in an additively separable form. With respect to utility from consumption and bequests, I assume that standard assumptions hold: it is strictly concave and satisfies the Inada condition for each argument.

Following CJ, consumption of the conspicuous good is limited for only one unit: the conspicuous good is indivisible and agents cannot buy more than one unit. In the dynamic model, it is also assumed that agents cannot bequest the conspicuous goods to offspring and that they perish within the period.[6]

The information structure of this economy is as follows. Agents cannot see the level of wealth, the amount of consumption and the amount of bequest with each other. The relative wealth position of each lineage and economic decisions are private information. On the other hand, the way in which wealth is distributed in the population is socially known. I assume that a firm that produces the conspicuous goods has enough information to distinguish those who will be *rich* from those who will not. Information superiority of the firm is attributable to its infinite life span: agents live only one period whereas the firm lives forever. Lastly, purchasing behavior of conspicuous goods is observable among consumers and denotes $\delta_{i,t}$ by a dummy variable taking value one if agent i buys a conspicuous good at time t; otherwise, it is zero.

With this information structure, conspicuous consumption behavior provides, in a Bayesian way, some value to agents in the economy with special social norm. The

[5]Hence, no population growth occurs in the model.

[6]From the standpoint of Veblen's view, it is essential for conspicuous consumers that they themselves purchase conspicuous goods rather than inherit them from their ancestors. Abstention from conspicuous consumption by a member of a lineage will give an impression of family decay.

social norm adopted in CJ and the chapter is called *wealth rank utility*: agents in front of the norm are assumed to extract utility from their relative wealth position in the economy.[7] Because they cannot directly compare their wealth rank with each other because of imperfect information, they must signal their wealth to extract status utility. For that reason, they might spend for conspicuous goods, which is by definition a meaningless but observable activity. That is, agents spend for conspicuous goods merely because they want to advertise their wealth and extract rank utility rather than utility from consumption (and bequests).

Define here *the rank utility function* $a(\cdot)$, which maps the wealth rank in the society into rank utility.[8] Next, denote $\Omega_t(i) : [0,1] \to [0,1]$ by the function mapping family index i (at the beginning of time t) into wealth order at the time. If a lineage with index $k \in [0,1]$ is the wealthiest in the society, then $\Omega_t(k) = 0$. Hereafter, I consider the mapped family index $j (= \Omega_t(i))$ in the analysis below.

Because the wealth rank of a family is unobservable, status utility ν is obtainable as the expected rank utility, which is conditional on the observable purchasing behavior of the conspicuous good (δ = 0,1) as

$$\nu_t = \nu_t(\delta_{i,t}) = E[a(j) \mid \delta_{i,t}]. \tag{6.1}$$

By this formulation, $\nu(\cdot)$ reflects the collective value aspect of conspicuous consumption under the condition that the wealth position of each lineage is unobservable.

Now the optimal behavior of agents can be depicted. First, the budget constraint of an agent reads as

$$c_{i,t} + \omega_{i,t+1} + \delta_{i,t} p_t \leq w_t + (1 + r_t)\omega_{i,t},$$

where prices (the relative price of conspicuous good; p_t, the wage; w_t and the rental price of capital; r_t) are given competitively. Define here $y_{i,t} = w_t + (1 + r_t)\omega_{i,t}$ by the total wealth for agent i at time t.

Agents determine their purchasing behaviors with respect to the conspicuous good such that the decisions are optimal, given their inferences related to their social status. With the additively separable utility function, this condition can be written as

$$\begin{aligned} \nu_t(1) - \nu_t(0) &= E[a(j) \mid 1] - E[a(j) \mid 0] \\ &\geq u(y_{i,t}) - u(y_{i,t} - p_t). \end{aligned} \tag{6.2}$$

In the analysis below, $s_t \equiv v(1) - v(0)$ is designated as the signaling value of conspicuous goods at time t.

The value of conspicuous consumption, in turn, can be determined in the Bayesian way after purchasing decisions of all lineages as to the conspicuous good.

[7] See 2.1 in CJ for a more detailed discussion.

[8] It is assumed here that the rank utility function is continuous on the interval $[0,1]$ and monotonously decreasing, with a finite lower and upper bound.

Specifically, through Eq. (6.1), s_t is given by the difference of expected rank utility between the conspicuous consumers and others:

$$s_t = \sigma(J_t) = \frac{\int_0^{J_t} a(j)dj}{J_t} - \frac{\int_{J_t}^1 a(j)dj}{1-J_t}$$
$$= \frac{1}{J_t(1-J_t)} \int_0^J \{a(j) - \bar{a}\}dj, \tag{6.3}$$

where $\bar{a}$ is the average rank utility over the whole population and J_t is the number of conspicuous consumers.[9] $\sigma(.)$ is called the signaling value function and the signaling value of conspicuous consumption depends only on the number of conspicuous consumers.

Here are three remarks on the signaling value function. First, the social norm in the model economy is regulated by the shape of $\sigma(.)$. When $d\sigma(.)/dJ < 0$ holds, the market demand for a conspicuous good decreases because others are purchasing the same good; consequently, the economy is "snobbish". On the other hand, when $d\sigma(.)/dJ > 0$ holds, the demand increases because others are buying the same good. Therefore, the economy is a "bandwagon". Following the literature, I will consider two social norm types in the dynamic analyses below: snobbish and bandwagon.

Second, the shape of the signaling value function is determined by the shape of the rank utility function through Eq. (6.3).[10] That is, *any continuous and differentiable function $\sigma(.)$, $\cdots$, can be rationalized by a rank utility of the form* of

$$a(x) = \bar{a} + (1-2x)\sigma(x) + x(1-x)\sigma'(x),$$

where x denotes the wealth rank of a lineage.

It is also noteworthy that, if the rank utility function is time invariant, then the signaling value function is unchanged through time, which I assume throughout the chapter: the social norm is assumed to be fixed for a society. In the application below, the rank utility function, the social norm, is assumed to be quadratic. Consequently, the assumption gives linear $\sigma(.)$ functions (see the r.h.s. of Eqs. (6.14) and (6.15)).

Third, the domain of $\sigma(.)$ is considered. The function is defined on the open interval of $J \in (0, 1)$:the case of the pooling equilibrium (either $J = 0$ or $J = 1$) for the conspicuous good market can be excluded. When $J = 0$ and no conspicuous consumers exist, the economy merely reduces to the Solow regime (see Eq. (6.13)). Hence, there is no need to define the signaling value function when $J = 0$.

[9] Throughout this chapter I assume that s_t is strictly positive and finite.

[10] As to how the social norm is regulated by the rank utility, CJ establishes that the snobbish society is attributable to the convexity of $a(.)$, whereas the bandwagon economy corresponds to the case in which $a(.)$ is concave. For an explanation of the intuition behind the argument, refer to CJ, page 61.

Furthermore, it will be shown that Lemma 2 in the snobbish economy excludes the case in which everyone buys the conspicuous good; in the bandwagon economy, J might converge asymptotically to one but never equals one. All in all, $\sigma(.)$ is defined on $J \in (0, 1)$.

2.2 Producers

The economy has two production sectors: sector 1 produces the intrinsic good and sector 2 produces the conspicuous good. In sector 1, the technology is given by a homogeneous function with two inputs, capital and labor, as

$$Y_t = F(K_t, N_t).$$

Competitive prices are given as

$$\begin{cases} r_t = \frac{\partial F(K_t, N_t)}{\partial K_t} \\ w_t^c = \frac{\partial F(K_t, N_t)}{\partial N_t}, \end{cases}$$

where r_t is the rental price of capital and w_t^c is the wage paid in sector 1.

The production of conspicuous goods requires only a labor force.[11] For that reason, it is written as

$$\hat{Y}_t = G(n_t),$$

where n_t is the labor force supplied to sector 2 and the competitive wage is given as

$$w_t^p - p_t G'(n_t) = 0,$$

where w_t^p is the wage paid in sector 2 and p_t is the price of the conspicuous good.

In equilibrium, $\hat{Y}_t$ coincides with the number of conspicuous consumers J_t because each lineage can buy the conspicuous good with one unit.

2.3 Equilibrium

Now the equilibrium of this economy can be defined. The equilibrium at time t is the set of prices $(w_t^p, w_t^c, r_t, p_t, s_t)$, and the set of allocations $(c_{j,t}, \omega_{j,t+1}, \delta_{j,t})$ and (n_t, N_t, K_t) so that (i) given $(w_t^p, w_t^c, r_t, p_t, s_t)$, $(c_{j,t}, \omega_{j,t+1}, \delta_{j,t})$ solves the

[11]The assumption that production in sector 2 depends only on labor is not essential to the following analysis, but it dramatically simplifies the analysis.

consumers' maximization problems; (ii) given (w_t^p, w_t^c, r_t, p_t), (n_t, N_t, K_t) solves the firms' maximization problems; and (iii) the market for labor, capital and the conspicuous goods markets clear.

Here functional forms must be specified to obtain explicit solutions. For production technology of the intrinsic good, I assume that

$$Y_t = \theta K_t^{\beta} N_t^{1-\beta},$$

where $\theta > 0$ is a productivity parameter and β is the capital share. As for sector 2, I assume that

$$\hat{Y}_t = n_t.$$

Assuming competitiveness and an interior solution, the first-order condition of the firm in sector 2 gives $p_t = w_t^p$. In turn, the competitive market also implies that $w_t^p = w_t^c (\equiv w_t)$ must hold in equilibrium. As for the utility from substantial expenditure, I assume that[12]

$$u = u(q(c_{j,t}, \omega_{j,t+1})) = (c_{j,t}^{1-\alpha} \omega_{j,t+1}^{\alpha})^{\eta},$$
$$\text{where} \quad \alpha \in (0,1) \quad \text{and} \quad \eta \in (0,1).$$

Here, the bequest motive is given by the impure altruism form adopted by Banerjee and Newman (1993), among others, and $u(\cdot)$ is strictly concave in q.

From the specification described above, the prices are obtained as

$$\begin{cases} r_t = \frac{\partial Y_t}{\partial K_t} = \theta\beta \left(\frac{K_t}{N_t}\right)^{\beta-1}, \\ w_t = \frac{\partial Y_t}{\partial N_t} = \theta(1-\beta) \left(\frac{K_t}{N_t}\right)^{\beta}, \\ p_t = w_t = \theta(1-\beta) \left(\frac{K_t}{N_t}\right)^{\beta}. \end{cases} \tag{6.4}$$

In addition, the signaling value, s_t, is given by the signaling value functions discussed later.

Market-clearing conditions of labor, capital and the conspicuous good market are given respectively as

$$N_t + n_t = 1, \tag{6.5}$$

[12]Here I might consider that ω is a term representing a value from investment: agents are now assumed to live forever and to derive substantial utility from today's consumption and from consumption plans onward. With this interpretation, the details of a full-fledged inter-temporal decision problem with rational expectations can be avoided.

$$K_t = \int \omega_{j,t} dj, \tag{6.6}$$

$$J_t = n_t. \tag{6.7}$$

Furthermore, from Eq. (6.2), the conspicuous consumption goods market clears such that the marginal gain from conspicuous consumption and the marginal gain of substantial expenditure are equal for the lineage with index $j = J$[13]:

$$\nu_t(1) - \nu_t(0) = s_t = u(y_{J,t}) - u(y_{J,t} - p_t). \tag{6.8}$$

In order to obtain maximum analytical scope in the problem, I must consider the special case of $\lim \eta \to 1$. By this condition and the functional form of utility from the substantial expenditure, it is readily apparent that $u(y_{j,t}) - u(y_{j,t} - p_t)$ converges to $(1-\alpha)^{1-\alpha}\alpha^{\alpha} p_t$ irrespective of the lineage index j. It is not needed to know explicitly the level of total wealth that lineage J has: $y_{J,t}$. Thereby, the equilibrium condition given by Eq. (6.8) reduces to

$$(1-\alpha)^{1-\alpha}\alpha^{\alpha} p_t = s_t = \sigma(J_t). \tag{6.9}$$

As described above, I can choose the shape of $\sigma(.)$ through an appropriate choice of the rank utility function. In this chapter, the signaling value function $\sigma(.)$ is assumed to be linear on $(0, 1)$ (equivalently, the rank utility function is assumed to be quadratic). Moreover, here I designate a new function of $S(N)$ such that $S(N_t) = [(1-\alpha)^{1-\alpha}\alpha^{\alpha}]^{-1}\sigma(J_t)$ for all $N_t = 1 - n_t = 1 - J_t \in (0, 1)$.[14] This is the signaling value function defined on N with market clearing conditions. Hence, Eq. (6.9) reduces to the following equilibrium relationship between capital and the number of workers in sector 1 as

$$S(N_t) = p_t = \theta(1-\beta)\left(\frac{K_t}{N_t}\right)^{\beta}. \tag{6.10}$$

From Eq. (6.10), N_t (and J_t) will be determined for a given level of capital as

$$N_t = N(K_t) = 1 - J_t.$$

With the impure altruism utility function and the condition that $\eta \to 1$, the following optimal allocation rules for each lineage can be obtained:

[13] In the economy, every agent can afford to buy a conspicuous good and the solution is always interior: from the individual budget constraint, it can be seen that the budget is always satisfied because $p_t = w_t$ in equilibrium. By this property in the general equilibrium model, I can concentrate on the analysis of accumulation path of aggregate level of capital. This exhibits a contrast with CJ in which the corner solutions are possible.

[14] $S(\cdot)$ is continuous, linear, bounded on $(0, 1)$ and positive by construction.

$$\begin{cases} c_{j,t} = (1-\alpha)\{(1+r_t)\omega_{j,t} + w_t - \delta_{j,t} p_t\} \\ \omega_{j,t+1} = \alpha\{(1+r_t)\omega_{j,t} + w_t - \delta_{j,t} p_t\} \\ \delta_{j,t} = 1 \quad \text{iff} \quad j \in [0, J_t] \\ \delta_{j,t} = 0 \quad \text{iff} \quad j \in (J_t, 1]. \end{cases} \tag{6.11}$$

Equations (6.4)–(6.7), (6.10) and (6.11) determine the equilibrium of the economy at time t. Furthermore, the dynamics of capital are also obtained because the utility function allows to aggregate the optimal allocation rule of ω for all lineages $j \in [0, 1]$: the social resource constraint is given as

$$\begin{aligned} \int \omega_{j,t+1} dj &= \int \alpha\{(1+r_t)\omega_{j,t} + w_t - \delta_{j,t} p_t\} dj \\ &= \alpha\{(1+r_t) \int \omega_{j,t} dj + w_t - p_t J_t\}. \end{aligned}$$

Because $K_t \equiv \int \omega_{j,t} dj$, it is obtained that

$$\begin{aligned} K_{t+1} &= \Psi(K_t) \\ &= \alpha\{(1 + r(N(K_t), K_t))K_t + w(N(K_t), K_t) - p(N(K_t), K_t)J(K_t)\} \\ &= \alpha\{K_t + \theta K_t^{\beta} N(K_t)^{1-\beta}\}. \end{aligned} \tag{6.12}$$

It is noteworthy that if no status utility pertains in the economy so that $N = 1$ for all t, Eq. (6.12) reduces to

$$\begin{aligned} K_{t+1} &= \psi(K_t) \\ &= \alpha\{(1 + \theta\beta K_t^{\beta-1})K_t + \theta(1-\beta)K_t^{\beta}\} \\ &= \alpha\{K_t + \theta K_t^{\beta}\}. \end{aligned} \tag{6.13}$$

Apparently, the property of the dynamics in Eq. (6.13) is essentially identical to that of the Solow model. That is, Eq. (6.13) has a unique globally stationary point of $K^* = \left(\frac{\alpha\theta}{1-\alpha}\right)^{\frac{1}{1-\beta}}$ because $\frac{dK_{t+1}}{dK_t} > 0$, $\frac{d^2 K_{t+1}}{dK_t^2} < 0$, $\lim_{K_t \to 0} \frac{dK_{t+1}}{dK_t} = \infty$, $\lim_{K_t \to \infty} \frac{dK_{t+1}}{dK_t} = \alpha \in (0, 1)$ and $\frac{dK_{t+1}}{dK_t} |_{K=K^*}= \alpha + \beta - \alpha\beta \in (0, 1)$ when $\alpha \in (0, 1)$ and $\beta \in (0, 1)$.

On the other hand, when the conspicuous consumption motive is present in society, capital accumulation is disturbed by two causes and the dynamics are governed by Eq. (6.12). One cause is the outflow of labor force from sector 1 to sector 2, which results in the reduction of productivity of capital. The other is the resource devoted to economically meaningless activity: conspicuous consumption.

Before closing the discussion in this section, here is a remark on the strategy of $\eta \to 1$. When I do not consider the special case of $\eta \to 1$, then it is necessary that the dynamic evolutions of the total wealth of all the lineages in the economy be tracked because J_t and the other endogenous variables depend on the wealth distribution at time t. That proposition is, of course, analytically intractable. Because this study is intended to show analytically that conspicuous consumption motivation has the potential to affect the capital accumulation path in interesting ways, I proceed with the assumption throughout the chapter.

I note here that an important implication in CJ will become completely invalid when I consider the case of $\eta = 1$. In this situation, the situation renders it ambiguous who in the economy actually spends for the conspicuous good, although J can still be determined, as in the case of $\eta \to 1$. This is because the willingness to purchase the conspicuous good becomes the same regardless of the wealth level. That situation, in turn, indicates that agents who buy the conspicuous good cannot advertise that their wealth positions are high.

Nonetheless, I might be able to proceed innocuously in the analyses with $\eta = 1$ because I assume that the conspicuous good firm has all information related to the wealth of agents: I can assume that the firm sells only to the *rich*. The firm has information superiority because it lives forever and has some screening skills that are acquired. That is, although all agents might have the same level of willingness to buy the conspicuous good and can indeed afford to do so, those who purchase are those who are qualified by the firm. Although this strategy might seem to be an arbitrary device to allocate conspicuous goods among agents, it has a great merit that J and the other endogenous variables are calculable without information of wealth distribution, as in the case of $\eta \to 1$.

3 Dynamics

This section presents a description of the patterns of the equilibrium dynamics of the economy, given by Eqs. (6.12) and (6.13), for the snobbish economy (where $dS(\cdot)/dN > 0$) and the bandwagon economy (where $dS(\cdot)/dN < 0$).

3.1 Snobbish Economy

In the snobbish economy, $dS(\cdot)/dN > 0$ (equivalently, $d\sigma(\cdot)/dJ < 0$) holds because an agent feels reluctant to see others buy the conspicuous goods and abstains from it. A linear signaling value function defined on N with zero intercept, $S(N) = bN$ ($b > 0$ and $N \in (0, 1)$), provides analytical solutions.[15] Equilibrium

[15]Qualitative implications obtained below will be unchanged as long as $S(N)$ is increasing monotonously on $(0, 1)$ and N_t is uniquely determined for every K_t from Eq. (6.10).

conditions for conspicuous consumption given by Eq. (6.10) are rewritten as the following

$$bN_t = \theta(1-\beta)\left(\frac{K_t}{N_t}\right)^{\beta}. \tag{6.14}$$

It can be readily seen from Eq. (6.14) that $N_t \in (0, 1)$ is uniquely determined for a given K_t. Notice also that when K_t is sufficiently high so that N_t in Eq. (6.14) is greater than or equal to one, there is no inner solution for the conspicuous good market and that capital accumulation is governed by Eq. (6.13) rather than Eq. (6.12). Here define $\bar{K}$ by $\bar{K} = \{K \mid \lim_{N\to 1} S(N) = \lim_{N\to 1} \theta(1-\beta)K^{\beta}N^{-\beta}\}$ so that $\bar{K} = \left\{\frac{b}{\theta(1-\beta)}\right\}^{\frac{1}{\beta}}$. This is the level of K, above which the demand for the conspicuous good disappears.

A simple algebra shows that N_t and p_t is strictly increasing in K_t in the snobbish economy. This provides the next lemma.

Lemma 1 *The demand curve of the conspicuous good is downward sloping in the snobbish economy.*

The equilibrium can be interpreted as follows. When K is small, the price of the conspicuous good is low so that the opportunity cost of conspicuous consumption is small. A smaller opportunity cost directly implies a smaller signaling value, which in the snobbish economy a implies more numerous conspicuous consumers in equilibrium. As K increases in $(\underline{K}, \bar{K})$, the price of the conspicuous good rises so that the number of conspicuous consumers must decrease to compensate for higher opportunity cost.

Because the demand curve of the conspicuous good is downward sloping in the snobbish economy, an important proposition proposed by CJ applies. The Proposition 2.2 in CJ argues that the demand for the conspicuous good becomes negligible when the price of the conspicuous good is sufficiently low and the number of the conspicuous consumers approaches asymptotically to full population, in the economy where the demand curve of the conspicuous good is downward sloping. The following lemma replicates Proposition 2.2 in CJ.

Lemma 2 *There exists a price* $\hat{p} > 0$ *such that the demand is negligible if* $p < \hat{p}$.

Proof See the proof in CJ. *Q.E.D.*

In this study, such a low price of conspicuous goods will be realized when K approaches zero.[16] Denote $\underline{K}$ by the level of capital below which the demand for the conspicuous good is nil.

[16]It is readily apparent that the convergence of K to zero occurs more rapidly than that of N. Consequently, prices are finite and p become close to zero when K approaches to zero from Eq. (6.4).

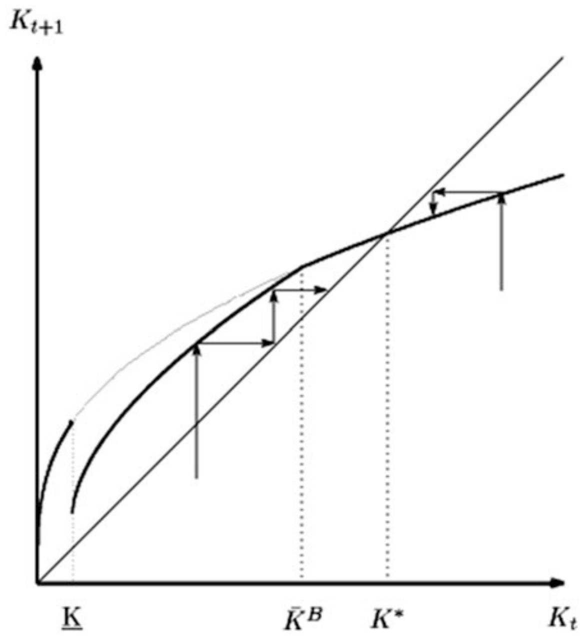


Fig. 6.1 Phase diagram in the snobbish economy: no conspicuous consumers in the steady state

Now I consider the phase diagram of K. Equation (6.12), in a relevant range, is readily apparent as strictly concave (see Appendix 1). On the other hand, when $K \notin (\underline{K}, \bar{K})$, Eq. (6.13) determines the evolution of capital. Finally, I obtain that Eqs. (6.12) and (6.13) are continuous at $\bar{K}$. The next proposition summarizes possible patterns of capital accumulation in the snobbish economy.

Proposition 1 *Two types of wealth accumulation path generate in the snobbish economy, depending on the value of b. For both cases, the economy has a unique steady state. The steady state is globally stable. In one case, conspicuous behavior disappears in the steady state and for the other case, the steady state is characterized by a lower level of capital and a positive number of conspicuous consumers.*

Proof See the proof in Appendix 1.

Figures 6.1 and 6.2 illustrate how the level of capital evolves in the snobbish economy. Realization depends on four parameters: $(\alpha, \beta, b, \theta)$. When the marginal decrease of the signaling value of the conspicuous good for an additional conspicuous consumer is sufficiently high; $b \geq \theta^{\frac{1}{1-\beta}}(1-\beta)(\frac{\alpha}{1-\alpha})^{\frac{\beta}{1-\beta}}$, conspicuous consumption remains in the steady state. On the other hand, when $b < \theta^{\frac{1}{1-\beta}}(1-\beta)(\frac{\alpha}{1-\alpha})^{\frac{\beta}{1-\beta}}$ holds, the steady state is that without conspicuous consumers.

The result indicates that, in the snobbish economy, disturbing effects of conspicuous consumption are weak: capital evolution is, despite the presence of conspicuous consumers, monotonic, as in the textbook Solow model.

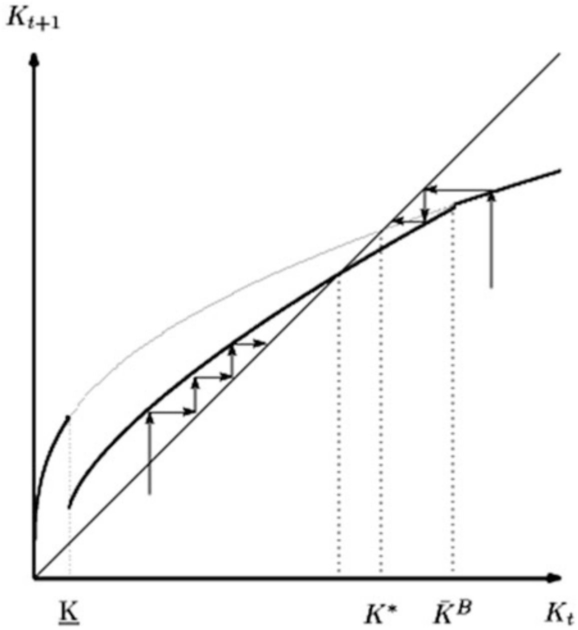


Fig. 6.2 Phase diagram in the snobbish economy: the steady state with conspicuous consumers

3.2 Bandwagon Economy

The analysis of the bandwagon economy is rather complicated because the demand for the conspicuous good might not be determined uniquely. To see this with clarity, define the signaling value function in the bandwagon economy as $S^B(N) = c - dN$, where $c > 0$, $d > 0$, and $N \in (0, 1)$. Hence, Eq. (6.10) now reads as

$$c - dN_t = \theta(1-\beta)\left(\frac{K_t}{N_t}\right)^{\beta}. \tag{6.15}$$

Figure 6.3 depicts the equilibrium condition in the labor market given by Eq. (6.15).

The figure suggests that when K is sufficiently small, there will be a unique intersection in $N \in (0, 1)$ for Eq. (6.15). Intuitively, as capital increases and *r.h.s.* of Eq. (6.15) shifts up for any level of N, another intersection will be generated. Hence, multiple equilibria generate in the bandwagon economy.

In the bandwagon economy, there is a point of $\bar{K}^B$, at which $S^B(N)$ and $\theta(1-\beta)K^{\beta}N^{-\beta}$ are tangent. For later reference, $\bar{K}^B$ is calculated here as

$$\bar{K}^B = \frac{\beta c^{\frac{1+\beta}{\beta}}}{(1-\beta)^{\frac{1}{\beta}}(1+\beta)^{\frac{1+\beta}{\beta}}\theta^{\frac{1}{\beta}}d}.$$

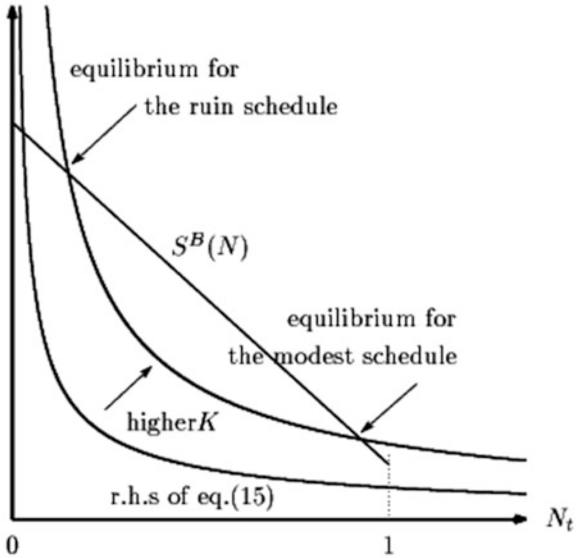


Fig. 6.3 Labor market equilibrium in the bandwagon economy

For a higher level of K than $\bar{K}^B$, there are no inner solutions for Eq. (6.10); consequently, Eq. (6.13) governs the dynamics. Note here that, in contrast to the snobbish economy, the number of conspicuous consumers jumps to zero discontinuously in the bandwagon economy. That is, when K overcomes $\bar{K}^B$, N will jump to one, which suggests that Eqs. (6.12) and (6.13) are discontinuous at $\bar{K}^B$

The other cases, in which a unique equilibrium always realize in the labor market for Eq. (6.10), might occur. That is, if $S^B(N)$ are tangent with $\theta(1-\beta)K^\beta N^{-\beta}$ for some given K at the level of $N > 1$, it is seen that there are no multiple equilibria. I impose a parametric restriction to exclude the equilibrium dynamics without multiple equilibria to concentrate on more interesting cases.

Condition:

$$\bar{N}^B = \frac{c}{d}\frac{\beta}{(1+\beta)} < 1,$$

where $\bar{N}^B$ is the equilibrium value of N when the economy has a capital level of $\bar{K}^B$. The derivation of the condition is explained in Appendix 2.

With that assumption, two equilibrium schedules are obtained for determination of N, as is apparent from Fig. 6.3. In one equilibrium schedule, an equilibrium level of N (J) decreases (increases) with K. Call this schedule a modest schedule. The modest schedule is generated when K is sufficiently high: denote $\underline{\mathrm{K}}^B$ such that $\underline{\mathrm{K}}^B = \{\underline{\mathrm{K}}^B \mid \lim_{N\to 1} S^B(N) = \lim_{N\to 1} \theta(1-\beta)(\underline{\mathrm{K}}^B)^\beta N^{-\beta}\}$. This is the level of K above which the modest schedule generates.

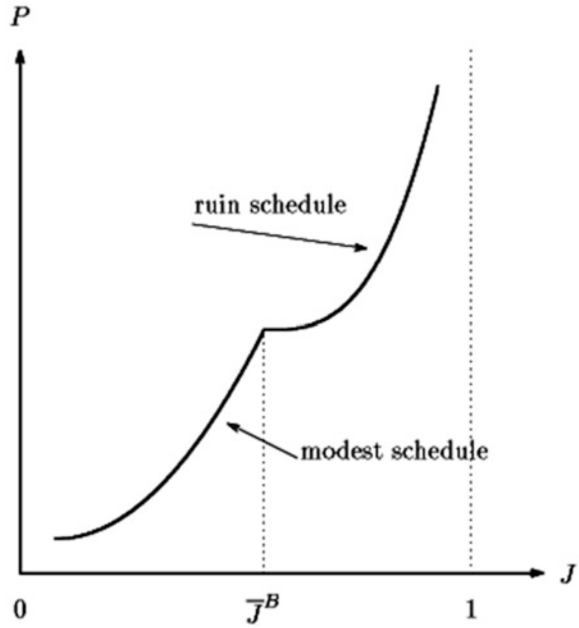


Fig. 6.4 Demand curve of the conspicuous goods in the bandwagon economy

For the modest schedule, it is easy to see that the price of the conspicuous good increases with $K \in (\underline{K}^B, \bar{K}^B)$ (and hence, with J) from Eq. (6.4). That is, the demand curve of the conspicuous good is upward sloping in this regime. Indeed, the possibility of an upward sloping demand curve is suggested by CJ for the bandwagon economy. As can be readily inferred, Eqs. (6.12) and (6.13) are continuous at $\underline{K}^B$ and K_{t+1} is strictly decreasing in K_t for the modest schedule (see Figs. 6.5–6.6(b)).

For the other equilibrium schedule, the equilibrium level of N (J) increases (decreases) with K. Call this schedule the ruin schedule. From Fig. 6.3, it can be deduced that the ruin schedule is generated when $K \in (0, \bar{K}^B)$. As to the demand curve in the ruin schedule, it might seem ambiguous that the demand curve is also upward sloping, as in the modest schedule, because both K and N rise in Eq. (6.4). The next lemma, however, shows that the demand curve is upward sloping in this regime as well.

Lemma 3 *The demand curve of the conspicuous good in the bandwagon economy is upward sloping.*

Proof See the proof in the Appendix 3.

Figure 6.4 portrays the demand curve of conspicuous goods in the bandwagon economy. As apparent there, the curve consists of two schedules, which have a kink at the point of $\bar{J}^B$ defined as

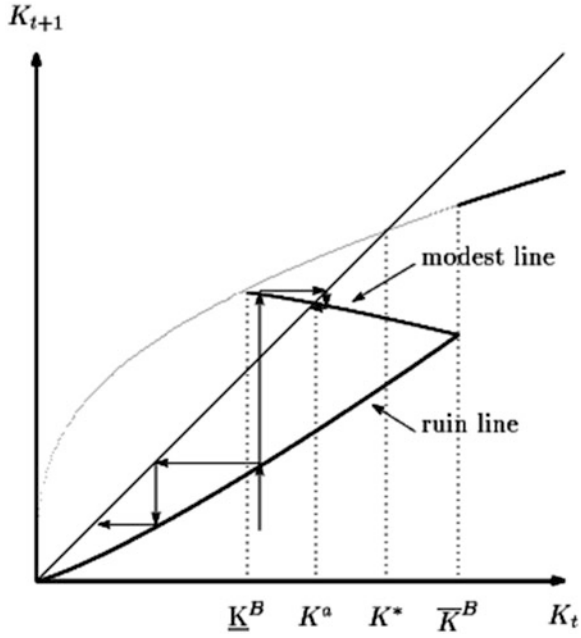


Fig. 6.5 Phase diagram in the bandwagon economy: an oscillating convergence path

$$\bar{J}^B \equiv 1 - \bar{N}^B = 1 - \frac{c\beta}{d(1+\beta)}.$$

In the bandwagon economy, the rationale of the multiple equilibria is given as follows. When the number of the conspicuous consumers increases, the price of the conspicuous good, which is equal to the opportunity cost of conspicuous consumption, rises (see Lemma 3). This higher opportunity cost will be compensated by higher signaling value driven by the larger number of conspicuous consumers. The economy then will attain the equilibrium. On the other hand, when the number of the conspicuous consumers decreases, the price of the conspicuous good, the opportunity cost of conspicuous consumption, declines. The signaling value is also reduced because of the smaller number of conspicuous consumers. Therefore, the economy will attain the equilibrium. Both of these situations are possible, so multiple equilibria generate in the bandwagon economy.

Now I consider the phase diagram of capital. In the bandwagon economy, multiple equilibria of conspicuous consumption generate when the economy has capital level of $K \in [\underline{K}^B, \bar{K}^B]$. On this regime, agents choose, taking the level of capital at the top of the period as given, which of two equilibria to be realized. The determination depends on the expectation of agents. The economy is governed by Eq. (6.13) when $K \in (\bar{K}^B, \infty)$ whereas the economy is on the ruin schedule when $K \in (0, \underline{K}^B)$. In the following analyses of the phase diagrams in (K_{t+1}, K_t) plane, I call the line for the ruin schedule the ruin line, and the line for the modest schedule

Fig. 6.6(a) Phase diagram in the bandwagon economy: polarization of the economy

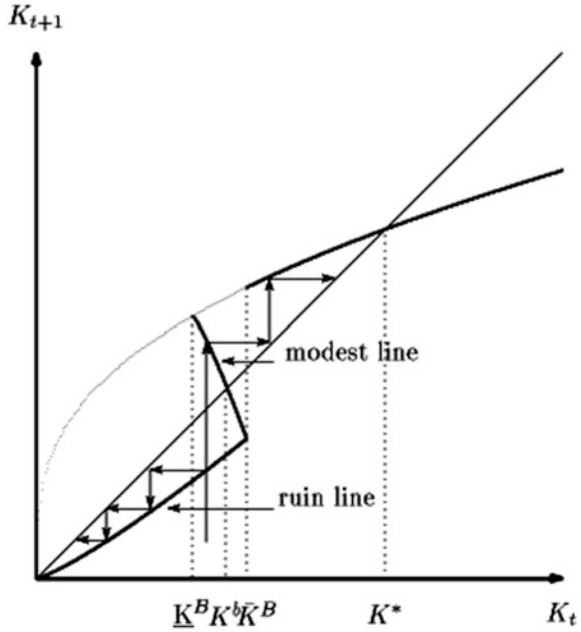
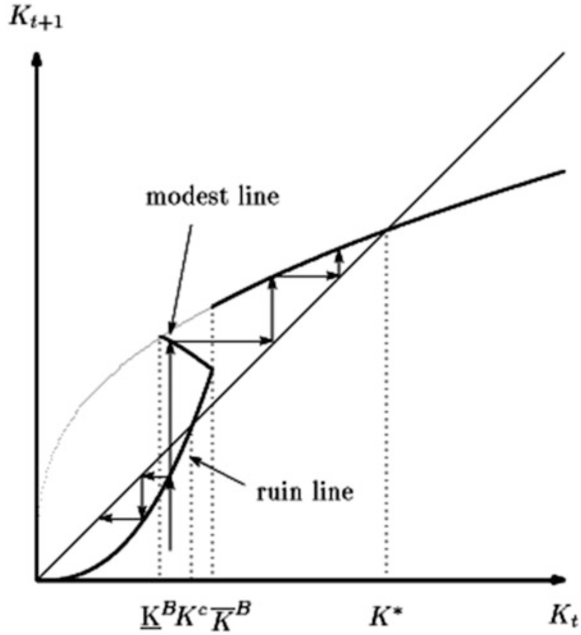


Fig. 6.6(b) Phase diagram in the bandwagon economy: polarization of the economy



the modest line. The following proposition classifies possible capital accumulation paths in the bandwagon economy.

Proposition 2 *There are three types of capital accumulation paths in the bandwagon economy under the assumption of* $\frac{c\beta}{d(1+\beta)} < 1$.[17] *For one case, a unique steady state, which might not be locally stable, exists (Fig. 6.5). For another case, there are two steady states, one of which is* K^*. *The other lies in* $K \in [\underline{K}^B, \bar{K}^B]$ *(Fig. 6.6(a)). The local stability of the latter steady state is ambiguous. In the third and final case, there are two steady states, one of which is* K^*. *The other is locally unstable and lies in* $K \in [\underline{K}^B, \bar{K}^B]$ *(Fig. 6.6(b)).*

Proof See the proof in Appendix 4.

Figure 6.5 depicts the case in which there is a unique steady state. When the steady state is stable as in the figure, then the economy can exhibit oscillating convergence to K^a along the modest line. This is the situation illustrated in chapter 3 of Sombart (1912): the rise of the economy feeds the decay; that decay subsequently brings prosperity. But when the economy goes along the ruin line, it shrinks monotonously.

Figures 6.6(a) and 6.6(b) show cases in which two steady states generate. One steady state is K^*; there are no conspicuous consumers in the steady state.

Regarding the case of Fig. 6.6(a), when the other steady state of K^b is unstable, it might be seen that no conspicuous consumers exist in the steady state even if there are some in transition. On the other hand, when K^b is stable (not shown), it can be deduced that if conspicuous consumers exist in the initial period, conspicuous consumers must exist not only in transition but also in the steady state. The economy will be caught in a poverty trap once capital becomes less than $\underline{K}^B$.

Finally, the case given by Fig. 6.6(b) is examined. Because there is a steady state K^c that is always unstable in this case, it can be suggested that the economy will converge to K^* if the initial level of capital is greater than K^c. In this case, conspicuous consumers disappear even if some exist in transition. Otherwise, the economy will be caught in the trap.

Among all the cases, it is most interesting that the economy might have an oscillating convergence path along the modest line.[18] This will happen when $\bar{N}^B = \frac{c\beta}{d(1+\beta)} < 1$ and $K_{t+1} - K_t \mid_{K_t \to \bar{K}^B} < 0$ are satisfied. These conditions require that c be smaller than d in the signaling value function, and that the level parameter in the production function θ should be small (see Appendix 4). Because these parameters are mutually independent, I might say that the situation of oscillating convergence will be indeed plausible. In that case, the rise of the economy feeds the decay through conspicuous consumption and the decay suppresses conspicuous consumption and engenders prosperity in the next period. Sombart (1912) predicted this situation.

[17] I exclude the origin, the poverty trap, from the definition of the steady state in this chapter.

[18] Mino (2006) shows that the oscillating convergence paths can be generated when we introduce the effect of keeping up with the Joneses into an overlapping-generations model.

Polarization of two economies, with identical economic fundamentals, initial levels of capital, and social norms, might occur. The realization depends on households' expectations. Cole et al. (1992) argue that two economies with identical economic fundamentals and initial conditions can have different growth rates when social norms of the two societies are different. The novelty here is that even if two economies have the same social norm, they can polarize. The results will have theoretical consequences on the convergence controversy, as summarized in Galor (1996).

I close this section by briefly discussing social welfare in the case where there are multiple equilibrium paths. Constructing adequate social welfare functions is indeed far beyond the scope of this chapter. Nonetheless, the social welfare function imposing equal weights to all the agents in the economy will be the simplest one and facilitates discussion.

In the bandwagon economy, two paths of capital accumulation can generate and the realization depends on the expectations of agents in the model. However, as to social welfare measured using an equally weighted social welfare function, it is easy to see that the modest schedule provides higher welfare than the ruin schedule. Then, it can be suggested that coordination failure occurs if the economy evolves along the ruin schedule when an alternative path of the modest schedule pertains.

To see this, merely note that the aggregate social rank utility is given as $\bar{a}$, irrespective of the number of conspicuous consumers: aggregate social rank utility is same in the ruin schedule and in the modest schedule. Therefore, utility from the substantial expenditure distinguishes the two schedules. It is readily apparent that the modest schedule dominates the ruin schedule in terms of the substantial expenditure because, in the modest schedule, the stock of capital becomes higher and the number of workers in sector 1 is larger.

4 Discussion

Above theoretical results show that when the social norm is of bandwagon type, conspicuous consumption seriously changes the pattern of capital accumulation from that of the Solow model. This outcome is indeed consistent with Hirschman (1984) and Basu (1989), who argue from philosophical points of view that the bandwagon effect in consumption is important, especially in the modern economy.[19] For further illustrations, here I will relate the results to anecdotal evidence that might be thought of as macroeconomic consequences of conspicuous consumption.

Example 1 The first example is from an anthropological study of Rao (2001) on rural India villages. McKim and Inden (1977) suggest that in rural India, personhood

[19]Mason (1998) is an excellent and comprehensive survey of the conspicuous consumption hypothesis. See chapter nine of Mason (1998) for a discussion of the contemporary importance of bandwagon effects.

is defined entirely in terms of one's relationships to others. For those individuals in those areas, it might be said that the bandwagon effect should be strong. Rao (2001) then investigates why very poor households in rural India spend large sums on celebrations and festivals. The study shows that their behaviors are explainable by their quest for higher social status within their villages.

This anecdotal evidence might provide a reason why underdeveloped regions in India have remained underdeveloped for a long time: poor households in rural India spend as much as six times their yearly incomes on celebrations. As a result, capital accumulation will be impeded seriously. This situation might be explained using the ruin line, or by an oscillating convergence path along the modest line to a steady state with conspicuous consumers and lower levels of capital.

Example 2 A second example is that of inefficient allocations of resources by Japanese firms.[20] Some economists argue that cooperate governance of Japanese firms was weak.[21] Especially during the bubble period, Japanese firms bought up Old Masters' paintings at high prices.[22] Because art will not function as capital, this could be conceptualized as conspicuous consumption by Japanese firms. Instead, their resources would have been better used for R&D activities to enhance productivity or for capital accumulation. As is well known, Japan's economic growth stagnated during the 1990s, which has come to be referred to as the lost decade. It might then be plausible that conspicuous consumption by firms engendered inefficient allocation of resources and caused business instability. An oscillating convergence path along the modest line might be captured by this situation.

5 Concluding Remarks

In this chapter, I described a macroeconomic model of capital accumulation with conspicuous consumption behaviors. The analyses should be regarded as a first attempt at investigating macroeconomic implications of conspicuous consumption because I consider the special case where the felicity function for the substantial expenditure converges to linearity. Nonetheless, the results obtained in the chapter are rich and suggest novel implications on the literature of status preference theory.

Especially, if the social norm is given by the bandwagon type, it is shown that the economy will be characterized by non-monotonic evolution of capital as well as by a multiplicity of equilibrium paths. The former result states that a Sombartian

[20]Slottje (1992) supports the existence of conspicuous consumption motivation in Japan with aggregate data for the pre-bubble period of 1974–1988.

[21]See, for discussion of cooperate governance of Japanese firms, Osano (2001).

[22]For example, it is well known that Yasuda Insurance Inc. bought a van Gogh painting for 5.8 billion yen.

economy is depicted with the model: the rise of the economy feeds decay through conspicuous consumption and that decay suppresses conspicuous consumption and brings the prosperity. Hence, I might say that conspicuous consumption is a cause of business fluctuation. The latter outcome indicates the possibility of polarization of two economies with identical economic fundamentals, initial conditions and social norms. Hence, the results imply that the literature of status preference theory will affect the convergence controversy of Galor (1996).

Appendices

Appendix 1

This appendix proves Proposition 1. When $K \in (\underline{K}, \bar{K})$ the wealth evolution in the snobbish economy is governed by

$$K_{t+1} = \alpha\{K_t + \theta K_t^{\beta} N(K_t)^{1-\beta}\}, \tag{6.16}$$

and

$$bN_t = \theta(1-\beta)\left(\frac{K_t}{N_t}\right)^{\beta}. \tag{6.17}$$

From these two equations, it is obtained that

$$K_{t+1} = \alpha\{K_t + \theta^{\frac{2}{1+\beta}} \frac{(1-\beta)}{b} K_t^{\frac{2\beta}{1+\beta}}\}.$$

As is apparent from that equation, this is a strictly concave function of K_t.

It is convenient to investigate the characteristics of dynamics in the snobbish economy globally by inquiring into two points. One is the relative position of $\bar{K}$ to K^*. When $\bar{K} < K^*$ holds, K^* must be the steady state, although there might be another steady state in $K \in (\underline{K}, \bar{K})$. Indeed, because (6.16) is strictly concave and, from the continuity argument of Eqs. (6.12) and (6.13) at $\bar{K}$, there might be another steady state in $K \in (\underline{K}, \bar{K})$ if $K_{t+1} - K_t < 0$ holds at $\underline{K}$. The other is the condition to generate the steady state in $N \in (0, 1)$. If there is a steady state of N in $(0, 1)$, then it must be the one with $K \in (\bar{K}, \underline{K})$.

For the first condition, it is simply obtained that $\bar{K} < K^*$ holds if and only if

$$b < \theta^{\frac{1}{1-\beta}}(1-\beta)(\frac{\alpha}{1-\alpha})^{\frac{\beta}{1-\beta}}. \tag{6.18}$$

As for the other condition, the steady state level of N, I impose the steady state condition on (6.16) and (6.17) to obtain the following.

$$N = \left(\frac{1-\alpha}{\alpha}\right)^{\frac{\beta}{1-\beta}} \theta^{\frac{1}{1-\beta}} \frac{1-\beta}{b}$$
$$\equiv \Lambda(b)$$

The last condition indicates that the number of steady states with N less than one is, at most, one. That is, when $\Lambda(b) < 1$, there will be a steady state of capital, $K \in (\underline{K}, \bar{K})$. This condition excludes the case in which there are two steady states in $K \in (\underline{K}, \bar{K})$ when condition (6.18) is violated.

The last question is whether or not $\Lambda(b) < 1$ and $\bar{K} < K^*$ simultaneously hold and two steady states are generated: one is for $K \in (\underline{K}, \bar{K})$ and the other is K^*. It is, however, readily apparent that (6.18) and $\Lambda(b) < 1$ are contradictory, so that, in any case, only one steady state generates globally. To sum up, there is a steady state without conspicuous consumption, K^*, when $b < \theta^{\frac{1}{1-\beta}}(1-\beta)(\frac{\alpha}{1-\alpha})^{\frac{\beta}{1-\beta}}$ holds (Fig. 6.1). Otherwise, conspicuous consumers exist in the steady state with a lower level of capital (Fig. 6.2). *Q.E.D.*

Appendix 2

The assumption in Sect. 3.2 is derived using the following three conditions:

(i) The condition for tangency

$$d = \theta\beta(1-\beta)\left(\frac{\bar{K}^B}{\bar{N}^B}\right)^{\beta}(\bar{N}^B)^{-1},$$

(ii) The condition for equilibrium

$$c - d\bar{N}^B = \theta(1-\beta)\left(\frac{\bar{K}^B}{\bar{N}^B}\right)^{\beta},$$

and

(iii) The condition for inner solution

$$\bar{N}^B < 1.$$

Q.E.D.

Appendix 3

The demand curve is upward sloping in the modest schedule. Consider then the ruin schedule, in which the equilibrium level of N increases with K, and assume that p decreases with K. In this case, J decreases as K increases. The decline of J implies a drop in the signaling value of conspicuous consumption in the bandwagon economy. This, however, contradicts the equilibrium condition given by Eq. (6.9). For Eq. (6.9) to be satisfied, p must increase with K in the bandwagon economy. Hence, it is shown that the demand curve is upward sloping in the ruin schedule, too. *Q.E.D.*

Appendix 4

This appendix proves Proposition 2. To analyze how K evolves in the bandwagon economy, it is sufficient to investigate three aspects: (i) the sign of $K_{t+1} - K_t$ in the ruin schedule when K is close to zero; (ii) the sign of $K_{t+1} - K_t$ at $\bar{K}^B$; and (iii) the relative position of $\bar{K}^B$ to K^*.

As to the first point, it is shown that

$$\begin{aligned} K_{t+1} - K_t \mid_{K_t \to 0} &= \alpha K_t - \theta K_t^{\beta} N_t^{1-\beta} - K_t \mid_{K_t \to 0} \\ &= -(1-\alpha)K_t - \frac{b}{\beta(1-\beta)} N_t^2 \mid_{K_t \to 0} \\ &< 0. \end{aligned}$$

This indicates that the ruin line lies below the 45° line near the origin. Notice also that the positive capital level is ensured by Eq. (6.11).[23] Furthermore, it can be seen that the ruin line is a monotonously increasing curve because N increases with K and the price of conspicuous good and the number of conspicuous consumers decline with K along the ruin schedule.

With respect to the second matter, remembering that the ruin line and the modest line intersect at $\bar{K}^B$, a little algebraic treatment leads to

$$\begin{aligned} K_{t+1} - K_t \mid_{K_t \to \bar{K}^B} &= \alpha K_t - \theta K_t^{\beta} N_t^{1-\beta} - K_t \mid_{K_t \to \bar{K}^B} \\ &= \bar{K}^B \left[\theta \frac{\bar{K}^B}{N_t}^{\beta-1} - (1-\alpha) \right] \end{aligned}$$

[23] See also footnote 13.

$$= \bar{K}^B \left[\frac{c^{\frac{\beta-1}{\beta}} \theta^{\frac{1}{\beta}} (1-\beta)^{\frac{1-\beta}{\beta}}}{(1+\beta)^{\frac{\beta-1}{\beta}}} - (1-\alpha) \right].$$

Hence, $K_{t+1} - K_t \mid_{K_t \to \bar{K}^B} \geq 0$ if and only if

$$\frac{c^{\frac{\beta-1}{\beta}} \theta^{\frac{1}{\beta}} (1-\beta)^{\frac{1-\beta}{\beta}}}{(1+\beta)^{\frac{\beta-1}{\beta}}} - (1-\alpha) \geq 0, \tag{6.19}$$

and *vice versa*.

Finally, as to the third point, $\bar{K}^B > K^*$ holds if and only if

$$\bar{K}^B = \frac{\beta c^{\frac{1+\beta}{\beta}}}{(1-\beta)^{\frac{1}{\beta}} (1+\beta)^{\frac{1+\beta}{\beta}} \theta^{\frac{1}{\beta}} d} > \left(\frac{\alpha\theta}{1-\alpha} \right)^{\frac{1}{1-\beta}} = K^*. \tag{6.20}$$

Under the condition that $\frac{c\beta}{d(1+\beta)} < 1$, (6.19) and (6.20) do not hold simultaneously.

Now the proposition will be best understood by graphical expositions. Figures 6.5–6.6(b) illustrate three possible patterns of capital accumulation. Figure 6.5 is a case in which $\bar{K}^B \geq K^*$ and $K_{t+1} - K_t \mid_{K_t \to \bar{K}^B} < 0$ hold. Hence there is a unique steady state (denoted as K^a in the figure). The local stability of the steady state depends on the slope evaluated at K^a. The figure depicts a stable case.

Figure 6.6(a) depicts a case where $\bar{K}^B < K^*$ and $K_{t+1} - K_t \mid_{K_t \to \bar{K}^B} < 0$ hold, whereas Fig. 6.6(b) illustrates the case where $\bar{K}^B < K^*$ and $K_{t+1} - K_t \mid_{K_t \to \bar{K}^B} \geq 0$ hold. For both cases, there is a stable steady state of K^*, in which conspicuous consumption disappears.

For the case of Fig. 6.6(a), there is a steady state denoted as K^b in the figure. The stability of K^b is ambiguous (the figure shows an unstable case). Finally, the case given by Fig. 6.6(b) shows a steady state K^c, which is always unstable. *Q.E.D.*

Addendum: A Caveat[24]

Based on Clark et al. (2008)'s review of the empirical literature of social preferences, recent trends in the field continue to grow. Theoretical contributions also continue to emerge, including Ravn et al. (2010) and Di Pace and Faccini (2012). In this note, I suggest a caveat regarding how we should interpret various types of social preferences; in particular, I have found that some studies use inappropriate definitions of preference externality related to consumption.

A classic view of social preference is found in Veblen (1899). Veblen started his discussion with a perspective of pecuniary emulation among citizens. When citizens compete with others in terms of monetary achievements, they must explicitly

[24]This addendum has been newly written for this book chapter.

demonstrate that they are superior to their peers in that regard. Thus, what matters when researchers think about the effects of pecuniary emulation on economic decisions is the information structure they introduce to their analyses. Also, it is important to remember that preference externality can be defined over several variables, including consumption, asset holdings, and income levels. In the literature on happiness, these three variables significantly affect happiness levels with the expected signs for coefficients.

Preference externality defined over consumption is the most convenient for theoretical analyses but it requires some caution. It is natural for researchers to introduce such effect in a reduced form utility function, such as $U(c, \bar{c})$ where c is own consumption and $\bar{c}$ is reference consumption. Behind this specification is the information structure in which citizens can recognize their peers' consumption levels. While such an assumption seems acceptable, one common misconception in the theoretical literature of social preferences is that $U(c, \bar{c})$ captures the effect of conspicuous consumption. It indeed captures only consumption externality, which includes not only the influence of consumption of others in the current period, but also the effects of habit formation due to past own consumption and future aspirations regarding own consumption levels.

So, what is conspicuous consumption? It actually appears when preference externality is defined over asset/saving levels. It is possible for researchers to introduce preference externality via saving as $U(c) + V(s, \bar{s})$, where c is own consumption, s is own saving, and $\bar{s}$ represents reference saving. The functional form of V would be the author's choice, but a recent experimental study by Ono and Yamada (2012) shows that a difference specification such as $V(s, \bar{s}) = v^1(s - \bar{s})$ fits the experimental data better than a ratio specification expressed as $V(s, \bar{s}) = v^2(s/\bar{s})$. Here again, an information structure plays a crucial and implicit role, which is that economic agents can observe the asset/saving levels of peers perfectly. Obviously, this is a strong assumption and it will make theoretical analyses more straightforward when introduced.

In real life, it is plausible that asset/saving levels of peers constitute private information. When information is private regarding $\bar{s}$, economic agents somehow must inform their peers that they are indeed better off than others. In this case, a greater amount of standard consumption, which would be visible to others, does not work as a device to advertise wealth levels. This is because even when the amount of standard consumption is greater than that of others, it is possible that people's asset/saving levels, over which social preference is defined, are smaller than their peers'. Such an observation can be explained by the differences in propensity to consume among people.

According to Veblen, this is where conspicuous consumption plays its role. By definition, people cannot derive utility from consuming conspicuous goods. Put differently, the definition of conspicuous goods includes that they do not provide the consumer with any value by consuming them. So, why do consumers spend money on seemingly useless goods? Veblen suggested that it is because they can signal their level of wealth in such a way that peers' inferences about their asset levels will be valid. A neoclassical economics theory validates this, as the marginal utility from standard consumption is smaller for greater amounts of consumption levels, just as Corneo and Jeanne (1997a) and Yamada (2008) showed. A typical example of consumption of conspicuous goods would be the use of aristocratic names. As Sombart argued, dropping aristocratic family names in conversation would signal wealth levels but would not increase utility levels by itself. Obviously, the signaling effect of conspicuous consumption is quite different from consumption externality, which is expressed as $\bar{c}$ in $U(c, \bar{c})$.

References

Abel AB (1990) Asset prices under habit formation and catching up with the Joneses. Am Econ Rev 80:38–42

Bagwell LS, Bernheim BD (1996) Veblen effects in a theory of conspicuous consumption. Am Econ Rev 86:349–373

Bakshi GS, Chen Z (1996) The spirit of capitalism and stock-market prices. Am Econ Rev 86:133–157

Banerjee AV, Newman AF (1993) Occupational choice and the process of development. J Pol Econ 101:274–298

Basmann RL, Molina DJ, Slottje DJ (1988) A note on measuring Veblen's theory of conspicuous consumption. Rev Econ Stat 70:531–535

Basu K (1989) A theory of association: social status, prices and markets. Oxf Econ Pap 41:653–671

Becker GS (1991) A note on restaurant pricing and other examples of social influences on price. J Pol Econ 99:1109–1116

Blanchflower DG, Oswald AJ (2004) Well-being over time in Britain and the USA. J Public Econ 88:1359–1386

Clark AE, Oswald AJ (1996) Satisfaction and comparison income. J Public Econ 61:359–381

Clark AE, Frijters P, Shields MA (2008) Relative income, happiness, and utility: an explanation for the Easterlin paradox and other puzzles. J Econ Lit 46:95–144

Cole HL, Mailath GJ, Postlewaite A (1992) Social norms, savings behavior, and growth. J Pol Econ 100:1092–1125

Cole HL, Mailath GJ, Postlewaite A (1995) Incorporating concern for relative wealth into economic models. Feder Reserve Bank Minneap Q Rev 19:12–21

Corneo G, Jeanne O (1997a) Conspicuous consumption, snobbism and conformism. J Public Econ 66:55–71

Corneo G, Jeanne O (1997b) On relative wealth effects and the optimality of growth. Econ Lett 54:87–92

Corneo G, Jeanne O (1997c) Snobs, bandwagons, and the origin of social customs in consumer behavior. J Econ Behav Organ 32:333–347

Di Pace F, Faccini R (2012) Deep habits and the cyclical behaviour of equilibrium unemployment and vacancies. J Econ Dyn Control 36:183–200

Easterlin RA (2001) Income and happiness: towards an unified theory. Econ J 111:465–484

Frank RH (1985) The demand for unobservable and other nonpositional goods. Am Econ Rev 75:101–116

Frijters P, Leigh A (2008) Materialism on the march: from conspicuous leisure to conspicuous consumption? J Behav Exp Econ (formerly J Socio-Econ) 37:1937–1945

Futagami K, Shibata A (1998) Keeping one step ahead of the Joneses: status, the distribution of wealth, and long run growth. J Econ Behav Organ 36:109–126

Gali J (1994) Keeping up with the Joneses: consumption externalities, portfolio choice, and asset prices. J Money Credit Bank 26:1–8

Galor O (1996) Convergence? Inferences from theoretical models. Econ J 106:1056–1069

Hayakawa H (2000) Bounded rationality, social and cultural norms, and interdependence via reference groups. J Econ Behav Organ 43:1–34

Hayes KJ, Molina DJ, Slottje DJ (1988) Measuring preference variation across North America. Economica 55:525–539

Hayes KJ, Slottje DJ, Ferrantino MJ, Wagner J (1992) Veblen effects and their impact on the new European community and some of their trading partners. Eur Econ Rev 36:51–70

Hirschman AO (1984) Against parsimony: three easy ways of complicating some categories of economic discourse. Am Econ Rev 74:89–96

Hollander H (2001) On the validity of utility statements: standard theory versus Duesenberry's. J Econ Behav Organ 45:227–249

Hume D (1739) A treatise of human nature (reprinted 1967). Clarendon Press, Oxford

Konrad KA (1992) Wealth seeking reconsidered. J Econ Behav Organ 18:215–227

Mason R (1998) The economics of conspicuous consumption – theory and thought since 1700. Edward Elgar, Cheltenham

McBride M (2001) Relative-income effects on subjective well-being in the cross-section. J Econ Behav Organ 45:251–278

McKim M, Inden R (1977) Toward an ethnosociology of South Asian caste system. In: David K (ed) The new wind: changing identities in South Asia. The Hague, Monton

Mino K (2006) Consumption externalities and capital accumulation in an overlapping generations economy. Mpra paper 17016, University Library of Munich

Ono Y, Yamada K (2012) Difference or ratio: implication of status preference on stagnation. Iser discussion paper 0856, Institute of Social and Economic Research, Osaka University

Osano H (2001) Economics of cooperate governance (in Japanese). Nihon-Keizai-Shinbun-Sha, Tokyo

Rao V (2001) Poverty and public celebrations in Rural India. The world bank policy research working paper 2528, The World Bank

Ravn MO, Schmitt-Grohe S, Uribe M (2010) Incomplete cost pass-through under deep habits. Rev Econ Dyn 13:317–332

Saito T, Sannabe A, Yamada K (2005) Satisfaction and comparison income in Japan: evidence from data of Japanese union workers. Econ Bull 4:1–20

Slottje DJ (1992) Is there conspicuous consumption in Japan? Jpn World Econ 4:333–342

Smith A (1759) Theory of moral sentiments. In: Raphael DD, Macfie AL (eds) Adam Smith: the theory of moral sentiments (reprinted 1976). Clarendon Press, Oxford

Sombart W (1912) Liebe, Luxus und Kapitalismus (reprinted 1967). Deutscher Taschenbuch Verlag, Munchen

Veblen T (1899) The theory of the leisure class (reprinted 1957). George Allen Unwin, London

Yamada K (2008) Macroeconomic implications of conspicuous consumption: a Sombartian dynamic model. J Econ Behav Organ 67:322–337

Zou H-F (1994) The spirit of capitalism and long-run growth. Eur J Pol Econ 10:279–293

Chapter 7
On Persistent Demand Shortages: A Behavioural Approach

Yoshiyasu Ono and Junichiro Ishida

Abstract We incorporate two sets of behavioural assumptions, fairness concerns and insatiable desire for money, into a dynamic optimization model to illuminate how they can generate persistent aggregate demand shortages. We obtain the conditions for persistent unemployment and temporary unemployment. Policy implications differ significantly between the two cases. A monetary expansion raises private consumption under temporary unemployment but not under persistent unemployment. A fiscal expansion may or may not increase short-run private consumption but crowds out long-run consumption under temporary unemployment. Under persistent unemployment, however, a fiscal expansion always increases private consumption. The "paradoxes of toil and flexibility" also appear.

Keywords Wage adjustment • Fairness • Phillips curve • Demand shortage • Persistent stagnation

1 Introduction

Despite the advances over recent years, it remains a daunting task to fully understand the underlying mechanism of persistent stagnations, such as the Great Depression of the 1930s and Japan's lost decade of the 1990s. In this chapter we contribute to this long-standing issue by presenting a "behavioural" model of persistent stagnations, especially focusing on why stagnations, once started, often persist for so long. We incorporate two sets of behavioural factors, fairness concerns and insatiable liquidity preferences, into a continuous-time dynamic model of a monetary economy as portrayed by Keynes (1936). In this setup, a liquidity trap occurs, which may

The original article first appeared in the Japanese Economic Review 65(1):42–69, 2014. A newly written addendum has been added to this chapter.

Y. Ono (✉) • J. Ishida
Institute of Social and Economic Research, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan
e-mail: ono@iser.osaka-u.ac.jp; jishida@osipp.osaka-u.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_7

create persistent shortages of aggregate demand and involuntary unemployment. Moreover, we obtain the necessary conditions for persistent unemployment and for temporary unemployment. These conditions turn out to be critical because the effects of fiscal and monetary policies on aggregate demand are shown to be quite different in the two cases.

To clarify the scope of our analysis, it is perhaps worth emphasizing at the outset that, despite the difference in approach, the present chapter shares a lot in common with the new Keynesian paradigm at its most fundamental level. We turn to Keynes as the source of insight on malfunctioning economies.[1] Since the publication of *The General Theory of Employment, Interest, and Money*, many attempts have been made to pick up the components of Keynes' insight and reduce them down to their essence. An emerging consensus seems to be that the crux of Keynesianism lies in various forms of market imperfection, especially nominal rigidity, whose effects are amplified through the effect of aggregate demand.[2]

We have no intention of arguing against this general consensus. Indeed, nominal wage sluggishness, which arises from fairness concerns, plays a critical role in the present study as well. It is our stance, however, that nominal rigidity alone does not exhaust all of Keynes's insight. We argue that another factor, insatiable liquidity preference, is equally, if not more, important in generating persistent unemployment. This chapter is an attempt to recast and revitalize this insight in a modern macroeconomic framework and to derive its dynamic policy implications. In particular, we argue that a model characterized by insatiable liquidity preferences, combined with fairness-based nominal rigidity, can go a long way in helping to explain important aspects of stagnant economies. Although this insight has somehow escaped economists' attention, especially in a rigorous context, we believe that the current framework sheds light on a different side of Keynesianism and provides a useful description of prolonged stagnations.

To be more precise, when analysing unemployment, whether temporary or permanent, we must consider sluggish price (or wage) adjustments as in the Phillips curve,[3] because any possibility of demand–supply imbalance would intrinsically be avoided without sluggish price adjustments. However, the original Phillips curve has a clear shortcoming: it lacks any microeconomic foundation for the nominal wage adjustment process. To overcome this problem, several attempts have been made to augment the Phillips curve in this direction. Among the most notable

[1] Several prominent economists make this point rather emphatically. Greg Mankiw (2008) notes: "If you were going to turn to only one economist to understand the problems facing the economy, there is little doubt that the economist would be John Maynard Keynes. ... His insights go a long way toward explaining the challenges we now confront" (NewYork Times, 28 November). "The Keynesians basically got it right", says George Akerlof (2007).

[2] The importance of nominal rigidity has recently gained some attention in other areas. In the search and matching literature, for instance, it has been identified as a possible resolution of the so-called Shimer puzzle (2005). See Cardullo (2010) for a survey.

[3] As an alternative to the sticky-price approach, Mankiw and Reis (2002) propose a sticky-information model where information diffuses slowly through the population.

examples of these attempts are the new classical Phillips curve, the New Keynesian Phillips curve and the hybrid of forward-looking and backward-looking Phillips curves.[4] Models along this line include Yun (1996), Woodford (2003), Gali (2008), Dotsey et al. (1999), Golosov and Lucas (2007) and Gertler and Leahy (2008). The first three assume Calvo's staggered pricing while the others adopt the menu cost approach. They examine firm and household reactions to a policy or parameter shock that occurs in the middle of the initial period. Because they assume that agents cannot revise their behaviour in the middle of any period, demand–supply imbalances temporarily arise (although they eventually disappear). Krugman (1998) also analyses Japan's stagnation in a two-period model in which prices are rigid only in the initial period and any demand shortages disappear in the second period.[5]

What is common among these studies is that they preclude the possibility of demand shortages in the steady state, almost by construction.[6] The logic of unemployment in this existing literature is indispensably based on the assumption that there is a period within which prices and wages cannot be revised.[7] According to this logic, unemployment is necessarily a temporary phenomenon that occurs only in the adjustment process, as any demand shortages would eventually dissipate once people's expectations are corrected and prices are adjusted completely. Therefore, although certainly insightful in terms of understanding short-run fluctuations, these existing settings face serious limitations in explaining long and persistent demand shortages because, for them to occur, it must be either: (i) that people continuously hold false expectations about prices; or (ii) that the price (or wage) adjustment process is extremely slow. Neither appears to be a likely cause of the persistent stagnations that we have so far observed. In many of those instances, the stagnations persisted for more than a decade, which should have been more than enough for people to adjust their expectations correctly. It is also implausible that the price adjustment process was so slow that even after all these years, prices and wages could not be adjusted to the equilibrium levels to clear the markets.

[4]See Woodford (2003) for properties of these Phillips curves.

[5]The current analysis is especially close in spirit to Krugman (1998) in that he also analyses a model of stagnant economies. Aside from the fact that his model assumes full employment in the end, however, there is a key distinction between our analysis and his in terms of policy implications. In Subsection 4.2, we explore how our analysis differs from his in more depth.

[6]In these settings, the existence of demand shortages implies that the deflation rate cumulatively expands. If this were to continue, prices would reach zero within a finite time and the feasibility condition would eventually be violated.

[7]Of course, unemployment may persist when the labour market suffers from some imperfections. For instance, the efficiency-wage rationale (e.g. Akerlof 1982; Shapiro and Stiglitz 1984), implies the presence of equilibrium unemployment. Also, unemployment obviously persists in the presence of some labour-market frictions; see Rogerson and Shimer (2011) for a survey on this approach. It should be noted that these types of unemployment, which stem fundamentally from some labour-market imperfections, are not our focus here. Although we also consider sluggish wage adjustment, our primary focus is on the shortage of aggregate demand, which, as we will later argue, is the fundamental force behind persistent unemployment in our model.

It is our view that to account for long and persistent deviations from full employment, we need an alternative framework where the presence of unemployment does not hinge on unexpected events or policies.[8] To explore this possibility, we adopt an approach that is distinct from the previous published literature: we attempt to find the fundamental cause of prolonged demand shortages in human nature, rather than in market imperfection. To this end, the model builds on two sets of behavioural assumptions, one on the demand side and the other on the supply side. On the demand side, we assume that agents (as consumers) derive utility from holding liquid assets, namely money, and, moreover, that the marginal utility from money holding never dissipates. On the supply side, we assume that agents (as workers) have a strong preference for being treated fairly, and their motivation and productivity depend heavily on this fairness concern, as in Akerlof and Yellen (1990). We will show that the latter assumption, that workers are concerned about fairness, has a profound impact on the wage dynamics, which eventually amount to what might be called the "behavioural Phillips curve".

We show that the interaction of insatiable liquidity preferences and the behavioural Phillips curve gives rise to persistent stagnations, thereby pointing to a specific route through which demand shortages arise as an equilibrium phenomenon. Of the two sets of assumptions, the driving force of the model is the insatiable nature of our intrinsic preferences for liquid assets, which embodies our perpetual craving for money. The importance of insatiable money demand in generating persistent demand shortages has been pointed out and formally analysed by Ono (1994, 2001). Aggregate demand falls short of its full-employment level because people hold on to liquid assets in the shadow of a stagnation, which, in turn, reduces aggregate demand even more. This setting, which captures Keynes's notion of a monetary economy,[9] has recently been used in various analyses of persistent stagnation.[10] All

[8]To illuminate this point, we construct a model of perfect foresight with no unexpected shocks where everyone understands the current state of the economy. Obviously, this is not to say that expectations play no role in aggregate fluctuations; they most certainly do. What we argue here is, rather, that unemployment could persist even under perfect foresight where there are no surprises. Moreover, we adopt a continuous-time setting which inherently has no border of periods: at any point in time, firms are free to set any prices and wages, so that there is no "adjustment" in this particular sense. We instead consider a nominal wage adjustment process that rests on workers' inherent concerns for fairness.

[9]Keynes (1936) defines a non-monetary economy as an economy where there is no asset such that its liquidity premium remains strictly positive.

[10]Ono (1994, 2001) presents a dynamic optimization model of a monetary economy with sluggish nominal wage adjustment and shows that a liquidity trap occurs and unemployment due to demand shortages persists in the steady state if there is a positive lower bound of the marginal utility of liquidity. Since then, many attempts have been made to extend this setting in various directions. Matsuzaki (2003) finds the effect of a consumption tax on effective demand in the presence of poor and rich people. Hashimoto (2004) examines the intergenerational redistribution effects of the public pensions system in an overlapping generations framework with the present type of stagnation. Ono (2006, 2014) extends the model into a two-country framework and analyses the international spillover effects of fiscal spending and trade policies on each country's aggregate demand. Johdo (2006) considers the relationship between R&D subsidies and unemployment.

of these previous studies, however, simply take the original Phillips curve as given, leaving the nominal wage adjustment process as a black box.

To complete this missing link in the existing argument, we add fairness concerns into the supply side of the economy, especially the wage formation process. We provide an explicit account of how workers' fairness perceptions form and evolve over time by incorporating various stylized facts in social psychology and behavioural economics into the model, and derive a wage adjustment process that exhibits a degree of nominal rigidity. We place this adjustment mechanism in a dynamic general equilibrium framework to provide our own version of the Phillips curve, in which the inflation rate is shown to be governed by the liquidity premium, the subjective discount rate and the unemployment rate. Around the steady state, the inflation rate depends only on the unemployment rate, as in the original Phillips curve. It should be noted, however, that the role of nominal wage rigidity is fundamentally different from in the previous models. In fact, while insatiable demand for money is the direct cause of persistent demand shortages, nominal wage rigidity works as a buffer, preventing the economy from falling into even more severe conditions. In Sect. 5, we will discuss this point more extensively.

Given this construction, we identify conditions that lead to persistent and temporary demand shortages and show that policy implications differ substantially between the two cases, suggesting that different sets of remedies may be needed to restore full employment once the economy becomes stuck in the state of persistent stagnation. We show that the fairness concerns, and the behavioural Phillips curve that results from them, give rise to a novel policy implication that is quite different from that in Ono (2001). As mentioned, one weakness of Ono (2001) is that he takes the Phillips curve as given and simply assumes, without any microeconomic foundation, that the monetary expansion rate is linearly related to the inflation rate, so that a rise in the monetary expansion rate always results in a one-to-one increase in the inflation rate. Because of this feature, expansionary monetary policy is always effective in that it stimulates consumption both in the short run and in the steady state, whether or not unemployment persists in the steady state. We view this conclusion as unsatisfactory because the recent Japanese experience seems to indicate that monetary expansions in the stagnation phase have little, if any, impact on aggregate economic outcomes. In contrast, the current analysis gives a result

Rodriguez-Arana (2007) examines the dynamic path with public deficit in the present stagnation case and compares it with that in the neoclassical case. Johdo and Hashimoto (2009) introduce foreign direct investment into a two-country model with the present stagnation mechanism and analyse the effect of the corporation tax on employment in each country. Ono (2010) applies this model to analyse various aspects of Japan's long-run stagnation and economic policies under it. Murota and Ono (2011) find that a preference for money holding as status is insatiable and, thus, generates persistent stagnation of the present type. Hashimoto and Ono (2011) examine the effects of various pro-population policies under this type of stagnation and show that they are opposite to those under full employment. Murota and Ono (2012) introduce a preference for deposit holdings to this setting and explain zero nominal interest rates and excess reserves held by commercial banks under persistent stagnation.

consistent with such experiences that a rise in the monetary expansion rate has no effect on the dynamic path that leads to persistent unemployment.[11] We argue that by elucidating the workings of stagnant economies, the current analysis provides a framework to understand why and how economic policies could work in a different way once an economy is trapped in the stagnation phase.

2 The Wage Dynamics

Before we set up a dynamic optimization problem, we first illustrate the wage adjustment process to describe how nominal wages are determined in this economy. As stated, the driving force behind this whole process is workers' inherent concern for fairness. Having derived the equilibrium wage dynamics we then incorporate it into a dynamic general equilibrium framework in the next section.

2.1 The Setup

There are continua of identical workers and firms, both with unit measure. At any instance, each worker is either employed or unemployed. Because firms are all identical, unemployed workers are randomly assigned to firms whenever vacancies exist. As a consequence, each firm hires workers of equal size in any equilibrium. Let $x(t)$ denote the number of workers newly hired at time t, which is the same across firms. Moreover, define $X(t)$ as the aggregate rate of employment $X(t) \in [0,\ 1]$, where $X(t) = 1$ means that the economy is in the state of full employment.

Although employed workers can, in principal, quit and leave their respective firms at will, workers must incur some flow cost while they are unemployed.[12] We assume that the cost of unemployment is prohibitively large, so that workers would not choose to voluntarily leave their firms when there is a positive probability of being unemployed. This means that their mobility is heavily influenced by the aggregate rate of employment: workers are perfectly mobile when $X(t) = 1$, but they are virtually forced to work for their respective firms when $X(t) < 1$. These assumptions are made for clarity and tractability rather than for realism.[13]

[11]Therefore, the current framework provides a framework to reassess the role of fiscal policy, which seems to have regained its recognition in the face of the recent crisis. For instance, Blanchard et al. (2010, p. 5) point out the importance of counter-cyclical fiscal policy as one of the important lessons of the recent crisis, although "fiscal policy took a backseat to monetary policy" in the past two decades.

[12]The cost of unemployment is meant to capture the physical costs of job search as well as the more psychological costs of anxiety or social stigma inherently attached to unemployment.

[13]All we need is the fact that workers' mobility is more limited when there are more unemployed workers waiting for job offers and the expected duration of unemployment is longer. Our

In this model, because both firms and workers are identical and there is no information asymmetry, job separations occur only for exogenous reasons (no voluntary job separations on any equilibrium path) as long as unemployment exists. We assume that each employed worker randomly separates from the current firm at Poisson rate α. The aggregate rate of employment is then obtained as

$$X(t) = \int_{-\infty}^{t} x(s)e^{\alpha(s-t)}ds. \tag{7.1}$$

The time differentiation of (7.1) yields

$$\dot{X}(t) = -\alpha X(t) + x(t). \tag{7.2}$$

Note that when $X(t) = 1$ and $x(t) = \alpha$, full employment continues.

At each instance t, each firm $i \in [0, 1]$ offers a (take-it-or-leave-it) wage $w(i, t) > 0$ that applies equally for all of its employed workers. Define

$$W(t) \equiv \frac{\int_0^1 \sigma(i)w(i,t)\,di}{\int_0^1 \sigma(i)di} \tag{7.3}$$

as the economy-wide average wage, where $\sigma(i)$ is the weight given to firm i. When unemployed, each worker receives unemployment benefits $b(t)$, which we later normalize to zero for all t.[14]

We suppose that workers in this economy have a strong preference for being treated fairly, and their motivation and productivity depend heavily on this fairness concern. To be more precise, let $\omega(t)$ denote the nominal wage level that is perceived as fair by employed workers at time t. In what follows, we simply refer to $\omega(t)$ as the fair wage. When $w(i,t) \geq \omega(t)$, each worker perceives that the current wage is "fair" and produces θ (per unit of time). When $\omega(t) > w(i,t)$, each worker perceives that the current wage is "unfair" and produces nothing by withholding work effort. Note that this specification is a variant of the fair wage–effort hypothesis, put forth by Akerlof and Yellen (1990), which posits that a worker's productivity falls when the wage level dips below what is perceived as fair.

assumptions should thus be viewed as a way to capture this aspect of reality in an analytically tractable manner.

[14]The presence of unemployment benefits plays no role in the analysis, but it helps in clarifying the definition of the fair wage, which we discuss next.

2.2 The Fair Wage

Under this setup, what becomes critical is how workers develop their perception of fairness. At a general level, because fairness is inherently a distributional concern, it should be subject to various kinds of social comparisons. However, the implications of social comparisons are potentially very broad and somewhat vague. More restrictions are thus necessary to pin down a tightly-specified process of fairness perception formation. To do this, we build on the following five popular views in social psychology and behavioural economics:

1. Concerns for nominal wages: Workers care about changes in nominal wages and use them as a basis for fairness perception (Bewley 1999).[15]
2. Opinion-based transmission of perceptions: Others' perceptions of fairness influence one's own perception of fairness (Folger and Kass 2000; Umphress et al. 2003).
3. The belief in a just world: People are motivated to help others who have been treated unfairly, to make the world fair and just again (Lerner 1980).
4. The entitlement effect: Once a high wage is offered, people develop a sense of entitlement which persists over time (Falk et al. 2006).
5. The anchoring effect: One's perceptions, preferences and valuations are initially malleable but, once imprinted, become fairly persistent over time (Ariely et al. 2003).

Among these, the first three effects are necessary to derive our results while the last two are not and can be used interchangeably. We obtain the same results when at least one of them is present.[16] We also assume throughout the analysis that all workers are subject to these behavioural tendencies to the exact same extent and are equally concerned about fairness, but this is only a simplifying assumption. Our main results are robust even when only a fraction of workers have fairness concerns, as long as each worker's behavioural type (attitude towards fairness) is his or her private information (see Appendix A1 for more detail).

For illustrative purposes, we momentarily consider a discrete-time version of the model, where each firm hires new workers and revises its wage contract at interval Δt, i.e. at time $t \in T \equiv \{\ldots, t_0 - \Delta t, t_0, t_0 + \Delta t, \ldots\}$. We posit that the fair wage in this economy is computed through the following two-stage process.[17]

---

[15]Bewley (1999) interviewed more than 300 business executives and labour leaders and found that employers avoided nominal pay cuts because they feared that doing so would demoralize workers and reduce their effort. Using the survey data, which cover 123 Japanese firms, Kawaguchi and Ohtake (2007) confirm this moral theory of nominal wage rigidity.

[16]The minimal set of assumptions is, therefore, either 1–4 or 1–3 and 5; we list those psychological effects simply to support and accommodate many different scenarios.

[17]Note that the current specification is a way, possibly among some others, to capture the five aforementioned views in a unified form.

First Stage: At any $t \in T$, workers can be classified into three classes: (i) remaining workers who continue to be employed; (ii) incoming workers who are newly hired; and (iii) unemployed. The fair wage is computed mainly from the viewpoint of the remaining workers. First, those remaining workers have a (common) nominal wage level in mind, denoted by $\nu(t)$, to which they believe they are rightfully entitled (concerns for nominal wages). With this as the basis, they also take into account the well-being of unemployed workers (the just-world hypothesis).[18] The fair wage at time t_0 is obtained as the average of these concerns, weighted by the number of each class of workers:

$$\omega\left(t_0\right)=\frac{\nu\left(t_0-\Delta t\right) X\left(t_0-\Delta t\right)(1-\alpha \Delta t)+b\left(t_0\right)\left(1-X\left(t_0\right)\right)}{1-x\left(t_0\right) \Delta t}. \tag{7.4}$$

At this point, the incoming workers, with no prior work experience, basically have no idea of what is supposed to be fair and simply accept and internalize their predecessors' view (opinion-based transmission of perceptions).

Second Stage: Immediately after time t_0, the incoming workers are assimilated into the remaining workers. With new wage contracts in effect, all the employed workers, both remaining and incoming, then readjust their perception, $\nu(t)$. Two factors enter into this readjustment process. First, their perception is influenced to some extent by the current average wage (the entitlement effect). Second, it is influenced by their past perceptions (the anchoring effect). Given this, their adjusted perception is obtained as

$$\delta W\left(t_0\right)+(1-\delta) \omega\left(t_0\right)=\nu\left(t_0\right) X\left(t_0\right)+b\left(t_0\right)\left(1-X\left(t_0\right)\right), \tag{7.5}$$

where $\delta \in [0, 1]$ measures the relative salience of the entitlement effect. As we will see shortly, however, δ has no impact on the equilibrium dynamics, meaning that only one of the two effects is actually necessary to obtain our results.

The fair wage is shaped by going through this two-stage process repeatedly over time. Because we normalize $b(t) = 0$ for all t, combining (7.4) and (7.5) yields

$$\omega\left(t_0\right)=\frac{\left(\delta W\left(t_0-\Delta t\right)+(1-\delta) \omega\left(t_0-\Delta t\right)\right)(1-\alpha \Delta t)}{1-x\left(t_0\right) \Delta t}, \tag{7.6}$$

which characterizes how the fair wage in this economy evolves over time.

[18]Because workers are all homogeneous and unemployed workers are simply unlucky to be in that state, workers are, to some extent, willing to take a wage cut by lowering their fair wage.

2.3 The Equilibrium Wage Adjustment

We assume perfect competition among firms, each of which acts as a price-taker in the goods market. Let $P(t)$ denote the aggregate price of the good. At time $t \in T$, each firm unilaterally offers a wage $w(i, t)$ to its employed workers, taking the fair wage $\omega(t)$ as given. Because the cost of unemployment is prohibitively large for each worker, there is a fine line between $X(t) < 1$ and $X(t) = 1$, and the nature of wage setting differs completely, depending on whether or not the economy achieves full employment.[19]

We first characterize the equilibrium wage adjustment process in the presence of unemployment, i.e. $X(t) < 1$. In this case, workers are completely immobile and have no choice but to accept the wages offered by their respective firms. While bargaining power is entirely in the hands of firms, workers can withhold work effort whenever they feel they are slighted. With this fairness concern as a credible threat, firms can lower wages only down to the fair wage level. The following is a formal representation of this fact.

Proposition 1 *When $X(t) < 1$, $w(i, t) = \omega(t)$ for all $i \in [0, 1]$.*

Proof First, if firm i chooses to hire a worker, it must be that $w(i, t) \geq \omega(t)$. To see this, note that if $\omega(t) > w(i, t) > 0$, the worker who never quits produces nothing and, consequently, yields a negative profit. It is then strictly better not to hire the worker in the first place. Second, it is also straightforward to see that firms have no incentive ex post to offer a wage that is strictly larger than the fair wage, because that would only decrease their profits. It follows from these observations that firms simply offer the fair wage at every instance when $X(t) < 1$. □

The proposition means that the evolution of the fair wage totally dictates the equilibrium wage dynamics in the presence of unemployment. First, because $w(i, t) = \omega(t)$ for all i, we have $W(t) = \omega(t)$ for any given weight $\{\sigma(i)\}_{i=0}^{1}$. It then follows from (7.3) and (7.6) that

$$W(t_0) = W(t_0 - \Delta t)(1 - \alpha \Delta t) + W(t_0)\, x(t_0)\, \Delta t.$$

The equilibrium wage dynamics are governed by this adjustment process. Two properties of this process are worth noting. First, the fair wage level depends heavily on the inflow of workers, $x(t)$, as they internalize their predecessors' fairness perceptions. Second, the speed of adjustment is also a function of α, which is the inverse of the average duration of employment. The fair wage adjusts more slowly

[19] It is important to note that although the current setup exhibits a sharp discontinuity as $X(t)$ approaches unity, this does not mean that agents, firms and workers, all of a sudden change the way they behave at this critical point. The fundamental principle behind each agent's behaviour remains the same, regardless of the aggregate unemployment rate. Each firm pays just enough to retain and induce effort from its workers; each worker simply chooses to exert effort if the wage exceeds the fair wage level and withhold it if not.

when workers, on average, stick with their respective jobs for a longer period of time. Letting $\Delta t \to 0$, we obtain

$$\frac{\dot{W}(t_0)}{W(t_0)} = \lim_{\Delta t \to 0} \frac{W(t_0) - W(t_0 - \Delta t)}{W(t_0 - \Delta t)\, \Delta t} = \lim_{\Delta t \to 0} \frac{x(t_0) - \alpha}{1 - x(t_0)\, \Delta t} = x(t_0) - \alpha. \quad (7.7)$$

The nature of wage setting drastically changes once the economy achieves full employment, i.e. $X(t) = 1$. Workers are now fully mobile in search of the best wage offer available to them without facing any risk of being unemployed. Consequently, the market price $P(t)$ dictates the equilibrium dynamics, and the fair wage may play no role. For simplicity, firms are assumed to produce commodities using only labour, and the labour productivity is θ, which is constant. With such production technology, because workers are mobile, competition among firms for workers drives up the wage offers to $\theta P(t)$. Therefore, we can establish the following result.

Proposition 2 *When* $X(t) = 1,\ w(i,\ t) = \theta P(t)$ *for all* $i \in [0, 1]$.

Proof First, it is evident that no firm offers $w(i,\ t) > \theta P(t)$ because it yields a strictly negative profit. Suppose that $\theta P(t) > w_{\max} \equiv \max_i w(i, t)$. Then, a firm can offer a wage that is slightly larger than $w_{\max}$ and attract all workers away from other firms. This means that the only equilibrium with perfect worker mobility is to offer $w(i,\ t) = \theta P(t)$. □

In either case, whether or not there is unemployment, $w(i, t) = W(t)$ for all $i \in [0, 1]$. Perfect market competition then forces all firms to break even and earn zero profit in equilibrium. Therefore, the following zero-profit condition always holds:

$$\theta P(t) = W(t). \quad (7.8)$$

3 General Equilibrium

In the previous section, we derived the nominal wage adjustment process that stems from workers' fairness concerns. We now incorporate this process into a dynamic general equilibrium framework that admits the possibility of equilibrium unemployment. The problem we consider is a standard dynamic money-in-utility optimization problem with perfect foresight. A conventional interpretation of the money-in-utility specification is that it is a reduced-form representation of the cash-in-advance constraint, i.e. consumers value money for what it can purchase. In contrast, we would like to emphasize more the possibility that consumers derive utility directly from holding money per se, which proves to be critical when we interpret the long-run consequences of the model.

The lifetime utility of a representative household is given by

$$U = \int_0^{\infty} [u(c) + v(m)] \exp(-\rho t)\, dt, \quad (7.9)$$

where ρ is the subjective discount rate, c is real consumption and m ($\equiv M/P$) is real money balances (in what follows, we abbreviate the time notation to simplify exposition). Throughout the analysis, we restrict attention to the case where ρ is relatively large so that $\rho > \alpha$.[20]

The representative household maximizes U subject to the flow budget equation:

$$\dot{m} = wXL - \pi m - c - z, \tag{7.10}$$

where L is the amount of labour that the household desires to supply, π is the inflation rate, z is the lump-sum tax-cum-subsidy and w is the real wage, W/P. Because all firms earn zero profit under perfect market competition, the only storable assets in this economy are the real balances m.[21] The first-order optimal condition of this problem is

$$\eta\frac{\dot{c}}{c} + \rho + \pi = \frac{v'(m)}{u'(c)}, \text{ where } \eta \equiv -\frac{u''(c)c}{u'(c)}, \tag{7.11}$$

and the transversality condition is

$$\lim_{t\to\infty} \lambda(t)m(t)\exp(-\rho t) = 0, \tag{7.12}$$

where $\lambda(t)$ is the costate variable of m, which equals $u'(c)$.

The government's budget constraint is

$$z + \mu m = g, \tag{7.13}$$

where g represents government purchases and μ is the monetary expansion rate:

$$\frac{\dot{M}}{M} = \mu. \tag{7.14}$$

Given the definition of m, this can also be written as

$$\frac{\dot{m}}{m} = \mu - \pi. \tag{7.15}$$

The general equilibrium properties of the model differ sharply, depending on whether or not there is unemployment. In the presence of unemployment, (7.7) and (7.8) give

[20] If the job-separation rate α is regarded as the rate of death as a worker, it is naturally less than the subjective discount rate ρ.

[21] Even if there are government bonds, the present analysis is unchanged because the Ricardian equivalence holds.

$$\pi = \frac{\dot{W}}{W} = x(t) - \alpha. \tag{7.16}$$

On any equilibrium path, all employed workers are motivated enough to exert effort and, hence, the total production is always θXL, as discussed in the previous section. Therefore, under perfect commodity price adjustment, we must have

$$c + g = \theta XL. \tag{7.17}$$

From (7.2), (7.11) and the time derivative of (7.17),

$$\frac{v'(m)}{u'(c)} - \pi - \rho = \eta \left(\frac{c+g}{c} \right) \left(\frac{x}{X} - \alpha \right). \tag{7.18}$$

Combined with (7.16) and (7.17), (7.18) yields

$$\frac{\dot{P}}{P} = \pi = \frac{c}{c + \theta L \eta} \left(\frac{v'(m)}{u'(c)} - \rho + \frac{\alpha \eta \theta L}{c} \left(\frac{c+g}{\theta L} - 1 \right) \right), \tag{7.19}$$

which is our version of the Phillips curve. Substituting (7.19) into (7.11) and (7.15) and rearranging the results produce

$$\begin{aligned} \frac{\dot{m}}{m} &= \mu - \frac{c}{c + \theta L \eta} \left(\frac{v'(m)}{u'(c)} - \rho + \frac{\alpha \eta \theta L}{c} \left(\frac{c+g}{\theta L} - 1 \right) \right), \\ \frac{\dot{c}}{c} &= \frac{\theta L}{c + \theta L \eta} \left(\frac{v'(m)}{u'(c)} - \rho - \alpha \left(\frac{c+g}{\theta L} - 1 \right) \right). \end{aligned} \tag{7.20}$$

These two equations constitute an autonomous dynamic system with respect to m and c in the presence of unemployment.

If full employment is realized ($X = 1$), on the other hand, we have

$$c = \theta L - g. \tag{7.21}$$

From (7.11), (7.15) and (7.21), we obtain

$$\frac{\dot{m}}{m} = \rho + \mu - \frac{v'(m)}{u'(\theta L - g)},$$

which is the same as the standard dynamics of the money-in-utility model (see e.g. Blanchard and Fischer 1989). From the beginning, therefore, P takes the level that satisfies

$$\rho + \mu = \frac{v'(M/P)}{u'(\theta L - g)}, \tag{7.22}$$

and rises at the same pace as μ, so that M/P remains at the constant level that satisfies (7.22). From (7.8), $\theta P = W$ and, hence, W also rises at the rate of μ.

Remark Under the current Phillips curve (7.19), the inflation rate is governed by the liquidity premium $v'(m)/u'(c)$, the subjective discount rate ρ and the unemployment rate $1 - (c + g) / (\theta L)$. More importantly, substituting (7.11) into this, we obtain

$$\pi = \frac{\dot{c}}{\theta L} + \alpha \left(\frac{c + g}{\theta L} - 1 \right) \qquad \text{if } X < 1,$$

$$\pi = \mu \qquad \text{if } X = 1. \tag{7.23}$$

This indicates that as the economy approaches the steady state ($\dot{c} \to 0$), the inflation rate depends solely on the unemployment rate. Hence, the current framework provides a microeconomic foundation for the relationship between the unemployment rate and the inflation rate, as the original Phillips curve posits. Note that the Poisson rate α of job separation determines the adjustment speed of prices in the neighborhood of the steady state. Once full employment is reached, from (7.8) and (7.22), P and W rise at the rate of μ, implying that the Phillips curve forms a vertical line as μ changes.

4 Temporary Unemployment

4.1 Dynamics Under Satiable Liquidity Preferences

Using the dynamic equations obtained in the previous section, we draw the phase diagram and analyse the dynamic properties of the model. This section first considers the case with no liquidity trap, where the marginal utility of liquidity has no positive lower bound, i.e. $\lim_{m \to 0} v'(m) = 0$, so that the demand for liquidity would eventually dissipate. Under this condition, any equilibrium path reaches the full-employment steady state and unemployment occurs only during the adjustment process.

From (7.20), the boundary curve of m dynamics and that of c dynamics are given by

$$\dot{m} = 0: \; v'(m) = \left(\rho - \frac{\alpha \eta \theta L}{c} \left(\frac{c + g}{\theta L} - 1 \right) + \mu \left(1 + \frac{\theta L \eta}{c} \right) \right) u'(c),$$

$$\dot{c} = 0: \; v'(m) = \left(\rho + \alpha \left(\frac{c + g}{\theta L} - 1 \right) \right) u'(c). \tag{7.24}$$

The right-hand side of the $\dot{m} = 0$ curve is obviously a decreasing function with respect to c, whereas the right-hand side of the $\dot{c} = 0$ curve can be sloped either positively or negatively. If the aggregate demand $c + g$ is less than the full-employment supply θL, and μ is non-negative, the right-hand side of the $\dot{m} = 0$

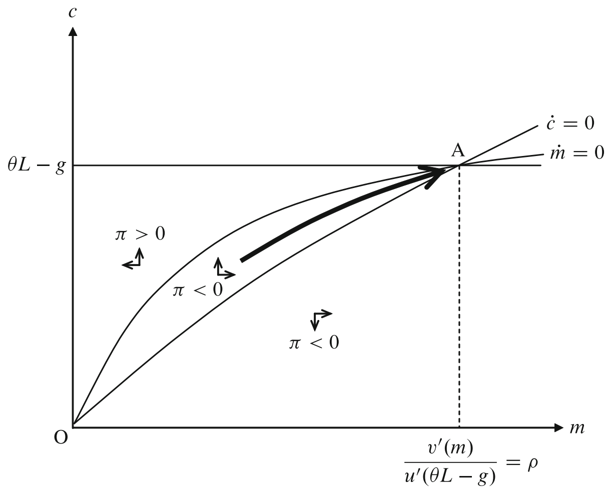


Fig. 7.1 Full-employment steady state

curve is larger than that of the $\dot{c} = 0$ curve. Thus, because $v''(m) < 0$, the $\dot{m} = 0$ curve is located on the left-hand side of the $\dot{c} = 0$ curve. Figure 7.1 illustrates the two curves in the case where $\mu = 0$ and the right-hand side of the $\dot{c} = 0$ curve is negatively sloped so that the $\dot{c} = 0$ curve is positively sloped.[22] In this case, the intersection point of the two curves is given by A, which satisfies (7.21) and (7.22), where $\mu = 0$. Therefore, there is a unique saddle path that converges to A, which we refer to as the full-employment path.[23] Along the path, the inflation rate gradually converges from negative to zero.

If $\mu > 0$, the steady state given by (7.20) does not exist within the range where $c + g \leq \theta L$. This is illustrated in Fig. 7.2. Then, the economy traces DE and reaches E, at which point (7.21) and (7.22) are valid, within a finite time and thereafter stays there with the inflation rate of μ.

4.2 *Policy Implications*

4.2.1 Monetary Policy

Using the phase diagram we now examine policy implications under temporary unemployment. We start with the effect of a monetary expansion on the dynamic

[22]When $g = 0$, it is valid if $\eta\rho > \alpha$ and $\rho > \alpha$. Note that a positively sloped unique saddle path obtains even if the boundary curve is negatively sloped.

[23]c is a jump variable whereas m is not, because M is predetermined and P is given by the zero-profit condition and is, hence, tied to W.

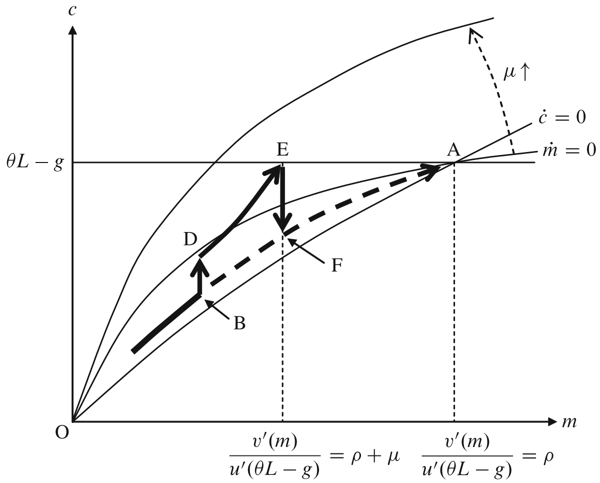


Fig. 7.2 Monetary expansion when approaching full employment

path. As Fig. 7.2 shows, an increase in μ lowers m but leaves c unaffected in the new steady state. Therefore, if μ unexpectedly increases when the economy passes through B, it jumps up to D and thereafter follows DE. Eventually, it reaches E, where the inflation rate is μ. Hence, an increase in μ is translated into a one-for-one increase in the inflation rate. If the monetary authority reduces the monetary expansion rate to zero in the new steady state so as to stabilize prices, private consumption stagnates in the short run, as the economy jumps down to F and moves along FA. Therefore, the monetary authority faces a short-run tradeoff between price stability and consumption. It should gradually and intermittently decrease μ without any notice so as to avoid a sudden downward spike in private consumption, which should be a very difficult task to carry out.

It is also important to note that the effect of a monetary expansion (an increase in the money expansion rate μ) is clearly different from a one-time increase in the money stock m in the current setup. To see this, suppose that the economy is at B. An increase in m then triggers a discrete jump from B towards A without affecting the steady-state levels of c and m. The economy moves along the same dynamic path and eventually reaches the same steady state. In particular, if the increase is large enough to move the economy all the way to A, full employment is immediately achieved.

This feature distinguishes the current analysis from Krugman (1998), who analyses Japan's stagnation in a two-period framework.[24] In his model, prices are rigid only in the initial period and any demand shortages disappear in the second, so that full employment is assumed in the end. He then argues that a monetary expansion in the second period, not in the first, stimulates current consumption

[24]Eggertsson and Krugman (2012) introduce an upper bound of debt into such a model and analyse debt-driven slumps.

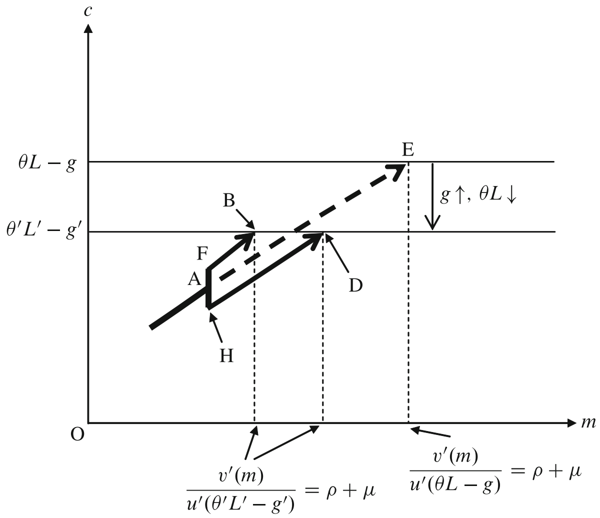


Fig. 7.3 Fiscal expansion when approaching full employment

because it raises the future price and, hence, lowers the real interest rate. However, this result is hard to interpret because his analysis does not differentiate the effect of a one-time increase in the money stock from that of an increase in the monetary expansion rate, thereby leaving some important policy questions unanswered.[25] This limitation stems from the fact that there are only two periods, and full employment is assumed in period 2, so that it fixes the endpoint of employment adjustment exogenously. In contrast, by working with an infinite-horizon model, we are able to endogenize the endpoint, which allows us to elucidate the effects of economic policy in a truly dynamic context. An increase in the money stock immediately stimulates consumption without affecting the steady state. A sustained increase in the money expansion rate also raises consumption but the change is triggered by a shift in the steady state and a consequent increase in the inflation rate, which yields an effect equivalent to that of what Krugman refers to as inflation targeting.

4.2.2 Fiscal Policy

The effect of an increase in the government purchases, g, is illustrated in Fig. 7.3. It decreases both c and m in the full-employment steady state given by (7.21) and (7.22). From (7.20), around the steady state the dynamics of m and c are represented by

[25]What is especially problematic, when it comes down to policy issues, is that it is not clear from his analysis whether a one-time increase in the money stock is sufficient or a continuous injection of money is necessary to stimulate aggregate demand.

$$\frac{\dot{m}}{m} = \frac{\eta\theta L\mu}{\theta L - g + \theta L\eta} \quad \text{and} \quad \frac{\dot{c}}{c} = \frac{\theta L\mu}{\theta L - g + \theta L\eta},$$

and, hence, on the equilibrium path c and m move so that they satisfy

$$\left.\frac{m}{c}\frac{dc}{dm}\right|_{\text{path}} = \frac{\dot{c}/c}{\dot{m}/m} = \frac{1}{\eta}. \tag{7.25}$$

From (7.21) and (7.22), on the other hand, the change in the steady state induced by an increase in g must satisfy

$$\left.\frac{m}{c}\frac{dc}{dm}\right|_{\text{steady state}} = \frac{\eta_m}{\eta} \quad \text{where} \quad \eta_m = -\frac{v''(m)m}{v'(m)}\,(>0)\,, \tag{7.26}$$

where η_m is the elasticity of the marginal utility of money. By comparing (7.25) and (7.26), we find that an increase in g shifts the steady state from E to B if $\eta_m < 1$ and to D if $\eta_m > 1$, as illustrated in Fig. 7.3. In either case, an increase in g leads to a one-for-one decrease in c in the long run (complete crowding out). In the short run, however, it may or may not stimulate private consumption, depending on the elasticity of money utility. Specifically, c increases if the elasticity is smaller than one and decreases if the elasticity is larger than one, because B and D are, respectively, located above and below the previous equilibrium path. If a fiscal expansion occurs at A, for instance, the path jumps to either F or H and thereafter traces the new saddle path.

4.2.3 Labour Supply and Productivity

From (7.21) and (7.22) it is also clear that an increase in labour productivity, θ, or in each household's labour supply, L, increases both c and m in the full-employment steady state. In fact, the steady-state level of c is directly given by $\theta L - g$, whereas that of m is obtained as a function of $\theta L - g$. This suggests that an increase in θL yields an effect that is equivalent to a fiscal contraction (a decrease in g). The same argument as above then applies, as illustrated in Fig. 7.3. Although an increase in θL unambiguously increases consumption in the steady state, it may or may not stimulate consumption in the short run, depending on the elasticity of money utility. The latter part of this result is closely related to Eggertsson and Krugman (2012), who find that an increase in labour supply could reduce aggregate employment in the short run (a phenomenon that is often referred to as the "paradox of toil"), by assuming a liquidity trap. In the current setup, the paradox of toil also arises in the short run when $\eta_m < 1$, i.e. liquidity preferences are not too responsive to the amount of money.

4.2.4 Summary

To sum up, along the full-employment path, any government interventions, either monetary or fiscal, affect aggregate demand in the short run, one way or the other. However, because the economy eventually reaches full employment, any policies have only a transitory effect on aggregate demand. To put it another way, market forces bring the economy back on track in the long run. We summarize our findings as follows.

Proposition 3 *If the marginal utility of liquidity has no positive lower bound, i.e.* $\lim_{m\to\infty} v'(m) = 0$*, there is a unique equilibrium path that reaches the full-employment steady state. Along the full-employment path:*

- *A monetary expansion raises private consumption in the short run, while it does not affect private consumption in the long run.*
- *A fiscal expansion totally crowds out private consumption in the long run, while it may or may not increase private consumption in the short run.*
- *An increase in labour productivity or each household's labour supply expands private consumption in the long run, while it may or may not increase private consumption in the short run.*

5 Persistent Unemployment

5.1 Dynamics Under Insatiable Liquidity Preferences

We have so far examined the case where the full-employment steady state, given by (7.21) and (7.22), exists and is eventually reached. Under certain conditions, however, our model also admits the possibility of persistent unemployment, where a full-employment steady state fails to exist. In this case, market forces alone are not sufficient to bring the economy back on track, and the government may have some role to play, even in the long run. It should be noted that the presence of persistent unemployment in this economy is not caused by real wage rigidity which tips the balance on the supply side of the market. Although wages are sluggish because of fairness concerns, prices are fully flexible so that the real wage rate is always equal to the marginal productivity, i.e. $\theta = w(= W/P)$, from (7.8). Persistent unemployment arises as a monetary phenomenon driven purely from the demand side.

The money demand curve is given by the relationship between m and the liquidity premium, $v'(m)/u'(c)$, which represents the nominal interest rate. It is negatively sloped because the liquidity premium decreases as m increases. We now introduce a liquidity trap to this, which arises if $v'(m)$ has a positive lower bound:

$$\lim_{m\to\infty} v'(m) = \beta > 0. \tag{7.27}$$

This property captures what Keynes envisioned as the essence of a monetary economy, i.e. an economy in which there is an asset whose marginal utility stays positive (Keynes 1936).[26] We take the presence of the lower bound literally as a formal representation of our perpetual craving for liquid assets; namely, money.[27] As demonstrated in Ono (1994, 2001) and also in this model, this constant and perpetual craving for money is the driving force behind persistent demand shortages.

We now characterize equilibrium dynamics under (7.27). If the full-employment supply θL is sufficiently large and satisfies

$$\frac{\beta}{u'(\theta L - g)} > \rho + \mu, \tag{7.28}$$

the liquidity premium exceeds the time preference for any m whenever c attains its full-employment level. This means that there exists no m that can satisfy (7.22) and, hence, the full-employment steady state fails to exist. In this case, along the boundary curve of c given in (7.24), c gradually approaches c^u, defined by

$$R^u \equiv \frac{\beta}{u'(c^u)} = \rho + \alpha\left(\frac{c^u + g}{\theta L} - 1\right). \tag{7.29}$$

If μ is sufficiently small, there exists an equilibrium path that leads to this stagnation steady state (the stagnation path, for short), which satisfies the transversality condition.

[26] Using aggregate quarterly data in Japan and the Japanese survey data called NIKKEI RADAR, Ono et al. (2004) empirically find this property to be well supported using both parametric and non-parametric methods.

[27] Although an assumption of this kind is somewhat non-standard in modern macroeconomics, we argue that it is not as provocative as some may think. To make a case for this, we focus on the nature of money as a reward as opposed to other rewards, called "primary rewards" in neuroscience, that are indispensable for the survival of the species. The satiable nature of the demand for primary rewards such as food is intuitively clear. For instance, we all know from our experiences that if we eat enough, the marginal value of food eventually diminishes to zero and even becomes negative. This does not necessarily mean, however, that we can apply this same logic to secondary rewards, which derive their value from their associated primary rewards. In fact, evidence in neuroscience suggests that the diminishing nature of marginal utility (called "devaluation" in the field) for primary and secondary rewards is represented in different brain areas. Evidence, based both on animal studies (Balleine and Dickinson 1998, 2000; Balleine and Ostlund 2007) and on human studies (Valentin et al. 2007), suggests that the orbitofrontal cortex is the region responsible for the devaluation of primary rewards, while Pine et al. (2009) find that the devaluation of secondary rewards is represented in the dorsal striatum. Although research along this line is still at its infant stage and there remains a lot to be discovered, these findings seem to confirm a belief that the valuation of secondary rewards is related more to human cognition, whereas that of primary rewards is related more to sheer instinct, indicating that simplistic analogy between primary and secondary rewards is not necessarily warranted. We thank S. Tanaka for enlightening us on the subject.

Proposition 4 *Suppose that the marginal utility of liquidity has a positive lower bound, i.e.* $\lim_{m\to\infty} v'(m) = \beta > 0$. *If (7.28) holds, the full-employment steady state does not exist. Moreover, if* μ *is small enough to satisfy*

$$\frac{\beta}{u'(c^u)} > \mu, \tag{7.30}$$

there arises a unique equilibrium path that leads to the stagnation steady state. Along the stagnation path, the steady-state consumption level converges to c^u*, which falls short of the full-employment level* $\theta L - g$*, and deflation occurs.*

Proof We can show that there exists a well-defined c^u such that

$$0 < c^u < \theta L - g.$$

To see this, define the difference of the two sides of (7.29) as

$$\Phi(c) \equiv \frac{\beta}{u'(c)} - \rho - \alpha\left(\frac{c+g}{\theta L} - 1\right),$$

which is continuous in c. The intermediate-value theorem then guarantees the existence of a well-defined c^u if

$$\Phi(\theta L - g) > 0 > \lim_{c\to 0}\Phi(c) = -\rho - \alpha\left(\frac{g}{\theta L} - 1\right).$$

The first inequality directly follows from (7.28), whereas the second inequality holds by assumption. This property also implies $\Phi'(c) > 0$ around the stagnation steady state (where $\Phi(c) = 0$) and, hence,

$$\frac{\eta\beta\theta L}{c^u u'(c^u)}\left(= \frac{\eta R^u \theta L}{c^u}\right) > \alpha. \tag{7.31}$$

Because $\Phi(c)$ is the inside value of the parenthesis of $\dot{c}/c$ in (7.20) when the stagnation steady state is reached, the boundary curve of the c dynamics is positively sloped with respect to c around the stagnation steady state, as illustrated in Fig. 7.4. The saddle-path stability around the stagnation steady state is proved to be valid in Appendix A2.

From (7.19) and (7.29), the steady-state inflation rate is then obtained as

$$\pi^u = \alpha\left(\frac{c^u + g}{\theta L} - 1\right) < 0. \tag{7.32}$$

From (7.20), (7.29) and (7.32), the transversality condition (7.12) is valid if and only if

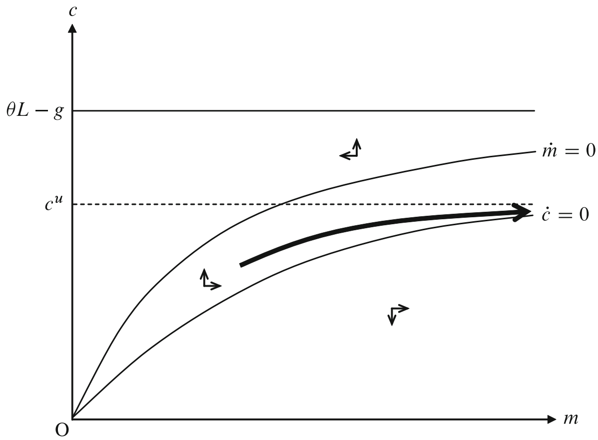


Fig. 7.4 Stagnation steady state

$$0 > \frac{\dot{m}}{m} - \rho = \mu - \frac{\beta}{u'(c^u)}.$$

Evidently, this condition holds when μ is so small as to satisfy (7.30). □

This situation is depicted in Fig. 7.4. In sum, on the equilibrium path, the economy never reaches the full-employment steady state. Along the stagnation path, the price level continuously falls, resulting in persistent deflation, which makes the real balances m keep expanding while satisfying the transversality condition. The consumption level falls short of the full-employment level $\theta L - g$ and, due to this insufficient level of aggregate demand, unemployment persists in the steady state.

5.2 Policy Implications

5.2.1 Monetary Policy

We are now in a position to draw policy implications in the stagnation phase with persistent unemployment. We again start with the effect of a monetary expansion on aggregate demand. An important observation here is that c^u is independent of μ, so that a change in μ does not affect either the upper bound of c or the steady-state inflation rate as long as μ satisfies (7.28) and (7.30). This fact leads to an implication of the effectiveness of monetary policies in the stagnation phase quite different from that in the case of temporary unemployment.

There are two cases we need to consider, but the implications are roughly the same. Suppose first that c^u is not much smaller than the full-employment level $\theta L - g$ such that

$$\frac{\beta}{u'(\theta L - g)} - \rho < \frac{\beta}{u'(c^u)}. \tag{7.33}$$

It then follows from (7.28) and (7.30) that

$$\mu < \frac{\beta}{u'(\theta L - g)} - \rho \Rightarrow \text{stagnation},$$

$$\frac{\beta}{u'(\theta L - g)} - \rho < \mu < \frac{\beta}{u'(c^u)} \Rightarrow \text{both stagnation and full employment},$$

$$\frac{\beta}{u'(c^u)} < \mu \Rightarrow \text{full employment}.$$

Only the stagnation path exists and, hence, the economy is trapped in the stagnation phase when μ is relatively small (less than $\beta/u'(\theta L - g) - \rho$). Because c^u is independent of μ, an expansion of μ is totally ineffective as far as μ remains within the range satisfying (7.28). This situation is illustrated in Fig. 7.5. Once the monetary authority raises μ enough to violate (7.28) and, hence, enable m to validate (7.22), the full-employment steady state, represented by E in Fig. 7.6, is restored. Consequently, there emerges a new equilibrium path given by BE, which leads to the full-employment steady state. Note, however, that the stagnation path AU also exists in this range. In the presence of multiple paths, there is no guarantee that the economy jumps to the new full-employment path, because it requires coordination among all agents. This means that the monetary authority needs an extra push, if it is to bring the economy out of the stagnation phase. This can, in fact, be done by raising μ even higher, in which case only the full-employment path survives. A policy intervention of this kind apparently comes with a cost, however, as the economy necessarily suffers from high inflation. The classic dilemma between inflation and unemployment surfaces, now with a long-run implication. If the monetary authority reduces μ to stabilize prices, the full-employment steady state disappears and the economy returns to the stagnation path AU, which results in persistent unemployment.

In contrast, if c^u is much smaller than $\theta L - g$ and

$$\frac{\beta}{u'(\theta L - g)} - \rho > \frac{\beta}{u'(c^u)},$$

we find

$$\mu < \frac{\beta}{u'(c^u)} \Rightarrow \text{stagnation},$$

$$\frac{\beta}{u'(c^u)} < \mu < \frac{\beta}{u'(\theta L - g)} - \rho \Rightarrow \text{no equilibrium path},$$

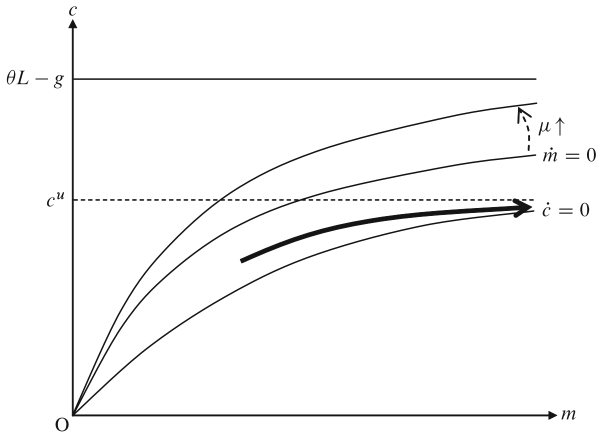


Fig. 7.5 Monetary expansion under persistent stagnation (when μ is small)

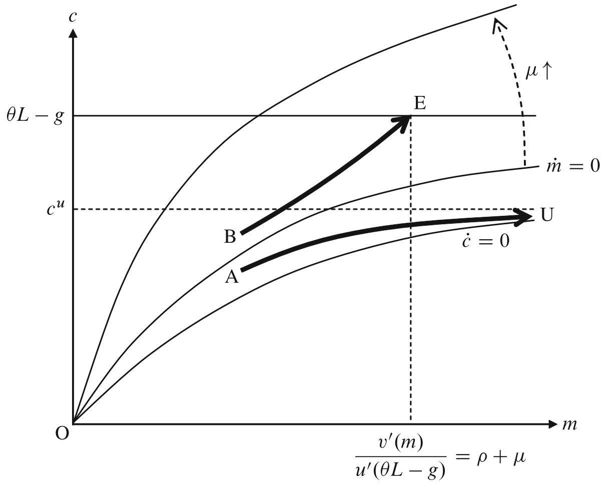


Fig. 7.6 Monetary expansion under persistent stagnation (when μ is large)

$$\frac{\beta}{u'(\theta L - g)} - \rho < \mu \Rightarrow \text{full employment.}$$

The situation is the same as the first case (where c^u is not much smaller) when μ is either relatively small (less than $\beta/u'(c^u)$) or relatively large (more than $\beta/u'(\theta L - g) - \rho$). Only the stagnation path exists in the former case, while only the full-employment path exists in the latter. For cases in between these two, no equilibrium path exists because the saddle path in Fig. 7.5 does not satisfy the transversality condition, which is different from the first case. Otherwise, we can draw the same conclusion that an increase in μ is totally ineffective and of no help to get the economy out of the stagnation phase as far as it is less than $\beta/u'(c^u)$.

This result is quite different from that of Ono (2001). Using a model with an exogenously given Phillips curve, in which it is simply assumed that an increase in μ, under any circumstances, results in a one-to-one increase in the inflation rate, Ono concludes that an increase in μ stimulates consumption both in the short run and in the steady state. Thus, the result implies that monetary policy is always effective or, at least, its effect is independent of the aggregate state of the economy. We view this conclusion as unsatisfactory because the recent Japanese experience suggests that monetary expansions have virtually no impact on deflation or national income while the economy is trapped in the stagnation phase.

In contrast, using a model with a microeconomic foundation of the wage adjustment mechanism, we show that the effect of monetary policy under persistent unemployment differs sharply from that under temporary unemployment. In the model, the deflation rate is independent of μ under persistent unemployment, while the inflation rate equals μ when the economy achieves full employment. What this suggests is that monetary expansions have no effect on the dynamic path as far as μ remains less than a certain level, and can stimulate consumption only if their magnitude is substantial enough to make μ greater than the critical level. Thus, if the economy is trapped in a serious stagnation, it requires very high inflation to get out of the situation. Moreover, the economy inevitably suffers from chronic high inflation once full employment is realized. This draws a clear contrast with the case of temporary unemployment in which monetary expansions are not required to reach full employment. Under temporary unemployment, the economy is bound to reach unemployment at some point and monetary expansions merely shorten the time to get there. We argue that this is one of the key insights of the model and provides an important policy implication for stagnant economies.

We would also like to note the difference between an increase in the money expansion rate and a one-time increase in the nominal money stock M, as in the case of temporary unemployment. Because a one-time increase in M affects neither of the two boundary curves in (7.24), it merely creates a discrete jump toward the steady state along the same dynamic path. Therefore, such a policy intervention can stimulate consumption, c, in the short run but its effect becomes negligibly small if the economy is in the vicinity of the steady state.

5.2.2 Fiscal Policy

We now turn to the effect of a fiscal expansion along the stagnation path, which is also drastically different from that along the full-employment path. This is illustrated in Fig. 7.7. Because (7.29) and (7.31) yield

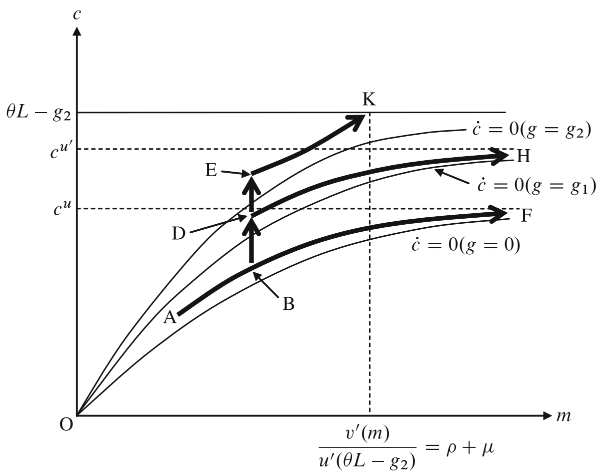


Fig. 7.7 Fiscal expansion under persistent stagnation

$$\frac{dc^u}{dg} = \frac{\alpha}{\left(\frac{\eta\beta\theta L}{c^u u'(c^u)} - \alpha\right)} > 0,$$

the equilibrium path shifts upward from AF to DH and, thus, c rises both in the short run (from B to D) and in the long run (from c^u to $c^{u\prime}$). If the government expands g sufficiently to satisfy

$$\rho > \frac{\beta}{u'(\theta L - g)},$$

as in the case of g_2, the full-employment steady state, K in the figure, is eventually reached. It should be noted that if the government reduces g to the previous level, consumption c returns to the previous level as well. To maintain a high level of consumption, the government must continuously hold g at the same level. Note that fiscal expansion stimulates consumption by creating new employment and thereby reducing the rate of decline in wages and prices. This mechanism is very different from the conventional Keynesian multiplier effect or that in the new Keynesian models (e.g. Mankiw 1988; Startz 1989; Christiano et al. 2011). The conventional multiplier effect is produced by an increase in income, while the new Keynesian multiplier effect is produced by an increase in labour supply. In contrast, the present multiplier effect is generated by a decrease in the deflation rate caused by new employment created by the government.

5.2.3 Labour Supply, Productivity and Flexibility

As in the case of temporary unemployment, an increase in labour productivity, θ, or each household's labour supply, L, is equivalent to a fiscal contraction (a decrease in g). Because c equals c^u given by (7.29), from (7.29) and (7.31) one finds

$$\frac{dc^u}{d(\theta L)} = -\frac{\alpha(c^u + g)}{\left(\frac{\eta\beta\theta L}{c^u u'(c^u)} - \alpha\right)\theta L} < 0.$$

Thus, an increase in θ or L worsens deflation and, hence, decreases private consumption both in the short run and in the long run. This means that the paradox of toil arises not only temporarily but also permanently in our framework. Note that this property is quite different from that in the case of temporary unemployment, where, as mentioned in Proposition 1, an increase in θ or L always increases consumption in the steady state while it may or may not increase short-run consumption.

Aside from these implications, our model also yields an implication that is closely related to what Eggertsson and Krugman (2012) term the "paradox of flexibility", which means that wage and price flexibility do not facilitate recovery from a recession during a liquidity trap. In our model, the speed of wage adjustment is governed by the rate of job separation, α, where a higher α means less inertia in wage adjustment. We can show that an increase in α lowers consumption both in the short run and in the long run. To see this, note that (7.29) and (7.31) yield

$$\frac{dc^u}{d\alpha} = \frac{c^u + g - \theta L}{\frac{\eta\beta\theta L}{c^u u'(c^u)} - \alpha} < 0,$$

which suggests that the paradox of flexibility appears both temporarily and permanently. The paradox of flexibility arises in our model because an improvement in the wage adjustment exacerbates deflationary pressures and makes holding money less costly. Note that the role of nominal rigidity differs fundamentally from the conventional new Keynesian models where nominal rigidity is the direct cause of demand shortages.[28] This also contrasts with the case of temporary unemployment in which an increase in α obviously shortens the adjustment period that is required to reach full employment and, hence, raises private consumption both temporarily and permanently.

5.2.4 Summary

Because market forces cannot be relied upon to restore full employment automatically, the role of the government is inherently different in the stagnation phase. Government interventions are not only effective, even in the long run, but also indispensable to realize full employment. Their effects in the stagnation phase are generally opposite to those under full employment. We summarize our findings as follows.

[28] As mentioned at the outset of this section, the real wage always equals the marginal productivity of labour, θ, due to perfect flexibility of commodity prices, so that there is no real rigidity in our model.

Proposition 5 *Along the stagnation path:*

- *A monetary expansion is totally ineffective and has no impact either on private consumption or on the inflation rate as long as the monetary expansion rate* μ *is less than* $\min\left(\frac{\beta}{u'(c^u)}, \frac{\beta}{u'(\theta L - g)} - \rho\right)$.
- μ *must be larger than* $\max\left(\frac{\beta}{u'(c^u)}, \frac{\beta}{u'(\theta L - g)} - \rho\right)$ *to be sure of restoring full employment.*
- *A fiscal expansion raises private consumption, both in the short run and in the long run.*
- *An increase in labour productivity or each household's labour supply decreases private consumption, both in the short run and in the long run (the paradox of toil).*
- *An increase in the probability of job separation decreases private consumption, both in the short run and in the long run (the paradox of flexibility).*

6 Conclusion

We introduce various concepts of fairness into the fair wage model of Akerlof and Yellen (1990) and relate it to the dynamic general equilibrium model of a monetary economy portrayed by Keynes (1936). In this "Akerlof–Yellen meets Keynes" framework, we derive a version of the Phillips curve with some microeconomic foundation. We show that the inflation rate is governed by the liquidity premium, the subjective discount rate and the unemployment rate. In particular, around the steady state, the inflation rate depends only on the unemployment rate, as in the original Phillips curve. More importantly, we show that unemployment caused by demand shortages may arise in the steady state. The effects of monetary and fiscal expansions are very different under temporary unemployment and under persistent unemployment.

An increase in the monetary expansion rate raises short-run aggregate demand when the economy is on the full-employment path, although the steady-state aggregate demand is unaffected. In contrast, if unemployment continues under persistent demand shortages, the monetary expansion rate must be higher than a certain level to locate the economy on the path that eventually reaches full employment. Otherwise, an expansionary monetary policy is totally ineffective in stimulating aggregate demand. Therefore, to get the economy out of persistent and serious stagnation, the monetary expansion rate must be significantly high. This obviously comes with a cost, however, because the economy inevitably suffers from high inflation once full employment is realized.

The effect of an increase in government purchases is also quite different under temporary unemployment and under persistent unemployment. An increase in government purchases can either increase or decrease private consumption in the short run, depending on the elasticity of the marginal utility of money, when the economy is on the full-employment path. If the elasticity is higher (lower) than one, it decreases (increases) private consumption in the short run. Because the economy eventually reattains full employment, however, an increase in government purchases totally crowds out private consumption in the long run. In contrast, under persistent stagnation, an increase in government purchases never fails to increase private consumption both in the short run and in the long run.

As a final note, we would like to point out that our model exhibits the two paradoxes discussed extensively in Eggertsson and Krugman (2012): the paradox of toil and the paradox of flexibility. First, we show that the effect of an increase in labour productivity, or an increase in each household's labour supply, is equivalent to a decrease in government purchases in our model, implying that the paradox of toil appears. Second, what is perhaps more intriguing is that nominal rigidity in our model directly implies the paradox of flexibility. A less frictional wage adjustment process, which results from an increase in the rate of job separation, raises private consumption under temporary unemployment by abridging the adjustment process, but reduces private consumption under persistent unemployment by exacerbating deflationary pressures. Thus, the model shows that the role of nominal rigidity differs sharply, and the paradox of flexibility arises once the economy becomes trapped in the stagnation phase.

Acknowledgements We would like to thank K. Hashimoto, S. Ikeda, T. Matsumura, R. Murota, K. Sugawara, S. Tanaka and the anonymous referee for their valuable comments. We are also grateful to D. Flath for carefully editing our manuscript. Helpful comments from seminar and lecture participants at the Tokyo Institute of Technology, Kobe University, Kyushu University, the University of Graz and the Paris School of Economics are gratefully acknowledged. This research is financially supported by the Global Center of Excellence Program, the Ministry of Education, Culture, Sports, Science and Technology; and Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science.

Appendices

A1. The Case Where Only a Fraction of Workers Have Fairness Concerns

Throughout the analysis, we have assumed that all agents are subject to behavioural biases to the same extent and, in particular, are equally concerned about fairness. Here, we relax this assumption and show that our results still hold under some conditions even when only a fraction of workers are concerned about fairness. More precisely, we show that the firm still chooses to pay the fair wage if: (i) the reservation wage for the workers is positive; and (ii) the fraction of workers with fairness concerns is sufficiently large.

Now consider a setting in which only a fraction $s \in (0, 1)$ of workers have fairness concerns and would withhold effort if the wage were below the fair wage level. We assume that the firm cannot observe each worker's behavioural type, at any point in time, so that there is no feasible way to discriminate among workers.[29] Finally, let $\underline{w}(t) \in (0, \omega(t))$ denote the reservation wage for the workers at time t, e.g. disutility of labour, which is strictly positive but lower than the fair wage.

Because each worker's behavioural type is his or her private information, and there is no feasible way to solicit this information, there are virtually only two choices for the firm, either to pay the reservation wage or pay the fair wage for all the workers. If the firm chooses to pay only the reservation wage, those workers with fairness concerns withhold work effort and the instantaneous profit (per worker) is $(1-s)\,\theta P(t) - \underline{w}(t)$. Because the firm can guarantee itself zero profit by offering the equilibrium fair wage, it is evident that the firm would pay the fair wage for all of its workers if s were sufficiently close to one.

A2. Stability Under Persistent Stagnation

This appendix shows the saddle-path stability of the dynamics. In the case where the full-employment steady state exists, there is a unique equilibrium path because the present model has essentially the same structure as the standard money-in-utility model. Thus, we focus on the stability when the economy is on the path that leads to the stagnation steady state.

For simplicity, we consider the case where $\mu = 0$ and $g = 0$. Because m diverges to infinity, we consider $h\ (= 1/m)$ instead of m and examine the stability of the two dynamic equations derived from (7.20):

$$\dot{h} = \frac{c}{c + \theta L\eta}\left(\frac{v'(1/h)}{u'(c)} - \rho + \alpha\eta\left(1 - \frac{\theta L}{c}\right)\right)h,$$

$$\dot{c} = \frac{\theta L c}{c + \theta L\eta}\left(\frac{v'(1/h)}{u'(c)} - \rho - \alpha\left(\frac{c}{\theta L} - 1\right)\right),$$

around the stagnation steady state where $c = c^u$ and $h = 0$. The partial derivatives of the above two equations around the stagnation steady state are

$$\frac{\partial \dot{h}}{\partial h} = \pi^u + \frac{c^u \eta_m R^u}{c^u + \theta L\eta},$$

$$\frac{\partial \dot{h}}{\partial c} = h \times d\left[\frac{c}{c + \theta L\eta}\left(\frac{\beta}{u'(c)} - \rho + \alpha\eta\left(1 - \frac{\theta L}{c}\right)\right)\right]/dc$$

$$\frac{\partial \dot{c}}{\partial c} = \frac{\eta R^u \theta L - \alpha c^u}{c^u + \theta L\eta},$$

[29] In other words, we assume that the firm cannot observe each worker's individual output either.

$$\frac{\partial \dot{c}}{\partial h} = \frac{\theta L c^u \eta_m R^u}{(c^u + \theta L \eta)\, h}.$$

Since $\eta_m = 0$ because of the liquidity trap, from (7.31) and (7.32), we find

$$\left(\frac{\partial \dot{h}}{\partial h}\right)\left(\frac{\partial \dot{c}}{\partial c}\right) - \left(\frac{\partial \dot{h}}{\partial c}\right)\left(\frac{\partial \dot{c}}{\partial h}\right) = \frac{(\eta R^u \theta L - \alpha c^u)\, \pi^u}{c^u + \theta L \eta} < 0,$$

i.e. one of the characteristic roots is positive and the other is negative. Note that c is jumpable, whereas $h\,(= 1/m = P/M)$ is not because W is not jumpable and $P = W/\theta$ from (7.8). We can thus conclude that the path is saddle-path stable.

Addendum: Liquidity Trap and Long-run Stagnation[30]

B1. Short-Run and Long-Run Stagnation Models

In the present model insatiable liquidity preference is assumed. It creates a liquidity trap with a strictly positive nominal interest rate. In contrast, Krugman (1998) regards the zero lower bound of the nominal interest rate as a liquidity trap. Besides the level difference in the lower bound of the nominal interest rate, the two models have basically the same structure. However, this difference leads to significantly different implications of stagnation in the two models; long-run stagnation arises only with a strictly positive lower bound of the nominal interest rate.

Of the optimal household condition (7.11)[31]:

$$\rho + \eta \frac{\dot{c}}{c} + \pi = R = \frac{v'(m)}{u'(c)} > 0, \tag{7.34}$$

the left-hand side represents time preference (i.e., a desire for present consumption) while the right-hand side implies liquidity preference (i.e., a desire to accumulate money). If the liquidity preference exceeds the time preference when the consumption level is large enough to attain full employment, the household lowers consumption and aggregate demand shortages arise.

Insatiable liquidity preference is mathematically expressed as $v'(m)$ having a strictly positive lower bound β. If β is so large, or full-employment output θL is so large, as to satisfy

$$\rho < \frac{\beta}{u'(\theta L)},$$

[30]This addendum has been newly written by Yoshiyasu Ono for this book chapter.

[31]This is a continuous version of Equation (7.3) in Krugman (1998, p. 145).

which is (7.28) in which $g = 0$ and $\mu = 0$, the left-hand side of (7.34) is less than the right-hand side when π is zero and c is constant at θL –i.e., aggregate demand shortages occur in the steady state. To obtain the steady state in this case, one has to model a price-wage adjustment mechanism that is consistent with the presence of aggregate demand shortages in the steady state.[32] Ono and Ishida (2014) formulate such a mechanism by extending the efficiency wage hypothesis of Akerlof (1982) and Akerlof and Yellen (1990) to a dynamic setting. This is mentioned in Sect. 2 of this chapter.

If $v'(m)$ reaches zero as m expands, as is the case in Krugman's model (1998), the liquidity premium $v'(m)/u'(c)$ reaches zero in (7.34). Furthermore, c is constant and π is zero in the steady state. Therefore, if P is low enough so that m is large enough, the left-hand side of (7.34) (which equals ρ) exceeds the right-hand side (which is $v'(m)/u'(c) = 0$), implying that a desire for consumption dominates a desire to accumulate money, and aggregate demand shortages disappear.

Moreover, in order for aggregate demand shortages to arise in the short run, Krugman adopts a period analysis and imposes two important assumptions –i.e., (1) the present price cannot be revised, and (2) the future price perfectly adjusts so that full employment is reached. In this setting the present price is historically given and the future price is determined by the future money supply. Therefore, if the future money supply is small, the future price is low and thus the present π is negative. If it is very negative, the left-hand side of (7.34) (which equals $\rho + \pi$) is smaller than the right-hand side (which is $v'(m)/u'(c) = 0$) and aggregate demand shortages arise in the present, although they disappear in the future.

Obviously, if the monetary authority supplies a sufficiently large amount of money in the future, the future price is high. Therefore, the present π can be large enough to make the left-hand side of (7.34) positive, which stimulates present consumption. As the future money supply is larger, the present consumption is higher and eventually full employment is achieved. This is so-called the inflation targeting policy.

Because of the period analysis structure the time span of each period is exogenously given. Therefore, the present price rigidity and the future price flexibility straightforwardly determine the present π without having an intertemporal price adjustment mechanism. Thus, Krugman's model must be a period analysis and cannot treat long-run stagnation. It fits short-run recession while the present model fits such long-run stagnation as Japan has been suffering for more than two decades.

B2. Wealth Preference and the Zero Interest Rate

In the present model the nominal interest rate stays strictly positive while in Krugman's model it is zero. In reality, the nominal interest rate has long been almost zero in Japan; how can it be explained in the present model?

[32] The new Keynesian price-wage adjustment assumes away aggregate demand shortages.

In the present model liquidity preference is insatiable. Thus, it should exhibit a desire to hold wealth rather than the transaction motive. In the text m is the only asset and hence utility of m may be reinterpreted as a desire to hold wealth. However, if there are both money m and interest-earning assets b, the household's utility must be

$$U = \int_0^\infty [u(c) + v(m) + \sigma(a)]\, e^{-\rho t} dt,$$

where $\sigma(a)$ implies a desire to hold wealth a $(= m + b)$ and $v(m)$ represents the transaction motive. Whereas $v(m)$ satisfies the normal condition:

$$v'(\infty) = 0, \tag{7.35}$$

$\sigma(a)$ satisfies the insatiability:

$$\sigma'(\infty) = \beta_a > 0. \tag{7.36}$$

The household maximizes U subject to the flow budget equation and the asset constraint:

$$\begin{aligned} \dot{a} &= ra + wx - c - Rm, \\ a &= m + b, \end{aligned} \tag{7.37}$$

and the first-order optimal conditions (7.34) are replaced by

$$\rho + \eta\frac{\dot{c}}{c} + \pi = R + \frac{\sigma'(a)}{u'(c)} = \frac{v'(m)}{u'(c)} + \frac{\sigma'(a)}{u'(c)}. \tag{7.38}$$

Along the money demand function:

$$R = \frac{v'(m)}{u'(c)}, \tag{7.39}$$

which is obtained from the second equality in (7.38), R approaches zero as deflation continues and m diverges to infinity, as is seen from (7.35). Nevertheless, aggregate demand shortages arise in the steady state as long as (7.36) is valid and β_a satisfies

$$\rho < \frac{\beta_a}{u'(\theta L)} < \frac{\sigma'(a)}{u'(\theta L)}.$$

Note that R in (7.39) is the nominal interest rate that is observed in the market, which approaches zero as m expands. The interest rate that stays strictly positive and yields aggregate demand shortages is $\sigma'(a)/u'(c)$, which is unobservable.

Therefore, the observed interest rate is zero and yet aggregate demand shortages arise in the steady state.

B3. Status Preference and Insatiable Wealth Preference

Ono and Yamada (2012) present a model of status preference with respect to asset holdings and show that the insatiable wealth preference represented by (7.36) is valid. They assume the following utility:

$$U = \int_0^\infty [u(c) + v(m) + \sigma(a, \overline{a})] e^{-\rho t} dt, \tag{7.40}$$

where status preference $\sigma(a, \overline{a})$ has either of the two forms:

$$\text{Case D (difference)} : \sigma(a, \overline{a}) = \sigma_D(a - \overline{a}),$$

$$\text{Case R (ratio)} : \sigma(a, \overline{a}) = \sigma_R(a/\overline{a}). \tag{7.41}$$

The representative household maximizes U in (7.40) subject to (7.37) and the first-order optimal conditions are

$$\rho + \eta\frac{\dot{c}}{c} + \pi = R + \frac{\sigma_a(a,\ \overline{a})}{u'(c)} = \frac{v'(m)}{u'(c)} + \frac{\sigma_a(a,\ \overline{a})}{u'(c)}. \tag{7.42}$$

Because $a = \overline{a}$, $\sigma_a(a,\ \overline{a})$ in each case of (7.41) satisfies

$$\text{Case D} : \sigma_a(a,\ \overline{a}) = \sigma'_D(0) > 0 \quad \text{for any } a,$$

$$\text{Case R} : \sigma_a(a,\ \overline{a}) = \sigma'_R(1)/\overline{a}, \quad \lim_{\overline{a}\to\infty} \sigma_a(a, \overline{a}) = 0. \tag{7.43}$$

In the presence of aggregate demand shortages, m expands to infinity and the optimal conditions given by (7.42) reduce to:

$$\rho + \eta\frac{\dot{c}}{c} + \pi = \frac{\sigma_a(a, \overline{a})}{u'(c)},\ \ R = 0.$$

Therefore, the status preference in case D of (7.43) plays the same role as the positive lower bound of the wealth preference given by (7.36) and yields persistent shortages of aggregate demand. In case R of (7.43), however, $\sigma_a(a, \overline{a})$ reaches zero and hence aggregate demand shortages do not arise in the steady state.

Ono and Yamada (2012) use data of an affluence comparison experiment by Yamada and Sato (2013) and show that the difference specification fits the data much better than the ratio specification. Therefore, persistent stagnation arises.

References

Akerlof GA (1982) Labor contracts as partial gift exchange. Q J Econ 97(4):543–569

Akerlof GA (2007) The new case for keynesianism. Challenge 50(4):5–16

Akerlof GA, Yellen JL (1990) The fair-wage effort hypothesis and unemployment. Q J Econ 105(2):255–284

Ariely D, Loewenstein G, Prelec D (2003) Coherent arbitrariness: stable demand curves without stable preferences. Q J Econ 118(1):73–105

Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37:407–419

Balleine BW, Dickinson A (2000) The effect of lesions of the insular cortex on instrumental conditioning: evidence for a role in incentive memory. J Neurosci 20:8954–8964

Balleine BW, Ostlund SB (2007) Still at the choice-point: action selection and initiation in instrumental conditioning. Ann N Y Acad Sci 1104:147–171

Bewley TF (1999) Why wages don't fall during a recession. Harvard University Press, Cambridge, MA

Blanchard OJ, G Dell'Ariccia, P Mauro (2010) Rethinking macroeconomic policy. IMF Staff Position Note SPN/10/03

Blanchard OJ, Fischer S (1989) Lectures on macroeconomics. MIT Press, Cambridge, MA

Cardullo G (2010) Matching models under scrutiny: an appraisal of the shimer puzzle. J Econ Surv 24:622–656

Christiano L, Eichenbaum M, Rebelo S (2011) When is the government spending multiplier large? J Polit Econ 119(1):78–121

Dotsey M, King RG, Wolman AL (1999) State-dependent pricing and the general equilibrium dynamics of money and output. Q J Econ 114(2):655–690

Eggertsson GB, Krugman P (2012) Debt, deleveraging, and the liquidity trap: a fisher-Minsky-Koo approach. Q J Econ 127(3):1469–1513

Falk A, Fehr E, Zehnder C (2006) Fairness perceptions and reservation wages -the behavioral effects of minimum wage laws. Q J Econ 121(4):1347–1381

Folger R, Kass EE (2000) Social comparison and fairness: a counter-factual simulations perspective. In: Suls J, Wheeler L (eds) Handbook of social comparison: theory and research. Plenum, New York, pp 423–441

Gali J (2008) Monetary policy, inflation, and business cycle: an introduction to the New Keynesian framework. Princeton University Press, Princeton

Gertler M, Leahy J (2008) A Phillips curve with an Ss foundation. J Polit Econ 116(3):533–572

Golosov M, Lucas RE Jr (2007) Menu costs and Phillips curves. J Polit Econ 115(2):171–199

Hashimoto K (2004) Intergenerational transfer and effective demand. Econ Bull 5(1):1–13

Hashimoto K, Ono Y (2011) Does pro-population policy raise per capita consumption? Jpn Econ Rev 62(2):151–169

Johdo W (2006) Geographical space and effective demand under stagnation. Aust Econ Pap 5(1):286–298

Johdo W, Hashimoto K (2009) International relocation, the real exchange rate and effective demand. Jpn World Econ 21(1):39–54

Kawaguchi D, Ohtake F (2007) Testing the morale theory of nominal wage rigidity. Ind Labor Relat Rev 61(1):59–74

Keynes JM (1936) The general theory of employment, interest and money. Macmillan, London

Krugman P (1998) It's Baaack!! Japan's slump and the return of the liquidity trap. Brook Pap Econ Act 2:137–187

Lerner MJ (1980) The belief in a just world: a fundamental delusion. Plenum, New York

Mankiw NG (1988) Imperfect competition and the Keynesian cross. Econ Lett 26(1):7–13

Mankiw NG (2008) What would Keynes have done? New York Times, 28 Nov, http://www.nytimes.com/2008/11/30/business/economy/30view.html?_r=0

Mankiw NG, Reis R (2002) Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve. Q J Econ 117(4):1295–1328

Matsuzaki D (2003) The effects of a consumption tax on effective demand under stagnation. Jpn Econ Rev 54(1):101–118

Murota R, Ono Y (2011) Growth, stagnation and status preference. Metroeconomica 62(1): 112–149

Murota R, Ono Y (2012) Zero nominal interest rates, unemployment, excess reserves and deflation in a liquidity trap. Metroeconomica 63(2):335–357

Ono Y (1994) Money, interest, and stagnation. Oxford University Press, Oxford

Ono Y (2001) A reinterpretation of chapter 17 of Keynes's general theory: effective demand shortage under dynamic optimization. Int Econ Rev 42(1):207–236

Ono Y (2006) International asymmetry in business activity and appreciation of a stagnant country's currency. Jpn Econ Rev 57(1):101–120

Ono Y (2010) Japan's long-run stagnation and economic policies. In: Bateman B, Hirai T, Marcuzzo MC (eds) The return to Keynes. Harvard University Press, Cambridge, MA, pp 32–50

Ono Y (2014) International economic interdependence and exchange-rate adjustment under persistent stagnation. Jpn Econ Rev 65(1):70–92

Ono Y, Ishida J (2014) On persistent demand shortages: a behavioral approach. Jpn Econ Rev 65(1):42–69

Ono Y, Ogawa K, Yoshida A (2004) Liquidity trap and persistent unemployment with dynamic optimizing agents: empirical evidence. Jpn Econ Rev 55(4):355–371

Ono Y, Yamada K (2012) Difference or ratio: implications of status preference on stagnation. ISER discussion papers 856, Osaka University, Revised 2014

Pine A, Seymour B, Roiser JR et al (2009) Encoding of marginal utility across time in the human brain. J Neurosci 29(30):9575–9581

Rodriguez-Arana A (2007) Inflation and the public deficit when the utility of money is insatiable. Jpn Econ Rev 58(2):238–254

Rogerson R, Shimer R (2011) Search in macroeconomic models of the labor market. In: Ashenfelter O, Card D (eds) Handbook of labor economics 4(A). North Holland, Amsterdam, pp 619–700

Shapiro C, Stiglitz J (1984) Equilibrium unemployment as a worker discipline device. Am Econ Rev 74(3):433–444

Shimer R (2005) The cyclical behavior of equilibrium unemployment and vacancies. Am Econ Rev 95(1):25–49

Startz R (1989) Monopolistic competition as a foundation for Keynesian macroeconomic models. Q J Econ 102(4):737–752

Umphress EE, Labianca G, Brass D, Kass EE, Scholten L (2003) The role of instrumental and expressive social ties in employees' perceptions of organizational justice. Organ Sci 14(6): 738–753

Valentin VV, Dickinson A, O'Doherty JP (2007) Determining the neural substrates of goal-directed learning in the human brain. J Neurosci 27(15):4019–4026

Woodford M (2003) Interest and prices. Princeton UP, Princeton

Yamada K, Sato M (2013) Another avenue for anatomy of income comparisons: evidence from hypothetical choice experiments. J Econ Behav Organ 89:35–57

Yun T (1996) Nominal price rigidity, money supply endogeneity, and business cycles. J Monet Econ 37(2):345–370

Part III
Time Preference in Macroeconomics

Chapter 8
Rate of Time Preference, Intertemporal Elasticity of Substitution, and Level of Wealth

Masao Ogaki and Andrew Atkeson

Abstract The rate of time preference (RTP) and the intertemporal elasticity of substitution (IES) are two important factors shaping intertemporal consumption decisions. Models in which the RTP and/or the IES differ systematically between rich and poor households have different empirical and policy implications for economic development, growth, and the distribution of income and consumption from those of standard models in which these parameters are constant across households. In this chapter, we estimate a model in which both RTP and IES are allowed to differ across rich and poor households using household level panel data from India. Our empirical results are consistent with the view that the RTP is constant across poor and rich households, but the IES is larger for the rich than it is for the poor.

Keywords Consumption growth • Wealth-varying RTP and IES models • India

1 Introduction

The rate of time preference (RTP) and the intertemporal elasticity of substitution (IES) are two important factors shaping consumers' intertemporal consumption decisions. In the theoretical literature, many authors have studied models in which the RTP changes with the level of wealth or consumption (see, e.g., Epstein 1983; Uzawa 1968). We call these wealth-varying RTP models. Others have studied

The original article first appeared in the Review of Economics and Statistics 79:564–572, 1997. A newly written addendum has been added to this book chapter.

M. Ogaki (✉)
Department of Economics, Keio University, 612 Mita-Kenkyu-shitsu, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
e-mail: mogaki@econ.keio.ac.jp

A. Atkeson
Department of Economics, University of California, Los Angeles, CA, USA
e-mail: andy@atkeson.net

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_8

models in which the IES changes with the level of wealth or consumption (see, e.g., Chatterjee 1994; Easterly 1994; Ogaki et al. 1996; Rebelo 1992). We call these wealth-varying IES models. In the context of models using the expected utility framework with time separable utility, the parameters governing the IES are intimately related to those governing risk aversion. Thus, in addition to the theoretical literature that discusses models in which the IES varies with the level of wealth, there is a large theoretical literature analyzing models in which the coefficient of relative risk aversion changes with the level of wealth which implicitly analyzes models in which the IES varies with the level of wealth.

In a variety of applications, models with different assumptions about how the RTP and the IES varies across households have strikingly different implications. For example, consider, in the context of a model economy with complete contingent markets, the implications of various assumptions about preferences for the evolution of the distribution of consumption. The assumption that all agents have the same RTP and IES yields the result that the ratio of the consumption of the rich to that of the poor is constant over time. The assumption that rich consumers are more patient than poor consumers, thus having a higher RTP, yields the result that the ratio of the consumption of the rich to that of the poor grows at a constant rate over time. Moreover, this growth rate of the ratio of the consumption of the rich to that of the poor does not depend on intertemporal prices. The assumption that rich consumers have a higher IES than poor consumers yields implication that depend on the path of intertemporal prices. If intertemporal prices exceed the rate of time preference, then the ratio of the consumption of the rich to that of the poor increases. If intertemporal prices fall below the rate of time preference, then the ratio of the consumption of the rich to that of the poor shrinks over time.[1] In an economy in which intertemporal prices are fluctuating both above and below the rate of time preference, the assumption that the IES rises with the level of wealth implies that the rich have more volatile consumption growth than the poor. Also, given the links between the IES and the coefficient of relative risk aversion, these models would also each have different implications for the evolution of the distribution of asset holdings or wealth as well as for the allocation of aggregate risk across consumers.

To date, there is little empirical work in which wealth-varying RTP models or wealth-varying IES models are estimated. Lawrance (1991) is a prominent exception. She estimated a wealth-varying RTP model with the Panel Study of Income and Dynamics (PSID) data set. In Lawrance's model, the IES is assumed to be constant as in the standard macroeconomic models with isoelastic utility functions. Atkeson and Ogaki (1996) estimate a wealth-varying IES model using panel data on the consumption of Indian households. In this empirical work, we assume that the RTP is constant across households.

The RTP measures the slope of the indifference curve between consumption at two adjacent dates when consumption at these dates are equal. Loosely speaking, this feature of preferences controls the mean of the consumption growth rate.

[1] We will explain these statements in detail in Sect. 3. Also see Ogaki (1992).

In a deterministic economy, the IES measures the extent to which consumption growth between two dates changes when the real interest rate between those dates changes.[2] Thus, the IES controls the consumption growth volatility in economies with fluctuating real interest rates. The common assumption that the RTP and IES are constant across poor and rich households implies that the mean and volatility of consumption growth are constant across poor and rich households, unless borrowing constraints or other market imperfections cause them to differ.

In this chapter, we use Indian panel data to estimate a model that allows both the RTP and IES to change systematically between rich and poor households. Our goal in this chapter is to provide evidence to distinguish separately the extent to which the RTP and the IES vary with the level of household wealth. As our example above illustrates, to distinguish variation in the RTP from variation in the IES, we will need to use data consistent with the hypothesis that intertemporal prices are fluctuating above and below the rate of time preference. In our application, this means that we require data in which aggregate consumption grows in some time periods and shrinks in others. One advantage of the Indian panel data we use is that they satisfy this requirement. Our empirical results with these data are consistent with the view that the RTP is constant across poor and rich households, but the IES is larger for the rich than it is for the poor.

The remainder of this chapter is organized as follows. In Sect. 2, we describe our data and present and estimate a statistical model which summarizes the extent to which the mean and volatility of consumption growth vary across households. In Sect. 3, we present economic models which might be used to interpret these statistical results. In Sect. 4, we conclude.

2 Data Description

In this section, we describe the household level panel data collected in India by the Institute for Crop Research in the Semi-Arid Tropics (ICRISAT) and our statistical model of household consumption growth. We discuss the data in detail in the Appendix.[3] We use panel data for three villages (Aurepalle, Shirapur, and Kanzara) from fiscal year 1975–1976 to fiscal year 1984–1985. (In what follows, we denote each fiscal year by its first calendar year.) Since the construction of

[2]We give formal definitions of the RTP and the IES in economies with uncertainty in Sect. 3.

[3]Following Townsend (1994), we use the consumption data in ICRISAT's summary data. There are two ways of estimating consumption using the ICRISAT data. The ICRISAT's method is to infer it from transactions. The other method is to retrieve consumption by applying flow accounting identities to the production and storage data, which Ravallion and Chaudhuri (1997) propose. Their consumption data are very different from the ICRISAT's consumption data, and the difference is correlated with income. We refer the reader to Ravallion and Chaudhuri (1997) and Townsend (pp. 554–555) for discussion of the suitability of these consumption data. It does not seem clear which consumption data set is more reliable.

food consumption was changed in 1976 and the data for nonfood consumption are missing for most categories after 1982, we choose 1976–1981 as our sample period.

2.1 The Choice of the Data

These Indian panel data have been used to study consumption smoothing and risk sharing models by many authors.[4] We find these data attractive for three reasons. First, the saving behavior of households in less developed countries is of general interest. Second, we suspect that the dependence of the IES on the level of wealth is more likely to be important for very poor households whose consumption level is near some subsistence level. Third, as we shall see, aggregate consumption in these data rises in some periods and falls in others, so that these data do not immediately contradict the hypothesis that intertemporal prices exceed the rate of time preference in some periods and fall below the rate of time preference in others.

In Table 8.1, we report average, minimum, maximum consumption per equivalent adult in terms of 1983 rupees for each of the three villages. These numbers are reported to facilitate the interpretation of the estimates of the subsistence levels that are reported below. From Table 8.1, we can see that average consumption fluctuates substantially over time in each village and that the maximum and minimum consumption levels across households are substantially different in our data.

2.2 A Statistical Model for Data Description

In this section, we present the statistical model that we use to summarize certain features of these panel data. We use this model to summarize the extent to which the mean consumption growth rate and the volatility of consumption growth varies systematically across rich and poor households. Later on, we provide an economic interpretation of this statistical model in which differences in the mean growth rate of consumption across rich and poor households is determined by differences in the RTP across rich and poor households and differences in the volatility of the growth rate of consumption across rich and poor households is determined by differences in the IES across rich and poor households. We present the statistical model first because the assumptions we need to interpret our model as an economic model are stronger than the ones we need to implement our statistical model. Hence, our statistical model may be consistent with a variety of economic models.

[4]See, e.g., Bhargava and Ravallion (1993), Jacoby and Skoufias (1988), Lim (1992), Morduch (1990, 1991), Rosenzweig (1998), Rosenzweig and Binswanger (1993), Rosenzweig and Stark (1989), Rosenzweig and Wolpin (1993), and Townsend (1994).

Table 8.1 Consumption per equivalent adult

	1976	1977	1978	1979	1980	1981	1976–1981
Average total consumption							
Aurepalle	502	490	544	750	738	660	614
Shirapur	1,063	980	749	869	787	664	852
Kanzara	852	847	758	993	937	815	867
Minimum total consumption							
Aurepalle	179	308	334	359	249	229	...
Shirapur	304	491	485	364	423	242	...
Kanzara	393	370	249	460	419	432	...
Maximum total consumption							
Aurepalle	1,300	913	984	1,574	1,945	1,510	...
Shirapur	1,703	1,723	1,974	1,525	1,596	1,676	...
Kanzara	2,694	1,509	1,290	2,189	2,085	1,945	...
Average food consumption							
Aurepalle	313	381	408	538	502	423	408
Shirapur	604	555	644	543	623	521	582
Kanzara	490	489	418	578	571	479	504
Minimum food consumption							
Aurepalle	221	178	173	214	288	187	...
Shirapur	238	137	274	221	308	102	...
Kanzara	221	178	173	214	288	187	...
Maximum food consumption							
Aurepalle	646	658	766	1,044	1,132	829	...
Shirapur	1,133	1,063	1,166	888	1,075	1,088	...
Kanzara	1,441	870	704	1,284	1,409	1,081	...
Average nonfood consumption							
Aurepalle	190	101	156	214	240	236	158
Shirapur	337	313	345	352	329	364	235
Kanzara	369	359	353	426	364	345	267
Minimum nonfood consumption							
Aurepalle	40	27	83	68	67	32	...
Shirapur	65	112	136	129	114	48	...
Kanzara	134	151	113	133	131	133	...
Maximum nonfood consumption							
Aurepalle	908	377	415	711	831	688	...
Shirapur	681	698	894	870	777	1,013	...
Kanzara	1,019	1,274	759	836	929	885	...

Consider the following statistical model of household consumption growth. Let $C_h(t)$ denote the consumption of household h at date t, and let household consumption growth be given by

$$\ln[C_h(t+1)-\gamma]-\ln[C_h(t)-\gamma]=\phi(t)+b_y y_h^c+v_h(t), \tag{8.1}$$

where $\phi(t)$ varies over time and across villages but is constant across the households in a village at date t, y_h^c is a proxy of permanent income, and $v_h(t)$ has zero mean and is uncorrelated with the level of the household's permanent income and the household's income growth.

Our focus in this chapter is on estimating the parameters γ and b_y. We can gain some intuition for the relationship between these parameters and systematic variation across poor and rich households in the mean and the volatility of their consumption growth as follow. First consider a case in which $\gamma = 0$ and $v_h(t) = 0$. If the parameter b_y is positive, then the growth rate of consumption is higher for rich households than for poor households. If b_y is negative, then the reverse is true.

Now consider a case in which $b_y = 0$ and $v_h(t) = 0$. In this case, household consumption growth is given by

$$\frac{C_h(t+1)}{C_h(t)}=\exp[\phi(t)]+\frac{\gamma(1-\exp[\phi(t)])}{C_h(t)}.$$

Thus, in this case, if the parameter γ is positive and the constant $\phi(t)$ is also positive, then the consumption growth rate of households with high levels of consumption is higher than that of households with low levels of consumption. On the other hand, if the parameter γ is positive and the constant $\phi(t)$ is negative, then the consumption of households with high levels of consumption shrinks faster than that of households with low levels of consumption. In this sense, if γ is positive, we say that the consumption growth of rich households is more volatile than that for poor households. If γ is negative, then the reverse is true.

We have chosen this statistical model for two reasons. First, this model is parsimonious. We are forced to use a parsimonious model here because our data set has only five time periods. Second, as we discuss later on, the parameters of this statistical model have a simple economic interpretation in the context of an economic model in which household consumption is chosen as if in an economy with complete markets.

We estimate this model as follows. Let y_h^p be another proxy of permanent income of household h, $y_h(t)$ be the current income of household h at date t, and $z_h(t) = (1, \ln(y_h^p), \ln[y_h(t+1)] - \ln[y_h(t)])'$ be a vector of instrumental variables. Because we assume that the error term in (8.1) is uncorrelated with income variables, these are valid instruments. As a statistical model, the model (8.1) may be misspecified in many dimensions. For the purpose of this chapter, one important type of misspecification is when the error term in (8.1) is correlated with income variables. If it is correlated with the level of household permanent income, then our parameters γ and b_y are not capturing the systematic differences in consumption growth across households. If the error term is correlated with household income growth, then household income growth will be an important missing variable. With these concerns in mind, we test the validity of the statistical model by testing whether or not the error term is correlated with the instrumental variables.

Let $p = (p_1, \ldots, p_{T+2})$ be $a(T+2)$-dimensional vector of unknown parameters. The true value of p is $p^0 = (\phi(1), \ldots, \phi(T), \gamma, b_y)'$. We define a3-dimensional vector $\xi_t^h(p)$, so that $\xi_t^h(p^0) = z_h(t)v_h(t)\exp(-\gamma/A)$, where A is a constant. Here we normalize the disturbance by $\exp(-\gamma/A)$ to avoid a trivial solution $\phi(t) = 0$ for $t = 1, \ldots, T, \gamma = -\infty, b_y = 0$. Let $\xi^h(p) = (\xi_1^h(p), \ldots, \xi_T^h(p))'$. Then we have $3T$ orthogonality conditions

$$E_H[\xi(p^0)] = \operatorname*{plim}_{N\to\infty}\left(\frac{1}{N}\right)\sum_{h=1}^{N}[\xi^h(p^0)] = 0, \tag{8.2}$$

where E_H is the expectation operator over households. A subscript H is attached to emphasize that the expectation is taken over households. We have these $3T$ orthogonality conditions for each village. We pool these orthogonality conditions for the three villages and estimate p for each village with the generalized method of moments (GMM).[5]

3 Empirical Results

In Table 8.2, we report results for real total consumption expenditure per equivalent adult. In the first panel, we report estimates of γ and b_y and test statistics. The first, second, and third rows report results when no restriction is imposed for alterative proxies of permanent income used as y_h^c; the fourth row, when one restriction $b_y = 0$ is imposed; the fifth row, when two restrictions $b_y = \gamma = 0$ are imposed. For the first and second rows, we use dummy variables based on land holding class as y_h^c. For the first row, we use the dummy variable that takes on the value of minus one for landless laborers and zero for the others. For the second row, the dummy variable takes on the value of one for large farms and zero for the others. For the third row, we use the level of real per capita total consumption in 1975 as y_h^c.

The J statistic reported in each row is Hansen's (1982) χ^2 test for the overidentifying restrictions. The C statistics reported in the first, second, third and fifth rows are the difference between the J of each row and the J of the fourth row, which

[5]See, e.g., Hansen (1982) and Gallant and White (1988). We assume that the regularity conditions of Gallant and White are satisfied. Hansen/Heaton/Ogaki's GAUSS GMM package (see Ogaki 1993b) is used for the GMM in the present paper. In pooling the data for three villages, we allow $\xi(p^0)$ to have different covariance matrices in different villages. Ogaki (1993a, Section 4.3) provides a more detailed explanation as to how the data for villages are pooled.

Table 8.2 GMM results for total consumption

Permanent income proxy	γ	s.e.	b_y	s.e.	J[a]	d.f.	p-value (%)	C[b]	d.f.	p-value (%)
Landless labor dummy	177.6	6.70	-0.021	0.049	34.44	28	18.7	0.255	1	61.4
Large farm dummy	177.4	7.22	-0.009	0.033	34.60	28	18.2	0.182	1	76.0
Consumption	177.6	6.70	-0.023	0.053	34.46	28	18.6	0.237	1	62.6
	177.2	7.45	0	...	34.70	29	21.5	...		
	0	...	0	...	98.89	30	0.0	64.428	2	0.0
	$\phi(1)$	s.e.	$\phi(2)$	s.e.	$\phi(3)$	s.e.	$\phi(4)$	s.e.	$\phi(5)$	s.e.
Aurepalle	0.017	0.198	0.163	0.041	0.475	0.052	-0.020	0.053	-0.150	0.068
Shirapur	-0.124	0.034	0.050	0.049	-0.129	0.062	0.147	0.057	-0.095	0.062
Kanzara	-0.008	0.036	-0.087	0.052	0.358	0.045	-0.139	0.034	-0.143	0.036

[a] χ^2 test statistic for overidentifying restrictions
[b] Loglikelihood ratio type test statistic for restrictions imposed

are called likelihood ratio type test statistics.[6] The C statistic in the fifth row tests the restrictions $b_y = \gamma = 0$ which corresponds with the hypothesis that there is no systematic difference in the consumption growth of the rich and the poor. The C test provides strong evidence against this hypothesis. The C statistics in the first, second, and third rows test the restriction $b_y = 0$ for the alternative proxies for permanent income used as y_h^c. There is little evidence against this hypothesis. Consistent with the C test results, γ is estimated to be statistically significantly positive, but b_y are not significantly different from zero. The J statistics in the fourth row tests the hypothesis that there exists no systematic component in consumption growth that can be explained by the income variables in the instruments once the effects of the parameters γ and $\phi(t)$ in (8.1) are taken into account. We do not reject this hypothesis, hence model (8.1) is a valid model of data description for our purpose.

We report estimates of $\phi(t)'s$ for Aurepalle, Shirapur, and Kanzara in the second panel of Table 8.2 when b_y is restricted to be zero. In this case, $\phi(t)$ is the average growth rate of $C(t) - \gamma$. We have both significantly positive values of $\phi(t)$ and significantly negative values of $\phi(t)$. This is important because, as we will discuss below, the wealth-varying IES and the wealth-varying RTP models can be

[6] See, e.g., Ogaki (1993a) for an explanation of the likelihood ratio type test in the GMM procedure. In order to compare J statistics with the C test, the same distance matrix needs to be used for unrestricted and restricted estimations. The distance matrix used is based on the estimation with the restriction $b_y = 0$. The initial distance matrix is an identity and the GMM estimation is iterated three times. The constant A for normalization was set to 200 for total consumption expenditure and food in Tables 8.2 and 8.3 and to 50 for nonfood consumption in Table 8.4. The final results were virtually the same when A was increased to 300 for total consumption and food and to 100 for nonfood but convergence for the initial distance matrix needed more iterations.

Table 8.3 GMM results for food consumption

Permanent income proxy	γ	s.e.	b_y	s.e.	J[a]	d.f.	p-value (%)	C[b]	d.f.	p-value (%)
Landless labor dummy	101.5	3.93	0.063	0.048	30.57	28	33.6	1.711	1	19.1
Large farm dummy	101.4	4.14	0.056	0.034	29.62	28	38.2	2.666	1	10.3
Consumption	101.4	4.30	-0.083	0.360	30.69	28	33.1	1.597	1	20.6
	101.5	3.70	0	...	32.28	29	30.8	...	...	
	0	...	0	...	56.93	30	0.0	64.428	2	0.0
	$\phi(1)$	s.e.	$\phi(2)$	s.e.	$\phi(3)$	s.e.	$\phi(4)$	s.e.	$\phi(5)$	s.e.
Aurepalle	0.362	0.077	0.057	0.034	0.383	0.050	-0.090	0.050	-0.274	0.049
Shirapur	-0.101	0.044	0.146	0.059	-0.193	0.063	0.158	0.058	-0.216	0.075
Kanzara	-0.025	0.040	-0.190	0.053	0.375	0.035	-0.051	0.043	-0.152	0.063

[a] χ^2 test statistic for overidentifying restrictions
[b] Loglikelihood ratio type test statistic for restrictions imposed

Table 8.4 GMM results for nonfood consumption

Permanent income proxy	γ	s.e.	b_y	s.e.	J[a]	d.f.	p-value (%)	C[b]	d.f.	p-value (%)
Landless labor dummy	26.8	1.45	-0.021	0.047	28.03	28	46.3	0.192	1	66.1
Large farm dummy	26.8	1.45	-0.010	0.036	28.15	28	45.7	0.073	1	78.7
Consumption	28.8	1.44	-0.014	0.059	28.17	28	45.6	0.053	1	81.9
	26.8	1.44	0	...	28.22	29	50.6	...	...	
	0	...	0	...	35.69	30	0.0	35.687	2	0.0
	$\phi(1)$	s.e.	$\phi(2)$	s.e.	$\phi(3)$	s.e.	$\phi(4)$	s.e.	$\phi(5)$	s.e.
Aurepalle	-0.970	0.124	0.828	0.083	0.294	0.072	0.047	0.065	0.124	0.118
Shirapur	-0.141	0.047	0.051	0.071	0.050	0.068	0.021	0.066	0.060	0.079
Kanzara	-0.027	0.039	0.043	0.054	0.234	0.056	-0.211	0.025	-0.153	0.039

[a] χ^2 test statistic for overidentifying restrictions
[b] Loglikelihood ratio type test statistic for restrictions imposed

discriminated sharply only when the data contain both periods in which aggregate consumption grows and those in which it shrinks.

We report results when $C_h(t)$ is taken as food in Table 8.3 and results when $C_h(t)$ is taken as nonfood in Table 8.4. The results for food and nonfood are qualitatively similar to those for total consumption. We find no evidence against the hypothesis b_y is zero, and that the estimates of γ are significantly positive.

Thus for each of these categories of consumption, we find that our results are consistent with the view that the consumption growth mean is the same between rich and poor households, and that rich households have more volatile consumption growth households than poor households.

4 Interpreting the Results

In this section, we discuss economic models that may be used to interpret our empirical results in the previous section. In the first subsection, we describe an economic model in which consumers have preferences with wealth-varying IES and RTP and in which markets allow for full risk sharing. We explain that the results are consistent with the hypothesis that the IES rises with the level of wealth and that the RTP is constant across wealth levels. In the second subsection, we show that this economic model can be used to formally motivate our statistical model. In the third subsection we speculate on how our results might be interpreted in the context of models with incomplete markets. In the fourth subsection, we report some test results concerning the suitability of our statistical model in the context of a model with incomplete markets.

4.1 A Model of Wealth-Varying RTP and IES

In this section, we present the model of consumers' intertemporal allocation of consumption expenditure that we use to motivate our estimation. In particular, we discuss the different implications of wealth-varying RTP and wealth-varying IES models for consumption growth in the context of a model with complete markets.

Consider an economy with H households, each of which consumes a good in each of T time periods. Let the consumer $h, h = 1, \ldots, H$, have time and state separable utility with an intratemporal utility function $u(C_h(t))$. Let a vector $s(t), s(t) = 1, 2, \ldots, S$, denote the state of the world in each period and the vector $e(t) = [s(0), s(1), \ldots, s(t)]$ be the history of the economy. The consumer h maximizes

$$U_h = \sum_{t=0}^{T} \sum_{e(t)} (\beta_h)^t \mathrm{Prob}(e(t)|e(0)) u(C_h(t, e(t))), \tag{8.3}$$

subject to a life time budget constraint

$$\sum_{t=0}^{T} \sum_{e(t)} \prod_{\tau=0}^{t} R[\tau - 1, e(\tau - 1), e(\tau)]^{-1} C_h(t, e(t)) \leq W_h(0). \tag{8.4}$$

Here $W_h(0)$ is the consumer $h's$ initial wealth and T can be either a finite number as in the life-cycle model or infinity as in the dynasty model. The term $Prob(e(t)|e(\tau))$ denotes the conditional probability of $e(t)$ given $e(\tau)$, and $R(t-1, e(t-1), e(t))$ is the (gross) asset return of the state contingent security for the event $e(t)$ in terms of the good in the event $e(t-1)$ at period $t-1$. We will often suppress $e(t)$ to simplify the notation below.

In (8.3), β_h is the consumer $h's$ discount factor. Following Lawrance (1991), we assume that β_h can be different across consumers, but is constant over time for each consumer. This constant discount factor assumption greatly simplifies the empirical work. One interpretation of this assumption is as an approximation to a model in which the discount factor actually changes as a consumer becomes wealthier as in Uzawa (1968), but is roughly constant for each household over the short sample period (6 years) covered in our data because household consumption is also roughly constant over this time period. This interpretation seems valid because the variation of consumption across households is generally much larger than the range of consumption fluctuations for each household in our data set.

In our model, the allocation of household expenditure over time is guided by the intertemporal first order condition

$$\frac{u'(C_h(t, e(t)))}{u'(C_h(t+1, e(t+1)))} = \beta_h R^*[t, e(t), e(t+1)].$$

Taking logs of both sides of this equation and using a first order Taylor approximation around $C_h(t)$ gives the result that the growth of consumption ($\hat{C}(t) = \log[C(t+1)] - \log[C(t)]$) is given approximately by

$$\hat{C}^h(t) \cong \sigma_h(t)\{r(t) - \delta_h\}, \tag{8.5}$$

where $\delta_h = \ln(1/\beta_h)$ is the RTP, $r(t) = \ln(R^*(t))$, $R^*(t) = R(t)Prob(e(t+1)|e(t))$, and $\sigma_h(t) = -u'(C_h(t))C_h(t)/u''(C_h(t))$ is the IES. From (8.5), $\sigma_h(t) \cong \partial C_h(t)/\partial r(t)$. If there is no uncertainty, $r(t)$ is the real interest rate.

The distinct implications for consumption growth of models in which the RTP varies systematically with wealth and models in which the IES varies systematically with wealth can be seen in Eq. (8.5). If δ_h falls systematically as wealth rises as Lawrance's (1991) estimates suggest, then the consumption growth rate of the poor is always lower than the consumption growth rate of the rich. As long as σ is constant, there will be no systematic difference in the consumption growth volatility between the rich and the poor. On the other hand, if σ_h rises systematically with wealth, then the consumption growth rate of the rich will be higher than that of the poor in the period in which $r(t) > \delta$ and the consumption of the rich will shrink faster than that of the poor in the in the period in which $r(t) < \delta$. Hence the rich will have more volatile consumption growth than the poor as r varies around δ. Note that, in an economy with constant $r(t)$, it is impossible to discriminate between a model in which δ_h falls with wealth and a model in which σ_h rises with wealth and δ_h is constant and is less than $r(t)$. However, we can discriminate between these models in economies in which $r(t)$ fluctuates by examining how the consumption growth volatility changes with wealth.

Thus our empirical results that the consumption growth mean is the same for rich and poor households and that consumption growth of rich households is more volatile than that of poor households are consistent with the view that the RTP is constant, but the IES rises as households become richer.[7]

4.2 Motivating the Statistical Model

The model in the previous subsection can be used to formally motivate (8.1) when we parameterize the utility function by the quasi homothetic Geary-Stone utility function:

$$u(C_h) = \frac{1}{1-\alpha}[(C_h - \gamma)^{(1-\alpha)} - 1], \tag{8.6}$$

where $\alpha > 0$. We will refer to the parameter γ as the subsistence parameter and the parameter α as the curvature parameter.

$$\sigma_h = \frac{1}{\alpha}\left(1 - \frac{\gamma}{C_h}\right). \tag{8.7}$$

If $\gamma > 0$, then the IES of the poor is smaller than that of the rich. For a poor household, C_h is close to γ and σ is close to zero. For a rich household, γ/C_h is close to zero and σ is close to $1/\alpha$. Thus the intertemporal elasticity of substitution rises with the level of wealth. On the other hand, the IES falls with the level of wealth if $\gamma < 0$.[8]

The intertemporal first order condition of the model is

$$\left[\frac{C_h(t, e(t)) - \gamma}{C_h(t+1, e(t+1)) - \gamma}\right]^{-\alpha} = \beta_h R^*(t, e(t), e(t+1)). \tag{8.8}$$

We assume that consumption $C_h(t)$ is measured with error in the following form:

$$C_h^m(t) - \gamma = [C_h(t) - \gamma]\epsilon_h(t), \tag{8.9}$$

[7]Since a positive γ implies more volatile consumption growth for the rich than that for the poor, our empirical results from the Indian data are in line with Mankiw and Zeldes's (1991) finding that consumption growth is more volatile for stockholders than nonstockholders in the PSID. In fact, because of the links between the IES and the coefficient of relative risk aversion, a model with wealth-varying IES would predict that the wealthy should hold a disproportionate share of aggregate risk and have more volatile consumption than the poor.

[8]It should be noted that there is no theoretical reason to exclude the case where $\gamma < 0$. In fact, in this context, this subsistence parameter is merely a convenient way to allow the curvature of the utility function to vary with the level of expenditure. Clearly, if $\gamma < 0$, then γ is not interpreted as the subsistence level. If $\gamma < 0$, then the consumption growth of the poor will be more volatile than that of the rich.

where $C_h^m(t)$ is measured consumption and $\epsilon_h(t)$ is a multiplicative measurement error, which can be serially correlated but is assumed to be independent across households. We assume that $\epsilon_h(t)$ is positive and $\ln(\epsilon_h(t))$ has mean zero. We assume that β_h satisfies

$$\ln(\beta_h) = \beta_0 + \beta_1 y_h^c + \epsilon_h^a, \tag{8.10}$$

where y_h^c is a proxy of permanent income, and ϵ_h^a reflects a measurement error in the proxy of permanent income that is assumed to be independent across households. Then from (8.8) to (8.10), we get Eq. (8.1), where $\phi(t) = (1/\alpha)(lnR^*(t)+\beta_0)$, $b_y = \beta_1/\alpha$, and

$$v_h(t) = \ln(\epsilon_h(t+1)) - \ln(\epsilon_h(t)) + (1/\alpha)\epsilon_h^a. \tag{8.11}$$

4.3 An Incomplete Market Interpretation

It is much more difficult to derive a statistical model of household consumption growth suitable for use in a short panel such as ours from an economic model in which incomplete risk sharing is assumed. We suspect, though, that our empirical results may also be consistent with a model with incomplete markets with borrowing constraints and homothetic preferences. In particular, imagine that agents have a constant relative risk aversion utility function. Then borrowing constraints will have two effects on the consumption growth volatility of households that are close to their borrowing constraints. Each of these effects work in opposite directions. First, households that are close to their borrowing constraints will try to avoid facing borrowing constraints in the future, and thus will be especially concerned about protecting themselves against negative shocks, while households with a lot of liquid assets may act as if they are not affected by the possibility of facing borrowing constraints in the future. Thus poor households may act as if they were more risk averse than rich households, even though they have identical preferences. Second, when a household actually hits its borrowing constraint, then its consumption growth depends more strongly on its current income growth than is the case for an unconstrained household. This effect works in the direction of increasing the consumption volatility of households that are borrowing constrained since these households cannot smooth consumption as much as they might wish. The answer to the question of which effect will dominate in equilibrium depends on many factors. In any case, as long as the first effect dominates, borrowing constraints can make rich households' consumption growth more volatile than that of poor households.

4.4 Additional Test Results

Our statistical model of household consumption growth can be motivated by an economic model with complete markets in which the parameters γ and b_y are preference parameters. This statistical model may also be consistent with the equilibrium of a model with incomplete markets in which agents' IES and RTP are constant across wealth levels. As we will discuss below, it is very difficult to discriminate between these two models with the data sample that we have available, and thus it is beyond the scope of this chapter to try to distinguish these models. Our primary interest in this chapter is to present evidence that casts doubt on Lawrance's Lawrance (1991) hypothesis that it is the RTP and not the IES that rises with the level of wealth.

Some authors, using virtually the same data set as ours, have found evidence against the null hypothesis of complete markets in favor of an alternative model with borrowing constraints and incomplete markets. Morduch (1990) and Bhargava and Ravallion (1993) in particular find statistically significant correlations between consumption and income and wealth variables. Morduch interprets his results as evidence for borrowing constraints and Bhargava and Ravallion (1993) interpret their results as evidence against the permanent income hypothesis. We note that tests such as theirs run on data generated from a complete-markets wealth-varying IES model such as ours may reject the null hypothesis of complete markets because, in our model, consumption growth can be correlated with the level or growth of household income. But, more importantly, their results that household consumption growth may be correlated with household income or income growth may raise some concern about the power of our J tests of the overidentifying restrictions given that we estimate a parameter (γ) that these authors did not estimate.

In order to address the issue of the power of the J test, we report some additional test results in Table 8.5. The first row for each of three consumption measures reports results when (8.1) is estimated with the assumption that full risk sharing is achieved across the three villages. Since there are virtually no direct trades across these villages, it is very unlikely that full risk sharing actually is achieved across the three villages. If the J test has power to detect correlation between consumption and income growth, then the J test should reject the null hypothesis of complete risk sharing. As reported in the table, the J test presents overwhelming evidence against complete risk sharing across the villages for each of three consumption measures. We also report the likelihood ratio type test statistic, C, for the restriction that the consumption over the subsistence level grows at the same rate for all villages. The C test also overwhelmingly rejects this hypothesis.

Even though these results are in favor of the complete markets hypothesis within each village, this hypothesis should not be taken literally. This hypothesis is used with an idea that the hypothesis may be a good enough approximation for consumption behavior to identify preference parameters with our model.

In our economic model with complete markets and wealth-varying IES, the parameter γ is a preference parameter and is assumed to be the same across all

Table 8.5 GMM test results

Error specification	Risk sharing	γ_a [a]	γ_s [a]	γ_k [a]	J [b]	d.f.	C [b]	d.f.
Total consumption								
Multiplicative error	Across villages	176.5 (6.7)	176.5 …	176.5 …	140.69 (0.0)	39	105.99 (0.0)	10
Multiplicative error	Within village	174.5 (20.1)	208.9 (208.6)	272.9 (28.8)	26.90 (46.9)	28	7.79 (2.0)	2
Additive error	Within village	306.5 (44.0)	306.5 …	306.5 …	30.93 (36.9)	29	…	
Additive error	Within village	320.4 (49.6)	414.7 (186.0)	448.2 (122.5)	29.49 (33.7)	27	0.88 (64.3)	2
Food consumption								
Multiplicative error	Across villages	102.3 (0.0)	102.3 …	102.3 …	193.38 (0.0)	39	161.09 (0.0)	10
Multiplicative error	Within village	102.3 (82.4)	101.5 (3.7)	102.6 (78.9)	32.21 (22.4)	27	0.07 (96.5)	2
Additive error	Within village	206.5 (37.2)	206.5 …	206.5 …	29.98 (41.5)	29	…	
Additive error	Within village	199.7 (45.2)	117.6 (232.9)	276.9 (60.8)	29.10 (35.6)	27	0.87 (64.6)	2
Nonfood consumption								
Multiplicative error	Across villages	26.8 (1.4)	26.8 …	26.8 …	181.74 (0.0)	39	153.52 (0.0)	10
Multiplicative error	Within village	26.6 (2.2)	32.5 (15.5)	104.2 (40.9)	20.32 (81.7)	27	7.90 (1.9)	2
Additive error	Within village	64.2 (9.1)	64.2 …	64.2 …	21.75 (83.0)	29	…	
Additive error	Within village	63.7 (9.2)	172.8 (60.2)	41.6 (191.5)	20.40 (81.4)	27	1.35 (50.9)	2

[a] Standard errors are in parentheses
[b] p-values in percentage are in parentheses

three villages. Since this interpretation suggests that this parameter is a structural parameter, tests for the hypothesis that γ is equal across the villages are of interest. In the second row for each consumption measure, we report the test results for this hypothesis when the multiplicative error specification of (8.9) is used. For total consumption and nonfood consumption, we reject this hypothesis at the 5 % level, but not at the 1 % level. Since the p-values reported are approximations based on asymptotic theory, these are not strong evidence against the hypothesis. For food consumption, we do not find any evidence against this hypothesis.

As discussed above, the statistical model in this chapter can be interpreted as a structural model in which consumption is measured with positive multiplicative measurement error as in (8.9). Observe that if the assumption of positive multiplicative measurement error is misspecified, however, it can bias our estimates of

γ downward. Ogaki and Zhang (2001) estimate subsistence levels with an additive measurement error model to allow estimates of γ to be greater than the minimum consumption level observed. For the purpose of comparing the estimates of γ in three villages, the additive error model may be better. For this reason, we also report results when we use the additive error specification. The third row for each consumption measure reports the results when γ is restricted to be the same across the villages. The fourth row reports the results when γ is allowed to be different, and the likelihood ratio-type tests for the hypothesis that γ is the same. We do not reject this hypothesis for any consumption measure.

These estimates for γ from the additive error model are much larger than those from the multiplicative error model, which suggests that the latter estimates of subsistence levels may be biased downward. Rosenzweig and Wolpin (1993) estimate the subsistence level in a single consumption good model by analyzing investments in bullocks in the ICRISAT data. The estimate of the subsistence level from the additive error model for total consumption multiplied by the average family size of 6 is closer to Rosenzweig and Wolpin's estimate of the subsistence level than that from the multiplicative error model. Similarly, the sum of estimates of γ for food and nonfood from the additive error model multiplied by the average family size of 6 is closer to Rosenzweig and Wolpin's estimate of the subsistence level than that from the multiplicative error model.

Thus we find only weak evidence against the hypothesis that γ is the same across villages when the multiplicative error model is used and no evidence against the hypothesis when additive error model is used. Because the additive error model is more reliable for the purpose of estimating the level of the subsistence level, we conclude that this test result also favors the model of wealth-varying IES.

5 Conclusions

In this chapter, we have estimated a model in which the RTP and the IES can rise or fall as a household become richer. Our empirical results are consistent with the view that the IES rises with the level of wealth, while the RTP does not vary with the level of wealth.

Our empirical results can also be interpreted as an atheoretical data description that rich households have more volatile consumption growth than poor households, while the consumption growth mean is constant across rich and poor households. A class of models with borrowing constraints may also be consistent with this data description.

Acknowledgements We thank three anonymous referees the *Review of Economics and Statistics*, James Stock, and seminar participants at many institutions for comments at various stages of this work and R. Anton Braun for his comments that motivated us to write this chapter. We gratefully acknowledge financial support by National Science Foundation grant no. SES-9213930. The additional research of Masao Ogaki for this chapter was partially supported by Grants-in-Aid for Scientific Research from Japan Society for the Promotion of Science.

Appendix

We use food including milk, sweets, and spices as the measure of food consumption. For nonfood consumption, we subtracted food and ceremonial expenses from total consumption expenditure. Ceremonial expenses are removed because they often jump from zero to large amounts. Nonfood consumption consists of narcotics, tea, coffee, tobacco, pan, and alcoholic beverages; clothing, sewing of cloth, other tailoring expenses, thread, needles, chap pals and other footwear etc.; travel and entertainment; medicines, cosmetics soap, barber service; electricity, water charges and cooling fuels for household use; labor expenses for domestic work; edible oils and fats (other than gee); and others, including complete meals in hotels, school and educational materials, stamps, stationery, grinding and milling charges, etc. Unfortunately, the ICRISAT consumption data do not include housing and transportation, because the market values of these categories of consumption are hard to measure in these villages. Total consumption expenditure is the sum of food and nonfood consumption.

To construct real consumption per male adult equivalent, nominal consumption at t is divided by the family size measure constructed by Townsend (1994) and the corresponding price index at t for each village. The price index for total consumption expenditure, food, and nonfood are the consumer price index, the price index for food, and the price index for nonfood, respectively. These real variables are valued at 1983 prices.

There are about forty households for each year in each of the three villages in the data. Some households drop out of the sample and others are added to the sample over years in the ICRISAT data. We exclude these households from our sample. There is one household in the village of Aurepalle with zero income in 1980. Because we take the log of income, this household is excluded. The number of households in our sample for the village of Aurepalle is 35; that for Shirapur, 33; and that for Kanzara, 36.

Addendum: Recent Developments[9]

The main purpose of this chapter's empirical work was to distinguish between the wealth-varying IES and RTP models in Sect. 3. It should be noted that the particular version of the wealth-varying IES model we study in this chapter is not a model of endogenous preferences: the IES varies with wealth because of the subsistence parameter in the model. On the other hand, the wealth-varying RTP model is a model of endogenous preferences. The empirical results of this chapter were more consistent with the wealth-varying IES model than with the wealth-varying RTP model.

[9]This addendum has been newly written for this book chapter.

We think that it is likely that the importance of consuming above the subsistence level dominates intertemporal behaviors for poor people. When a household is near the subsistence level, other considerations such as low or high real interest rates may be of much less importance. This view is consistent with our empirical results. After the publication of our paper, Ogaki and Zhang (2001) used data from Pakistani villages as well as the same data as we used for Indian villages. They found that the subsistence parameter is important in understanding risk sharing behaviors in these Indian and Pakistani villages.[10]

Bhatt and Ogaki (2012), a paper included as Chap. 2 in this book, briefly discussed our empirical evidence. As they argue, it is possible that the RTP is wealth-varying for richer households as we focused on poor households in Indian villages. For example, consumption decisions of richer households in developed countries are not likely to be affected by the subsistence level considerations.

References

Atkeson A, Ogaki M (1996) Wealth-varying intertemporal elasticities of substitution: evidence from panel and aggregate data. J Monet Econ 38:507–534

Bhargava A, Ravallion M (1993) Does household consumption behave as a Martingale? A test for rural South India. Rev Econ Stat 128:500–504

Bhatt V, Ogaki M (2012) Tough love and intergenerational altruism. Int Econ Rev 53:791–814

Chatterjee S (1994) Transitional dynamics and the distribution of wealth in a neoclassical growth model. J Public Econ 54:97–119

Easterly W (1994) Economic stagnation, fixed factors, and policy thresholds. J Monet Econ 33:525–557

Epstein LG (1983) Stationary cardinal utility and optimal growth under uncertainty. J Econ Theory 31:133–152

Gallant AR, White H (1988) A unified theory of estimation and inference for nonlinear models. Basil Blackwell, New York

Hansen LP (1982) Large sample properties of generalized method of moments estimators. Econometrica 50:1029–1054

Jacoby H, Skoufias E (1998) Testing theories of consumption behavior using information on aggregate shocks: income seasonality and rainfall in rural India. Am J Agric Econ 80:1–14

Lawrance EC (1991) Poverty and the rate of time preference: evidence from panel data. J Politi Econ 99:54–77

Lim Y (1992) Disentangling permanent income from risk sharing: a general equilibrium perspective on the rural market in developing countries. Ph.D. dissertation, University of Chicago

Mankiw NG, Zeldes SP (1991) The consumption of stockholders and nonstockholders. J Financ Econ 29:97–112

Morduch J (1990) Risk, production, and saving: theory and evidence from Indian households. Manuscript, Harvard University

Morduch J (1991) Consumption smoothing across space: tests for village-level responses to risk. Manuscript, Harvard University

[10]Zhang and Ogaki (2004) found further evidence for the importance of the subsistence level for risk sharing in the Indian villages.

Ogaki M (1992) Growth in open economies: a comment. Carnegie-Rochester Conf Ser Public Policy 36:47–56

Ogaki M (1993a) Generalized method of moments: econometric applications. In: Maddala GS, Rao CR, Vinod HD (eds) Handbook of statistics, vol 11: econometrics. North-Holland, Amsterdam, pp 455–488

Ogaki M (1993b) GMM: a user's guide. Rochester Center for Economic Research Working Paper No. 348

Ogaki M, Atkeson A (1997) Rate of time preference, intertemporal elasticity of substitution, and level of wealth. Rev Econ Stat 79:564–572

Ogaki M, Zhang Q (2001) Decreasing relative risk aversion and tests of risk sharing. Econometrica 69:515–526

Ogaki M, Ostry JD, Reinhart CM (1996) Saving behavior in low- and middle-income developing countries: a comparison. Int Monet Fund Staff Pap 43:38–71

Ravallion M, Chaudhuri S (1997) Risk and insurance in village India: comment. Manuscri Econometrica 65:171–184

Rebelo S (1992) Growth in open economies. Carnegie-Rochester Conf Ser Public Policy 36:5–46

Rosenzweig MR (1988) Risk, implicit contracts and the family in rural areas of low-income countries. Econ J 98:1148–1170

Rosenzweig MR, Binswanger HP (1993) Wealth, weather risk, and the composition and profitability of agricultural investments. Econ J 103:56–78

Rosenzweig MR, Stark O (1989) Consumption smoothing, migration, and marriage: evidence from rural India. J Politi Econ 97:905–926

Rosenzweig MR, Wolpin KI (1993) Credit market constraints, consumption smoothing and the accumulation of durable production assets in low-income countries. J Politi Econ 101:223–244

Townsend RM (1994) Risk and insurance in village India. Econometrica 62:539–591

Uzawa H (1968) Time preference, the consumption function, and optimum asset holdings. In: Wolfe JN (ed) Value, capital and growth: papers in honour of Sir John Hicks. Edinburgh University Press, Edinburgh

Zhang Q, Ogaki M (2004) Decreasing relative risk aversion, risk sharing, and permanent income hypothesis. J Bus Econ Stat 22:421–430

Chapter 9
Economic Development and Time Preference Schedule: The Case of Japan and East Asian NICs

Kazuo Ogawa

Abstract The constancy of the time preference rate has been a traditional assumption of the literature on intertemporal choice of consumers. This chapter examines empirically the validity of the constancy hypothesis of the time preference schedule. Three alternative hypotheses concerning the time preference rate are investigated. The first hypothesis posits that the time preference rate decreases as the economy develops (Fisher hypothesis). The second hypothesis posits that the time preference rate increases as the economy develops (Uzawa hypothesis). The third hypothesis, which is a combination of the second and third hypotheses, posits that as a nation develops, the time preference rate decreases up to a certain point, and thereafter, the time preference rate increases (Fukao-Hamada hypothesis). Post-war annual time series data from Taiwan, Japan, and Korea are employed to examine these hypotheses. I find that the Fukao-Hamada hypothesis is supported for Taiwan and Japan, while the time preference rate is constant for Korea.

Keywords Time preference schedule • Economic development • Liquidity constraint

1 Introduction

Capital accumulation plays a vital role in economic development. It not only stimulates an economy as a component of final demand in the short run, but also expands the production capacity, leading to an increase of factor productivity. To

---

The original article first appeared in Journal of Development Economics 42: 175–195, 1993. A newly written addendum has been added to this book chapter.

K. Ogawa (✉)
Institute of the Social and Economic Research, Osaka University,
6-1 Mihogaoka, Ibaraki, Japan
e-mail: ogawa@iser.osaka-u.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_9

support capital accumulation, the supply of capital is indispensable. The supply of capital is mainly provided by consumers in the form of savings. The savings pattern of a consumer is determined by an intertemporal choice of consumption.

One celebrated theory of consumption and savings is the life-cycle permanent income hypothesis, abbreviated as LCY-PIH, according to which consumption and saving behavior is determined in such a manner that the time path of consumption is smoothed out. Among other things, time preference rate is an important factor in determining the consumption- savings pattern of the consumer. In a perfect foresight world, the time path of consumption crucially hinges on the size of the time preference rate relative to the real interest rate. When the time preference rate is equal to the real interest rate, the consumption path should be flat. On the other hand, when the time preference rate is smaller (larger) than the real interest rate, the consumption path should tilt toward the future (the present), implying that more consumption will be made in the future (the present). This illustrates the importance of the time preference schedule in capital accumulation. However, although it has been recognized as significant, most studies on consumption and savings in the past have not paid much attention to this factor and have treated it as an exogenous fixed parameter.

This chapter focuses on the role of the long-neglected time preference rate in consumption and savings patterns. Our main goal is to examine the hypothesis that the time preference rate of a nation varies, depending on the stage of economic development. This hypothesis is based upon the naive conjecture that at the early stage of economic development, a nation often exhibits a low saving rate, for too much saving leads to starvation, which might reflect a high time preference rate for the nation. This conjecture is well documented by Fisher (1907). As the economy develops and more capital is accumulated, consumers can afford to plan for the future and more supply of capital will be provided. This suggests that the consumer's time preference rate might tend to decline as the economy develops.

On the contrary, Uzawa (1968) holds the opposite view on the path of the time preference rate. He argues that the time preference rate rises as the consumer's consumption level increases. Fukao and Hamada (1991) combine both Fisher's hypothesis and Uzawa's and assume that as a nation accumulates wealth, the time preference rate declines up to a certain point and thereafter the time preference rate increases. Although these economists lay emphasis on the relationship between the time preference rate and the degree of economic development, no empirical studies have been undertaken on this issue, using the aggregated data.[1]

[1] Hong (1988) is the only exception. He proposes the hypothesis that trade is likely to decrease the time preference rate of a developing country, inducing the developing country to save more in an attempt to catch up with the higher living standards of advanced countries, By cross section study of 42 developing countries he finds evidence that opening to trade has a statistically significant positive effect on the aggregate saving rate. As for microeconometric studies on this issue, Lawrance (1991) estimates the time preference rate of rich and poor consumers in the United States, respectively, based on the Panel Study of Income Dynamics. They are identified from

This chapter tackles this problem empirically. However, in analyzing empirically the dependence of the time preference rate on the degree of economic development based on the aggregated data, attention should be paid to the existence of liquidity-constrained consumers who cannot smooth out their consumption paths. This is especially so for the countries in the early stage of economic development where access to the financial market by consumers is severely restricted. As the economy develops, the financial market becomes more accessible for consumers, which will mitigate the degree of liquidity constraint in the economy.

Therefore, there are two channels through which economic development affects the aggregate consumption-savings pattern of the economy: one is by affecting the time preference rate of LCY-PIH type consumers, and the other is by changing the composition of LCY-PIH type consumers and liquidityconstrained consumers in the economy. This study focuses on the former channel, but the latter is also taken into consideration by testing the hypothesis that the proportion of liquidity-constrained consumers is time-varying.

Three countries are chosen as our sample: Japan, Taiwan, and Korea, the latter two of which are typical newly industrializing countries (NICs) in East Asia. The countries covered by our study are well-known for their high personal saving rates as well as high economic growth in the post-war period. Using post-war annual time series data on consumption in these countries, we test the constant time preference hypothesis. When the constant hypothesis is not supported by the data, we further investigate statistically whether the time preference schedule is well represented by Fisher's hypothesis or by Uzawa's or by Fukao and Hamada's.

The summary of our findings is as follows: the constant time preference hypothesis is valid only for Korea. For Japan the Fukao-Hamada hypothesis is appropriate. For Taiwan the empirical result is not yet decisive as to which hypothesis is most suitable. The post-war data of Taiwan is consistent with the Fukao-Hamada hypothesis as well as the Fisher's hypothesis with the proportion of liquidity-constrained consumers declining as the economy develops. Section 2 presents the basic model of consumption utilized in estimation. Section 3 reports estimation results and interprets them. Section 4 is a concluding remark.

2 Basic Model of Consumption

Our formulation of consumers' behavior is based on the assumption that the economy consists of two types of consumers: those following the LCY-PIH and those who are liquidity-constrained. The first type of consumer determines their consumption-saving plan on the basis of current nonhuman wealth and the present discounted value of future expected labor earnings without any constraints in the

---

estimation of Euler equation of LCY-PIH type consumers. It is reported that the time preference rate of poor consumers are three to five percentage points higher than those of rich consumers.

capital market. The consumption level of the second type of consumers is totally restricted by current disposable income, because they neither hold sufficient liquid assets nor are able to borrow against expected future labor earnings to sustain the consumption plan determined from lifetime resources.

The consumption levels of these two types of households are combined to form the aggregate consumption.[2] This dual specification of consumption behavior is necessary for two reasons. First, in the early post-war period capital markets were not well developed in Japan and East Asian NICs. This prevented consumers from allocating lifetime resources optimally over time and bound their consumption behavior by the conditions of liquidity constraints. Second, without incorporating the liquidity-constrained consumers into the aggregate consumption, it is almost impossible to interpret the coefficient of disposable income. It can be interpreted either as evidence for the existence of liquidity-constrained consumers or as evidence in favor of the dependence of the time preference rate on income level.

Formally, the consumption-savings pattern of the LCY-PIH type consumers is characterized by the following celebrated Euler equation, originally derived by Hall (1978):

$$U'\left(C_{it-1}\right) = \mathrm{E}\left[\left(\left(1 + r_t\right) / \left(1 + \delta_{it-1}\right)\right) U'\left(C_{it}\right) \middle| \Omega_{t-1}\right], \tag{9.1}$$

where

$U(\cdot)$ is a utility function with $U'(\cdot) > 0$ and $U''(\cdot) < 0$,
C_{it} is the consumption level of LCY-PIH type consumer i at time t,
r_t is the real interest rate at time t,
δ_{it-1} is the time preference rate of LCY-PIH type consumer i at time $t-1$,
$\mathrm{E}\left[\cdot \middle| \Omega_{t-1}\right]$ is the mathematical expectation operator conditioned on the information set Ω_{t-1} available to the consumers at time $t-1$.

In (9.1) we take account of the possibility that the time preference rate might vary over time. It is assumed that the factor driving the time preference rate is exogenous to the consumers.[3]

[2]The idea that the economy consists of two types of consumers: LCY-PIH type consumers and liquidity-constrained consumers was first utilized by Hayashi (1982) in specifying an aggregate consumption function. Campbell and Mankiw (1989, 1990) and Ogawa (1990) are on the same line.

[3]In Uzawa's original formulation the time preference schedule depends positively on the current level of utility. In other words, a higher level of consumption increases the time preference rate. Therefore, the intertemporal Euler equation should be derived, taking the endogeneity of the time preference rate into consideration. Our formulation is different from Uzawa's in that what affects the time preference rate is taken as exogenous to the consumers. However, Uzawa's spirit is incorporated in our model since the time preference rate is specified as an increasing function of income level or the level of living standard.

In order to make (9.1) operational, we assume that the utility function of the LCY-PIH type consumers is represented by the constant absolute risk aversion utility function[4]:

$$U\left(C_{it}\right) = \left(-1/\alpha\right) \mathrm{e}^{-\alpha C_{it}}, \tag{9.2}$$

where α is the degree of absolute risk aversion. We assume that the degree of absolute risk aversion is common to all the LCY-PIH consumers. Then (9.1) can be written as

$$\left(\left(1 + r_t\right) / \left(1 + \delta_{it-1}\right)\right) \mathrm{e}^{-\alpha \Delta C_{it}} = 1 + \varepsilon_{it}, \tag{9.3}$$

where

$$\varepsilon_{it} \equiv \left(\left(1 + r_t\right) / \left(1 + \delta_{it-1}\right)\right) \mathrm{e}^{-\alpha \Delta C_{it}}$$
$$-\mathrm{E}\left[\left(\left(1 + r_t\right) / \left(1 + \delta_{it-1}\right)\right) \mathrm{e}^{-\alpha \Delta C_{it}} \middle| \Omega_{t-1}\right].$$

The forecast error ε_{it} has mean zero and is not correlated with any elements in the information set Ω_{t-1}.

Taking the logarithm of the both sides of (9.3) and rearranging the terms, we obtain

$$\Delta C_{it} = (1/\alpha) \left\{\sigma_{i\varepsilon}^2 /2 - \ln\left(1 + \delta_{it-1}\right)\right\} + (1/\alpha) \ln\left(1 + r_t\right) + \upsilon_{it}, \tag{9.4}$$

where

$$\upsilon_{it} = (1/\alpha) \left\{\varepsilon_{it}^2 /2 - \sigma_{i\varepsilon}^2/2 - \varepsilon_{it}\right), \ \mathrm{E}\left[\upsilon_{it} \middle| \Omega_{t-1}\right] = 0.$$

As for the liquidity-constrained consumers, the consumption level is totally dependent on their current disposable income:

$$C_{it} = \beta(YD)_{it}. \tag{9.5}$$

[4]It is well known that the constant absolute risk aversion utility function has some drawbacks including the decrease in the proportion of wealth invested in risky assets as the wealth level increases. Nonetheless, we adopted this type of utility function because under this utility function exact aggregation of consumers to form aggregate consumption is guaranteed. Campbell and Mankiw (1989, 1990) assume an isoelastic utility function, which has more appealing implications on the portfolio choice, but the exact aggregation is not possible. In a later section we will present estimation results of a consumption function based on the iso-elastic utility function and discuss the robustness of our findings that the time preference schedule is dependent on the real income level.

Expressing (9.5) in terms of difference form, we obtain

$$\Delta C_{it} = \beta \Delta (YD)_{it}. \tag{9.6}$$

Summing (9.4) and (9.6) and dividing them by the total population yields the total consumption change (ΔC_i) per capita as[5]

$$\begin{aligned}\Delta C_t = \gamma - ((1-\lambda_1)/\alpha) \ln (1+\delta_{t-1}) + ((1-\lambda_1)/\alpha) \ln (1+r_t) \\ + \beta\lambda_2 \Delta (YD)_t + u_1,\end{aligned} \tag{9.7}$$

where

C_t is aggregate consumption per capita,
YD_t is aggregate disposable income per capita,
δ_{t-1} is the geometric mean of the individual time preference rate,
λ_1 is the proportion of liquidity-constrained consumers in the total population $(= N_2/N)$,
N_2 is the number of liquidity-constrained consumers,
N is the number of total consumers,
λ_2 is the proportion of the disposable income earned by the liquidityconstrained consumers,
$\gamma = (1/\alpha N) \sum_{i \in \mathrm{I}} \left(\sigma_{i\varepsilon}^2/2\right)$ (type I consumers refer to unconstrained consumers).

$$u_t = (1/N) \sum_{i \in \mathrm{I}} \upsilon_{it},$$

When the constant absolute risk-averse utility function characterizes the preference of the LCY-PIH consumers, the *arithmetic* consumption change of the LCY-PIH consumers is expressed as a function of the time preference rate and the real interest rate. On the other hand, the *arithmetic* consumption changes of the liquidity-constrained consumers is proportional to the arithmetic change of their disposable income. Thus exact aggregation can be attained to express aggregate consumption change as a function of the time preference rate, the real interest rate, and the disposable income. This consumption function is similar to that specified

[5]In deriving (9.7) we implicitly assume that the proportion of the disposable income earned by the liquidity-constrained consumers (λ_2) is invariant over time. This assumption appears stringent, since the extent of the liquidity constraint in developing countries may change over time, depending on the development stage of their financial system. In a later section we come back to this point and examine whether relaxation of the time-invariant assumption on the degree of liquidity constraint might improve the estimation of the consumption function.

by Campbell and Mankiw (1989, 1990) which incorporates not only unconstrained consumers, but also liquidity-constrained consumers. The difference lies in that their consumption and income variables are all expressed in terms of logarithmic differences, while ours are all in terms of arithmetic differences.

Now we proceed to specify the time preference rate. The simplest case is to assume that the time preference rate is constant, independent of time. We call this case Type I. In this case the time preference rate is included in the constant term in (9.7). The other cases assume that the time preference rate depends on the stage of economic development. We measure the degree of economic development by the real labor income per capita (Y). Inclusion of this term is justified when the labor supply decision of the consumer is independent of the consumption-savings decision, such as the case where the utility function of leisure and consumption are additively separable, or the labor supply is exogenously given.

We formulate the dependence of the time preference rate on the real labor income in such a manner that the hypotheses on the evolution of the time preference rate stated above can be incorporated. The following specification serves to embody the hypotheses made by Fisher, Uzawa or Fukao-Hamada:

$$\text{In}\,(1 + \delta_{t-1}) = \rho_1 + \rho_2 Y_{t-1} + \rho_3 Y_{t-1}^2. \tag{9.8}$$

When $\rho_2 < 0$ and $\rho_3 = 0$, then Fisher's hypothesis holds, while in the case of $\rho_2 > 0$ and $\rho_3 = 0$ Uzawa's hypothesis holds. We call this case Type II. The Fukao-Hamada hypothesis is held when $\rho_2 < 0$ and $\rho_3 > 0$. The turning point is attained when the real labor income is equal to $-(\rho_2/2\rho_3)$. The case is called Type III.

Substitution of (9.8) into (9.7) yields the following consumption function to be estimated:

$$\begin{aligned}\Delta C_t = {} & \gamma - ((1-\lambda_1)/\alpha)\,\rho_1 - ((1-\lambda_1)/\alpha)\,\rho_2 Y_{t-1} - ((1-\lambda_1)/\alpha) \\ & \times \rho_3 Y_{t-1}^2 + ((1-\lambda_1)/\alpha)\ln(1+r_t) + \beta\lambda_2 \Delta(YD)_t + u_t\end{aligned} \tag{9.9}$$

3 Empirical Evidence

In this section we investigate empirically the hypotheses concerning the time preference rate by estimating the consumption functions derived in the previous section. We use the post-war annual time series data of three countries: Taiwan, Korea, and Japan. As for the sample periods for estimation, they are 1952–1985 for Taiwan, 1956–1987 for Korea, and 1955–1986 for Japan. These countries are similar in two aspects. First, they have all attained rapid economic growth in the post-war period. For the period of 1956–1986, the average annual GDP growth rate was 6.99 % for Japan, 8.69 % for Taiwan, and 7.80 % for Korea. Second, these countries have high personal savings rates. The average personal saving rate during the sample period is 13.77 % for Taiwan, 8.59 % for Korea, and 16.88 % for Japan.[6]

Some explanations of the data used for estimation are now in order. Detailed description of the data is given in the appendix. The consumption data we need is ideally that excluding expenditures on durable goods. Therefore, we constructed a consumption series that excluded expenditure on items classified as durables. The consumption and income variables are measured on a per capita basis. As for interest rates, we use the interest rate on 3-month time deposits for Taiwan, that on time deposits of 1 year or more for Korea, and that on 1-year time deposits for Japan.

3.1 Econometric Methodology

The disturbance terms in the consumption functions derived in the previous section might exhibit heteroskedasticity, which is caused by the exponential trend of the arithmetic difference of consumption variables. We assume that the variance of the error terms is proportional to the squared current disposable income. Hence all the variables in the consumption functions were divided by the current disposable income in estimation to secure homoskedasticity. Furthermore, we compute the heteroskedasticity consistent estimates of the standard errors of the regression coefficients suggested by White (1980) so that we may draw proper inferences even when heteroskedasticity cannot be completely wiped out.

We applied the instrumental variables method to estimate the consumption functions. The reason lies in the error term u_t which is not correlated with any variables at time $t-1$, but correlated with the real interest rates at time $t(r_t)$ and the difference variables $(\Delta(YD)_t)$. In general any variables prior to time t included in the information set Ω_{t-1} can be valid instruments. However, we did not use the variables dated as of time $t-1$ as instruments to avoid the problems caused by delays in the information available to consumers, the partly durable nature of the goods labelled non-durables, and the white noise errors remaining in the levels of consumption and income variables.[7]

3.2 Estimation Results

We will first present the results of the traditional Type I consumption function, where the time preference rate is invariant over time. We estimated the consumption functions for two cases. One case assumes that the real interest rates are perceived

[6] The average personal saving rate of Korea jumps to 11.80 % if the period 1956–1965 is excluded.

[7] See Campbell and Mankiw (1990) for a more detailed discussion why variables lagged more than one period are desirable instruments.

Table 9.1 Estimation of consumption function with constant time preference rate (Taiwan)

Instruments list		Constant	ΔYD_t	$\ln(1+r_t)$	$R^2/D.W.$
1.	#1	1.0414 (1.41)	0.5578 (9.81)		0.5557 1.9197
2.	#2	0.6916 (0.97)	0.5869 (9.95)		0.5636 1.9555
3.	#3	0.8849 (1.19)	0.5709 (9.54)		0.5595 1.9387
4.	#4	0.6596 (0.97)	0.5895 (9.75)		0.5643 1.9577
5.	#5	4.3124 (2.41)	0.4366 (7.41)	−58.410 (−1.79)	0.2389 1.9510
6.	#6	2.5566 (1.93)	0.5153 (10.17)	−29.674 (−1.09)	0.4500 2.0771
7.	#7	2.2532 (1.95)	0.5179 (12.37)	−18.448 (−0.87)	0.4813 2.0309
8.	#8	1.5286 (1.72)	0.5466 (12.85)	−4.7639 (−0.26)	0.5078 1.9789

R^2 is an adjusted coefficient of determination. *D. W.* is Durbin-Watson statistics. The values in *parentheses* are asymptotic *t*-values

Table 9.2 Estimation of consumption function with constant time preference rate (Korea)

Instruments list		Constant	ΔYD_t	$\ln(1+r_t)$	$R^2/D.W.$
1.	#1	0.0591 (1.28)	0.3293 (2.88)		0.4203 2.0436
2.	#2	0.0629 (1.42)	0.3086 (2.84)		0.4158 1.9394
3.	#3	0.0464 (0.95)	0.3786 (3.50)		0.4297 2.2600
4.	#4	0.0546 (1.23)	0.3406 (3.44)		0.4237 2.1005
5.	#5	0.1043 (2.48)	0.2043 (1.75)	0.8440 (2.11)	0.2972 2.2068
6.	#6	0.1183 (2.61)	0.1540 (1.39)	1.0412 (2.05)	0.2478 2.0414
7.	#7	0.1094 (2.39)	0.1849 (1.55)	0.9511 (2.00)	0.2798 2.1377
8.	#8	0.1171 (2.52)	0.1573 (1.39)	1.0691 (2.06)	0.2497 2.0458

See the note to Table 9.1 for notations

to be constant by consumers. Then the interest rates are subsumed into the constant terms. The other case assumes that the interest rates are perceived to vary over time. The estimated results are shown in Table 9.1 for Taiwan, Table 9.2 for Korea, and Table 9.3 for Japan. The first column gives information on the instrumental variables employed. The list of instrumental variables is summarized in Table 9.4.

Table 9.3 Estimation of consumption function with constant time preference rate (Japan)

Instruments list		Constant	ΔYD_t	$\ln(1+r_t)$	$R^2/D.W.$
1.	#1	0.1669 (2.53)	0.3167 (2.08)		0.7655 1.8972
2.	#2	0.1459 (2.49)	0.3622 (2.64)		0.7783 1.8762
3.	#3	0.1665 (2.42)	0.3176 (2.01)		0.7658 1.8970
4.	#4	0.1369 (2.29)	0.3817 (2.73)		0.7823 1.8616
5.	#5	0.0626 (1.32)	0.5534 (5.96)	1.7199 (2.12)	0.8525 1.7917
6.	#6	0.0642 (1.49)	0.5502 (6.47)	1.7759 (2.41)	0.8541 1.7943
7.	#7	0.0674 (1.49)	0.5432 (6.14)	1.7125 (2.40)	0.8540 1.8037
8.	#8	0.0761 (1.94)	0.5253 (6.71)	1.7704 (2.57)	0.8577 1.8206

See the note to Table 9.1 for notations

The coefficient of the interest rate variables measures the degree of intertemporal substitution.[8] The coefficients are statistically significant at the 5 % level for Japan and Korea and satisfy the sign condition irrespective of the instruments to be chosen. For Taiwan it picks up a wrong sign. Poor estimates on the intertemporal substitution of consumption might result from the fact that the capital markets accessible for consumers are not sufficiently sophisticated in Taiwan or from the inadequacy of interest rate variables as a true signal employed by consumers to allocate lifetime resources over time. Our findings partially support the evidence documented by Giovannini (1985) which is unfavorable to the intertemporal substitution theory of consumption for the less developed countries. However, it should be noted that coexistence of the LCY-PIH consumers and liquidity-constrained consumers in the economy is not taken into consideration in his study.

The disposable income variable turns out to be significant for all the countries examined, indicating that some fraction of consumers are liquidity-constrained. In Taiwan the proportion of consumers facing liquidity constraints measured by the disposable income is 56–59 % for the constant interest rate case.

In Korea it is 31–38 % for the constant interest rate case, while it is reduced to 15–20 % for the variable interest rate case. However, it should be noted that the

[8]The readers should be careful enough to interpret the estimated coefficient of the interest rate, degree of intertemporal substitution, since the interest rates chosen here for Taiwan and Korea might be held below their market equilibrium levels due to repressed financial regime and thus do not function properly as a signal to allocate consumer's resources intertemporally. It was only after 1959 that the data on the interest rate of one-year time deposits for Taiwan became available, which made us adopt the interest rate on three-month time deposits. However, this choice will affect estimation results little since two series of the interest rates exhibit a parallel movement.

Table 9.4 List of instrumental variables

Instruments	
#1	Constant, $\Delta YD_{t-2}/YD_{t-2},\ \Delta YD_{t-3}/YD_{t-3}, 1/YD_{t-2},\ 1/YD_{t-3}$
#2	$\#1 + \Delta C_{t-2}/YD_{t-2},\ \Delta C_{t-3}/YD_{t-3}$
#3	$\#1 + \Delta i_{t-2}/YD_{t-2},\ \Delta i_{t-3}/YD_{t-3}$
#4	$\#1 + \Delta C_{t-2}/YD_{t-2},\ \Delta C_{t-3}/YD_{t-3},\ \ \Delta i_{t-2}/YD_{t-2},\ \Delta i_{t-3}/YD_{t-3}$
#5	$\#1 + \ln(1+r_{t-2})/YD_{t-2},\ \ln(1+r_{t-3})/YD_{t-3}, \ln(1+r_{t-4})/YD_{t-4}$
#6	$\#5 + \Delta C_{t-2}/YD_{t-2},\ \Delta C_{t-3}/YD_{t-3}$
#7	$\#5 + \Delta i_{t-2}/YD_{t-2},\ \Delta i_{t-3}/YD_{t-3}$
#8	$\#5 + \Delta C_{t-2}/YD_{t-2},\ \Delta C_{t-3}/YD_{t-3},\ \ \Delta i_{t-2}/YD_{t-2},\ \Delta i_{t-3}/YD_{t-3}$
#9	$\#1 + Y_{t-3}/YD_{t-2}$
#10	$\#9 + \Delta C_{t-2}/YD_{t-2},\ \Delta C_{t-3}/YD_{t-3}$
#11	$\#9 + \Delta i_{t-2}/YD_{t-2},\ \Delta i_{t-3}/YD_{t-3}$
#12	$\#9 + \Delta C_{t-2}/YD_{t-2},\ \Delta C_{t-3}/YD_{t-3},\ \ \Delta i_{t-2}/YD_{t-2},\ \Delta i_{t-3}/YD_{t-3}$
#13	$\#9 + Y_{t-3}^2/YD_{t-2},\ Y_{t-4}^2/YD_{t-3}$
#14	$\#13 + \Delta C_{t-2}/YD_{t-2},\ \Delta C_{t-3}/YD_{t-3}$
#15	$\#13 + \Delta i_{t-2}/YD_{t-2},\ \Delta i_{t-3}/YD_{t-3}$
#16	$\#13 + \Delta C_{t-2}/YD_{t-2},\ \Delta C_{t-3}/YD_{t-3},\ \ \Delta i_{t-2}/YD_{t-2},\ \Delta i_{t-3}/YD_{t-3}$
#17	$\#13 + \ln(1+r_{t-2})/YD_{t-2},\ \ln(1+r_{t-3})/YD_{t-3}, \ln(1+r_{t-4})/YD_{t-4}$
#18	$\#17 + \Delta C_{t-2}/YD_{t-2},\ \Delta C_{t-3}/YD_{t-3}$
#19	$\#17 + \Delta i_{t-2}/YD_{t-2},\ \Delta i_{t-3}/YD_{t-3}$
#20	$\#17 + \Delta C_{t-2}/YD_{t-2},\ \Delta C_{t-3}/YD_{t-3},\ \ \Delta i_{t-2}/YD_{t-2},\ \Delta i_{t-3}/YD_{t-3}$
#21	Constant, $\Delta\ln(YD_{t-2}),\ \ \Delta\ln(YD_{t-3}),\ Y_{t-3},\ YD_{t-2},\ YD_{t-3},\ Y_{t-3}^2,\ Y_{t-4}^2$
#22	$\#21 + \Delta\ln(C_{t-2}),\ \Delta\ln(C_{t-3})$
#23	$\#21 + \Delta\ln(i_{t-2}),\ \Delta\ln(i_{t-3})$
#24	$\#21 + \Delta\ln(C_{t-2}),\ \Delta\ln(C_{t-3}),\ \ \Delta\ln(i_{t-2}),\ \Delta\ln(i_{t-3})$
#25	$\#21 + \ln(1+r_{t-2}),\ \ln(1+r_{t-3}),\ \ln(1+r_{t-4})$
#26	$\#25 + \Delta\ln(C_{t-2}),\ \Delta\ln(C_{t-3})$
#27	$\#25 + \Delta\ln(i_{t-2}),\ \Delta\ln(i_{t-3})$
#28	$\#25 + \Delta\ln(C_{t-2}),\ \Delta\ln(C_{t-3}),\ \Delta\ln(i_{t-2}),\ \Delta\ln(i_{t-3})$

explanatory power of the equation is reduced to a large extent once the interest rate variable is added as an explanatory variable. Our estimates of the liquidity-constrained households for the variable interest rate case are comparable with the one estimated by Haque and Montiel (1989), which is 18.2 %.

In Japan the proportion is 32–38 % for the constant interest rate case, but it rises to 53–55 % for the more credible case where the real interest rate varies over time. Our estimates for the variable interest rate case are very close to the one obtained by Campbell and Mankiw (1989) where a quarterly data series for the period of 1959–1986 is employed.[9] Their estimate is 55.3 %. What is common to every country is

[9]Campbell and Mankiw (1989) use total consumption spending as a consumption variable and GDP as a proxy of disposable income.

Table 9.5 Estimation of consumption function with the preference rate dependent quadratically on real labor income (Taiwan)

Instrument list		Constant	ΔYD_t	Y_{t-1}	Y_{t-1}^2	$R^2/D.W.$
1.	#13	−4.1276	0.2626	0.0883	−0.0001	0.5806
		(−2.05)	(2.27)	(3.08)	(−2.62)	1.8321
2.	#14	−3.6049	0.3367	0.0745	−0.0001	0.5927
		(−1.81)	(3.04)	(2.54)	(−2.23)	1.9264
3.	#15	−4.0175	0.2818	0.0850	−0.0001	0.5861
		(−1.98)	(2.62)	(3.02)	(−2.57)	1.8578
4.	#16	−3.4944	0.3511	0.0717	−0.0001	0.5928
		(−1.70)	(3.47)	(2.46)	(−2.14)	1.9422

See the note to Table 9.1 for notations

that the non-negligible proportion of consumers, varying from 15 % to 59 %, are trapped by liquidity constraints.

Now we examine the results of estimating the consumption functions where the time preference rate depends on the level of real labor income. For Taiwan, it turns out that the effects of labor income on the time preference rate are significant only in the quadratic specification without the real interest rates variable. The estimated results are shown in Table 9.5. The linear term of real labor income is positive, and the squared term is negative. It implies that at the early stage of development the time preference rate declines as the income level increases. but beyond a certain point it starts to rise with the income level. This conforms to the Fukao-Hamada hypothesis on the time preference schedule. When the real interest rate is included as an explanatory variable, it either picks up a wrong sign or the coefficient is statistically insignificant even if it satisfies the sign condition. Therefore, it seems to me that the quadratic specification without the real interest rate characterizes the consumption pattern of Taiwan most suitable. The proportion of liquidity-constrained consumers ranges from 26 % to 35 %, which is much smaller than those obtained under the assumption that the time preference rate is constant.

For Korea we fail to find evidence that the time preference rate hinges on the real labor income irrespective of the specification of the time preference rate. None of the coefficients on the real labor income and the squared real labor income are statistically significant. Moreover, the standard errors of the estimates on $\Delta(YD)_t$ increases to a large extent, which makes most of the parameter estimates quite unstable. To sum up, judging from the explanatory power of the equations, we may conclude that for Korea the most suitable specification of the consumption pattern is the one with the constant time preference schedule and the constant real interest rate.

For Japan the dependence of the time preference rate on the real labor income cannot be detected if the real interest rate is not included as an explanatory variable. However, once the real interest rate is added to the explanatory variables list, a nonlinear time preference schedule emerges with statistical significance. The estimated results are summarized in the upper part of Table 9.6. The linear term of real labor income is positive, and the squared term is negative, which is similar to the Taiwanese case. The FakaoHamada hypothesis on the time preference

Table 9.6 Estimation of consumption function with the time preference rate dependent quadratically on real labor income (Japan)

Instrument list		Constant	ΔYD_t	Y_{t-1}	Y_{t-1}^2	$\ln(1+r_t)$	$R^2/D.W.$
One-year time deposits rate							
1.	#17	−0.1647 (−1.29)	0.2345 (2.68)	0.1084 (3.97)	−0.0079 (−4.03)	3.1134 (5.03)	0.8937 2.2110
2.	#18	−0.0654 (−1.40)	0.2553 (2.91)	0.1046 (3.86)	−0.0076 (−3.91)	3.1378 (5.54)	0.8946 2.1674
3.	#19	−0.0545 (−1.11)	0.2455 (2.85)	0.1023 (3.60)	−0.0074 (−3.64)	2.9525 (4.97)	0.8949 2.2590
4.	#20	−0.0563 (−1.18)	0.2642 (3.14)	0.0993 (3.61)	−0.0072 (−3.63)	2.9901 (5.35)	0.8956 2.2131
Bank debenture rate							
5.	#17	−0.1851 (−2.56)	0.5200 (4.64)	0.0560 (1.77)	−0.0033 (−1.40)	2.7956 (4.36)	0.8623 1.6604
6.	#18	−0.1757 (−2.69)	0.5302 (4.62)	0.0515 (1.55)	−0.0030 (−1.20)	2.7169 (4.87)	0.8600 1.6655
7.	#19	−0.1684 (−2.10)	0.3903 (5.94)	0.0779 (2.80)	−0.0050 (−2.61)	2.5990 (4.17)	0.8752 1.8402
8.	#20	−0.1817 (−2.49)	0.4078 (6.41)	0.0778 (3.02)	−0.0050 (−2.78)	2.7255 (4.81)	0.8754 1.7875

See the note to Table 9.1 for notations

schedule is supported for Japan as well as for Taiwan. The estimated results using the instrumental variables list #17 to #20 in Table 9.6 are the most suitable characterization of the Japanese consumers' behavior. The proportion of liquidity-constrained consumers is 23–26 %. As is the case with Taiwan, the proportion is overestimated without the nonlinear dependence of the time preference schedule on the income variable. The consumption response to the real interest rate is almost twice as large as in the case of constant time preference rate.

To summarize the estimated results, we detect a nonlinear dependence of time preference rate on real income for Taiwan and Japan, giving empirical support to the Fukao-Hamada hypothesis on the evolution of the time preference schedule. For Korea no evidence has been obtained about the time-dependence hypothesis on the time preference rate. Our final estimates on the proportion of liquidity-constrained consumers are 26–35 % for Taiwan, 31–38 % for Korea, and 23–26 % for Japan. These values are of reasonable magnitude judging from the different stage of development in the capital market of each country.

3.3 Nonlinear Time Preference Schedule and Turning Point

For Taiwan and Japan, we obtained evidence favorable to the Fukao Hamada hypothesis on the time preference schedule. The linear term of real labor income in the consumption function turns out to be positive, while the squared term is negative for both countries. It guarantees that the real income level producing the turning point in the time preference schedule is positive. We can now pinpoint the dates when the turning point occurred for Taiwan and Japan, and give some interpretations to them.

Let the coefficients of Y_{t-1} and Y_{t-1}^2 be θ_1 and θ_2. Then the level of real labor income yielding the turning point of the time preference schedule is given by $-(\theta_1/2\theta_2)$. Based on the estimated results in Table 9.5 (instrumental variables list #13 to #16) for Taiwan and those in Table 9.6 (instrumental variables list #17 to #20) for Japan, we can calculate the levels of real labor income and obtain the years when the turning points occurred. For Taiwan, the turning point is uniquely estimated to have occurred between 1977 and 1978. As for Japan, the turning point is also consistently calculated as between 1969 and 1970, irrespective of the choice of the instrumental variables.

The turning points coincide with the points at which the growth pattern of real labor income changes. For Taiwan, the growth rate exceeded 5 % in the 1960s and 1970s except 1974 and 1975, while it decreased significantly to 3–4 % in the 1980s, except in 1984. The average growth rate between 1952 and 1979 was 7.33 % and it declined to 4.75 % during 1979–1985. The growth pattern of real labor income in Japan changed dramatically. The growth rate hovered around 10 % before the first oil crisis. Thereafter, it decreased substantially to the level of no more than 4 %. The average growth rate was 9.67 % and 3.84 % for the period of 1955–1970 and 1970–1986, respectively. Thus our evidence indicates that the turning point of the time preference schedule emerges around the periods when the high growth pattern of income is replaced by the low growth pattern of income.

Finally, we will examine whether the turning points of the time preference rate bring forth any changes in savings pattern. The saving rates show an increasing trend up to the late 1970s in Taiwan and it stays around 20 % in the 1980s. For Japan, the saving rate also shows an increasing trend up to the early 1970s, but it switches to a decreasing trend thereafter.[10] For Taiwan and Japan, the periods of increasing saving rates correspond to the phase of decreasing time preference and a constant or decreasing pattern of saving rates emerges once the time preference rate starts to rise.

[10]The saving rates are notably high in Japan from 1974 to 1976. These high saving rates are interpreted as the consumers' provision against the uncertainty of real income in the future. Ogawa (1991) estimates the proportion of precautionary savings out of total savings amounts to 2–4 %. If these precautionary parts are subtracted from the total personal saving rates, then the hump of the saving rates during this period will disappear.

3.4 Robustness of the Results

We now examine the robustness of our findings obtained so far from various angles. In particular we are interested in seeing whether our main findings of nonlinear time preference schedule in Taiwan and Japan emerge if we pick up different variables in estimation or assume a different set of assumptions underlying the model.

Firstly, we use the bank debenture rate instead of the 1-year time deposits rate as an interest rate variable in estimating the consumption function of Japan. This enables us to test whether the estimates are sensitive to the choice of the interest rate variables. The estimated results are shown in the lower part of Table 9.6. We still detect the quadratic time preference schedule and the interest rate exerts a positive effect on ΔC_t significantly in that specification. The turning point is estimated to have occurred between 1971 and 1972 for the instrumental list #19 and #20 or between 1972 and 1973 for the instrumental list #17 and #18. These estimates of the turning point are close to those derived above. The estimates of the proportion of liquidity-constrained consumers are 39–53 %, which is higher than those obtained when the 1-year time deposits rate is employed as an interest rate variable. Findings of higher proportion of liquidity-constrained consumers in the case of the bank debenture rate may reflect the characteristics of the portfolio choice by Japanese consumers. Most of the household financial wealth is held in the form of time deposits.[11] This implies that while most of the consumers respond to a change in the 1-year time deposits rate, few respond to that in the bank debenture rate. Therefore, when the bank debenture rate is employed in the regression, most consumers look as if they were liquidity-constrained.

Secondly, we can modify our model in such a manner that the proportion of liquidity-constrained consumers changes over time. It might be argued that the extent of liquidity constraint in a developing country varies over time as its financial system goes through changes. As the capital market becomes more accessible for consumers, the proportion of liquidity-constrained consumers will become smaller. In general the higher the real income level of a country, the more advanced its financial system is and therefore the smaller the proportion of liquidity constraint is. This idea can be embodied in our model in the following way. As is discussed above, a change of consumption level for the liquidity-constrained consumers is expressed as

$$\begin{aligned}\Delta C_{it} &= \beta \Delta (YD)_{it} \\ &= \beta \left(YD_t \left(YD_{it}/YD_t\right) - YD_{t-1}\left(YD_{it-1}/YD_{t-1}\right)\right).\end{aligned} \tag{9.6a}$$

The term (YD_{it}/YD_t), denoted by λ_{2t} is the proportion of the disposable income earned by the liquidity-constrained consumers. We assume that the parameter λ_{2t} is

[11] In 1955, 58.4 % of the net household financial assets were held in the form of time deposits, and the proportion of time deposits out of net financial assets was 67.0 % in 1986.

a function of real disposable income[12]:

$$\lambda_{2t} = \theta_0 + \theta_1 YD_t,\ \theta_1 < 0. \tag{9.10}$$

Then (9.6a) can be simplified as

$$\Delta C_{it} = \beta\left(\theta_0 + \theta_1\left(YD_t + YD_{t-1}\right)\right)\Delta(YD)_t. \tag{9.11}$$

Combining (9.11) with (9.4) yields the total consumption change as

$$\begin{aligned}\Delta C_t = \gamma &- \left(\left(1-\lambda_1\right)/\alpha\right)\rho_1 - \left(\left(1-\lambda_1\right)/\alpha\right)\rho_2 Y_{t-1} - \left(\left(1-\lambda_1\right)/\alpha\right)\\ &\times \rho_3 Y_{t-1}^2 + \left(\left(1-\lambda_1\right)/\alpha\right)\ln\left(1+r_t\right) + \beta\theta_0\Delta(YD)_t\\ &+ \beta\theta_1\left(YD_t + YD_{t-1}\right)\Delta(YD)_t + u_t.\end{aligned} \tag{9.12}$$

Equation 9.12 is estimated for Taiwan, Korea, and Japan. For Taiwan, once the time-variant feature of liquidity constraint is allowed for in estimation, all the coefficients of the income variables except $\Delta(YD)_t$ lose significance. However, when (9.12) is estimated without the squared term all the coefficients turn out to be significant at the 5 % level. The estimated results without the squared term are shown in the upper part of Table 9.7. The coefficient of the linear labor income is positive, which supports the Fisher hypothesis on the time preference schedule. The coefficient of $\left(YD_t + YD_{t-1}\right)\Delta(YD)_t$ is negative, as is expected. It implies that the proportion of disposable income earned by liquidity-constrained consumers becomes smaller as the consumer's income level rises. In fact, if the marginal propensity to consume for liquidity-constrained consumers is assumed to be unity, the proportion of disposable income earned by liquidity-constrained consumers is 65.5 % in 1953 for the case #12 which has the highest explanatory power among four cases, and falls monotonically over time to 37.4 % in 1985. The proportion of liquidity constrained consumers for the other cases exhibits a similar trend and it falls by 36.3 (case #10) to 46.4 (case #11) percentage points from 1953 to 1985.

The evidence obtained here for Taiwan seems to be incompatible with that obtained above, which supports the existence of a quadratic time preference schedule. These mixing results might be due to multicollinearity among the income variables. Multicollinearity prevents us from disentangling the dependence of the time preference rate and the proportion of liquidity-constrained consumers on income variables.

For Korea inclusion of the term of $\left(YD_t + YD_{t-1}\right)\Delta(YD)_t$ fails to improve the estimation of the consumption function. It turns out that none of the coefficients on

[12]The proportion of liquidity-constrained consumers may also depend on the wealth accumulated by the consumers. The property income received by consumers can measure the extent to which the consumers accumulate assets, so that we formulate the proportion of liquidity-constrained consumers as a function of disposable income rather than labor income.

Table 9.7 Estimation of consumption function with time preference rate and liquidity-constraint dependent on real income level

Instrument list		Constant	ΔYD_t	Y_{t-1}	Y_{t-1}^2	$(YD_t + YD_{t-1})_x \Delta YD_t$	$\ln(1+r_t)$	$R^2/D.W.$
Taiwan								
1.	#9	−3.8481 (−1.63)	0.6516 (2.46)	0.0649 (2.77)		−0.00065 (−1.85)		0.3723 1.8477
2.	#10	−3.7187 (−1.82)	0.8443 (3.72)	0.0429 (2.19)		−0.00058 (−1.99)		0.5971 2.0478
3.	#11	(−4.5592) (−2.30)	0.7746 (4.38)	0.0645 (3.02)		−0.00075 (−2.57)		0.4292 1.9243
4.	#12	−2.7785 (−1.59)	1.7190 (5.09)	0.0395 (2.05)		−0.00045 (−1.98)		0.0633 2.0626
Japan (1-year time deposits rate)								
5.	#17	−0.1709 (−1.08)	0.4974 (1.37)	0.1257 (3.00)	−0.0083 (−3.57)	−0.0126 (−0.67)	2.9268 (3.91)	0.8701 2.1619
6.	#18	−0.0521 (−0.39)	0.2222 (0.78)	0.1025 (2.72)	−0.0075 (−3.54)	0.0016 (0.11)	3.1586 (5.46)	0.8911 2.1758
7.	#19	−0.1300 (−0.84)	0.4355 (1.30)	0.1142 (2.69)	−0.0077 (−3.27)	−0.0091 (−0.52)	2.8058 (4.18)	0.8776 2.2256
8.	#20	−0.0210 (−0.15)	0.1743 (0.58)	0.0940 (2.48)	−0.0071 (−3.30)	0.0042 (0.28)	3.0538 (5.42)	0.8942 2.2336

See the note to Table 9.1 for notations

the income variables but $\Delta(YD)_t$ are significant, whether the dependence of the time preference schedule on the income level is linear or quadratic.

The estimation results of the Japanese consumption function are shown in the lower part of Table 9.7. The significance of the coefficients of the income variables Y_{t-1} and Y_{t-1}^2 as well as the interest rate remains intact even after allowing for dependence of the liquidity-constraint on the income level. We still find the quadratic time preference schedule. The turning point in the time preference schedule is estimated between 1969 and 1970 for the instruments list #18 and #20, while it is between 1970 and 1971 for #17 and #19. This evidence is consistent with that obtained above. None of the coefficients on $(YD_t + YD_{t-1})\,\Delta(YD)_t$ are significant. It implies that the proportion of liquidity-constrained consumers is independent of the income level.

Finally, we estimate the consumption function derived under the different preference structure. We assume that the preference of the LCY-PIH type consumers is characterized by the utility function with constant relative risk aversion. Then the consumption change of the LCY-PIH consumers is written as

$$\Delta \ln(C_{it}) = \eta - (1/\tau)\ln(1+\delta_{it-1}) + (1/\tau)\ln(1+r_t) + \upsilon_{it}, \tag{9.13}$$

where

τ is the degree of relative risk aversion,
υ_{it} is an error term with $\mathrm{E}\left[\upsilon_{it} \middle| \Omega_{t-1}\right] = 0$.

For liquidity-constrained consumers, the consumption level is given by (9.5). Taking the logarithm and expressing in terms of difference form, we obtain

$$\Delta \ln (C_{it}) = \Delta \ln (YD)_{it}. \tag{9.14}$$

As is discussed in footnote 4, exact aggregation of consumption over different types of consumers is no longer possible, since consumption change is expressed in *logarithmic* difference. However, if we assume that the log of an average is approximated by an average of logs, combination of (9.13) and (9.14) yields total consumption change per capita as follows:

$$\begin{aligned} \Delta \ln (C_t) = \eta_1 &- ((1-\lambda_1)/\tau) \ln (1+\delta_{t-1}) + ((1-\lambda_1)/\tau) \ln (1+r_t) \\ &+ \lambda_2 \Delta \ln (YD)_t + u_t, \end{aligned} \tag{9.15}$$

where u_t is an error term. Substituting (9.8) into (9.15), we obtain the total consumption change as

$$\begin{aligned} \Delta \ln (C_t) = \eta_1 &- ((1-\lambda_1)/\tau)\, \rho_1 - ((1-\lambda_1)/\tau)\, \rho_2 Y_{t-1} \\ &- ((1-\lambda_1)/\tau)\, \rho_3 Y_{t-1}^2 + ((1-\lambda_1)/\tau) \ln (1+r_t) \\ &+ \lambda_2 \Delta \ln (YD)_t + u_t. \end{aligned} \tag{9.16}$$

Equation 9.16 is an equation to be estimated. For Taiwan and Japan we still detect the quadratic time preference schedule, while there is no correlation between the time preference rate and the income level in Korea. The estimation results of Taiwan and Japan are shown in Table 9.8. For Taiwan the turning point of the time preference schedule is estimated between 1975 and 1976, irrespective of the choice of the instruments. For Japan the interest rate affects the LCY-PIH consumer's allocation of the lifetime resources over time in a significant manner. The coefficient of relative risk aversion is estimated slightly below three, which is consistent with the findings obtained by other studies. The turning point is estimated to have occurred between 1965 and 1966 except for the case with the instruments list #26, where it is estimated between 1966 and 1967.

4 Concluding Remarks

This chapter examined empirically several hypotheses on the time preference schedule which has been traditionally assumed to be constant. The countries examined were Taiwan, Korea, and Japan. These three countries are characterized

Table 9.8 Estimation of consumption function with the time preference rate dependent quadratically on real labor income (logarithmic version)

Instrument list		Constant	$\Delta \ln(YD_t)$	Y_{t-1}	Y_{t-1}^2	$\ln(1+r_t)$	$R^2/D.W.$
Taiwan							
1.	#21	0.0004 (0.04)	0.4274 (2.91)	$0.1762 \cdot 10^{-3}$ (2.16)	$-0.2734 \cdot 10^{-6}$ (−2.04)		0.6489 1.8528
2.	#22	−0.0013 (−1.14)	0.4764 (3.31)	$0.1665 \cdot 10^{-3}$ (1.99)	$-0.2567 \cdot 10^{-6}$ (−1.85)		0.6477 1.9161
3.	#23	0.0003 (0.03)	0.4294 (2.93)	$0.1755 \cdot 10^{-3}$ (2.12)	$-0.2717 \cdot 10^{-6}$ (−2.00)		0.6489 1.8553
4.	#24	−0.0016 (−0.17)	0.4841 (3.35)	$0.1647 \cdot 10^{-3}$ (1.92)	$-0.2524 \cdot 10^{-6}$ (−1.78)		0.6473 1.9243
Japan (1-year time deposits rate)							
5.	#25	0.0308 (1.69)	0.3321 (2.33)	0.0065 (1.66)	−0.00070 (−2.78)	0.3439 (4.19)	0.8716 2.1582
6.	#26	0.0259 (1.69)	0.3835 (2.83)	0.0066 (1.95)	−0.00068 (−2.92)	0.3452 (4.73)	0.8746 2.0879
7.	#27	0.0297 (2.02)	0.3500 (2.62)	0.0063 (1.73)	−0.00068 (−2.64)	0.3394 (4.29)	0.8730 2.1521
8.	#28	0.0269 (1.89)	0.3771 (3.04)	0.0064 (1.88)	−0.00068 (−2.83)	0.3418 (4.58)	0.8744 2.1086

See the note to Table 9.1 for notations

by high economic growth and high saving rates. We obtained some evidence against the constancy time preference rate hypothesis. The Fukao-Hamada hypothesis holds for Taiwan and Japan. The turning points in the time preference schedule were also analyzed and it was found that they are closely related to the turning points in the growth pattern of the country.

The dependence of the time preference rate on the income level of a country provides us with a fresh viewpoint on the dynamic relationship between savings behavior and the pattern of economic development. Personal savings are the main source of supply of capital. They are channeled through financial markets to firms and used for capital accumulation. New investment expands the production capacity embodying high productivity. More output is produced with more efficient technology, generating more income. When the time preference rate hinges on the income level, higher levels of income will affect the time preference rate and hence the consumption-savings pattern. If the intertemporal consumption tilts toward the future, as is suggested by Fisher, more savings will be available for capital accumulation and economic growth might be accelerated. On the other hand, if consumers prefer present consumption to that in the future, as is suggested by Uzawa, the speed of economic growth will slow down. More empirical evidence is to be accumulated for further understandings of the dynamic interaction of savings behavior with the pattern of economic development via the time-dependent time preference rate.

Acknowledgments Part of this work was financially supported by a research project 'Studies of the Prospective International Capital Ownership' sponsored by the National Institute for Research Advancement (NIRA). I am grateful for extremely useful comments and suggestions on earlier drafts to Kyoji Fukao, Vittorio Grilli, Koichi Hamada, Kazumasa Iwata, Masaaki Kuroyanagi, Makoto Sakurai, T.N. Srinivasan, Clifford Tan, and two anonymous referees.

Appendix: Data Description

Table 9.9 Data description

Variable	Description
Consumption[a]	Private consumption expenditure – expenditure on furniture, furnishings, and household equipment – expenditure on transport and communication
Price deflator	Implicit deflator corresponding to the consumption series constructed above
Labor income	Compensation of employees
Disposable income	Private consumption expenditure + personal savings
Interest rate	Interest rate on 3-month time deposits (Taiwan); interest rate on time deposits 1-year or more (Korea) interest rate on 1-year time deposits; bank debenture rate (Japan)

Data source: Taiwan – All the data but interest rate and population are taken from *National Income in Taiwan Area*, *The Republic of China* (Directorate-General of Budget, Accounting and Statistics, Executive Yuan). Population is taken from *Monthly Statistics of the Republic of China* (Directorate-General of Budget, Accounting and Statistics, Executive Yuan) and interest rate is taken from *Statistical Abstract of the Republic of China* (Directorate-General of Budget, Accounting and Statistics, Executive Yuan). *Korea* – All the data are taken from *Economic Statistics Yearbook* (The Bank of Korea)
Japan – All the data but interest rate and population are taken from *Report on National Accounts from 1995 to 1969* and *Annual Report on National Accounts* (Economic Planning Agency, Government of Japan). Population and interest rate are taken from *Economic Statistics Annual* (The Bank of Japan)
[a]For Japan, it is only since 1970 that separate figures have been available for consumption expenditure on non-durables, services, semi-durables, and durables

Addendum: Further Evidence[13]

In Ogawa (1993), I investigated the hypothesis that the time preference rate of a nation depended on the stage of economic development. Specifically, I examined four hypotheses on the time preference schedule. The first hypothesis posited that the time preference rate is time-invariant. The second hypothesis posited that the

[13]This addendum has been newly written for this book chapter.

time preference rate decreases as the economy develops (Fisher hypothesis). The third hypothesis posited that the time preference rate increases as the economy develops (Uzawa hypothesis). The fourth hypothesis, which is a combination of the second and third hypotheses, posited that as a nation develops, the time preference rate decreases up to a certain point, and thereafter, the time preference rate increases (Fukao-Hamada hypothesis). I tested these hypotheses based on the aggregate time series data of three countries, Japan, Taiwan and Korea, all three of which experienced high economic growth during the post-war period. The sample period was from 1952 to 1985 for Taiwan, from 1956 to 1987 for Korea and from 1955 to 1986 for Japan. I found that the Fukao-Hamada hypothesis was supported for Taiwan and Japan, while the time preference rate was constant for Korea.

Based on the evidence from that earlier study, an interesting research question is whether the findings from 20 years ago still hold when the sample period is extended to cover the recent turbulent decades. The purpose of this addendum is to reexamine the four hypotheses with respect to the time preference schedule based on the updated aggregate data of Japan. The sample period covers 1955–2006, thus including the bubble period of the late 1980s and the lost decade of the late 1990s to the 2000s. Fortunately, as the aggregate household dataset for this extended period is available in Ogawa (2010), I use this dataset to reexamine the four hypotheses for the time preference rate.

There is a caveat in comparing the evidence in Ogawa (1993) with the new evidence reported in this addendum, as the variables employed for estimation in this addendum are slightly different from the original variables. First, my new consumption series uses private final consumption expenditures, while the original series excluded expenditures on durable goods.[14] Second, new labor income contains both employee compensation and net mixed income excluding imputed values of owner-occupied dwellings, while the original series contained only the former. Third, I use contracted interest rates of loans as the interest rate, while the original interest rates were those for 1-year time deposits and bank debentures. Fourth, I add exogenous variables such as real government final consumption expenditures, general government real gross fixed capital formation, and real money supply (M2 + CD) to the list of instruments. Finally, a new series except the interest rate is measured on a per household basis, while the original series was on a per capita basis.

I estimate the consumption function with the time preference schedule (δ_{t-1}) specified as a linear or quadratic function of real labor income (9.16). Furthermore, I estimate six variants of the consumption function specified above: three that do not take liquidity-constrained households into consideration and three that take liquidity-constrained households into consideration. The instruments include thrice-lagged growth rates of private final consumption, government final consumption

[14] Consistent consumption series by type of consumption are available only after 1970.

Table 9.10 Estimation results of consumption function with the time preference rate dependent on real labor income

Equation number	Constant	$\ln(1+r_t)$	Y_{t-1}	$Y^2{}_{t-1}$	$\Delta\ln(YDt)$	R^2/p-value of overidentifying restrictions test
1.	0.0276*** (4.45)	−0.1534 (−0.74)				0.0116/0.005
2.	0.1098*** (9.46)	0.1027 (0.80)	−0.0196*** (−7.79)			0.5586/0.212
3.	−0.0293 (−0.55)	0.2083 (1.52)	0.0556** (1.98)	−0.0095*** (−2.67)		0.5846/0.268
4.	0.0042 (1.30)	0.1418* (1.68)			0.7526*** (10.48)	0.8348/0.348
5.	0.0159 (0.55)	0.1660** (2.10)	−0.0025 (−0.46)		0.7021*** (3.36)	0.8426/0.231
6.	−0.0174 (−0.54)	0.2024** (2.41)	0.0219 (1.10)	−0.0035 (−1.24)	0.5676*** (3.23)	0.8434/0.377

Notes: See the note to Table 9.1 for notations
*, **, *** significant at 10 %, 5 % and 1 % level, respectively

expenditures, general government gross fixed capital formation, and real money supply in addition to thrice-lagged interest rates and disposable income and twice-lagged labor income and squared labor income.

The estimation results are shown in Table 9.10. I find that the linear term of labor income is significantly positive and that the quadratic term is significantly negative. These findings imply that the Fukao-Hamada hypothesis is supported. That is, as the economy develops the time preference rate declines up to a certain point, and thereafter, the time preference rate increases. The value of real labor income at the turning point, which is estimated to be between 1963 and 1964, is 2.92 million yen.[15] However, the dependence of the time preference schedule on labor income is no longer significant once liquidity-constrained households are taken into account. While this might be due to multicollinearity between labor income and the growth rate of disposable income, we cannot obtain precise estimates of the proportion of liquidity-constrained households and the dependence of the time preference schedule on labor income simultaneously. There are two research avenues to overcome this problem of multicollinearity. One is to use country panel data that combine the time series data of several countries, and the other is to use household-based panel data.

[15]In Ogawa (1993), the turning point was estimated to be between 1965 and 1967 under the specification of the utility function with constant relative risk aversion, which is quite close to my new estimate.

References

Campbell JY, Mankiw NG (1989) Consumption, income, and interest rates: reinterpreting the time series evidence, NBER macroeconomics annual. NBER, Cambridge, pp 185–216

Campbell JY, Mankiw NG (1990) Permanent income, current income, and consumption. J Bus Econ Stat 8:265–279

Fisher I (1907) The rate of interest. Macmillan, London

Fukao K, Hamada K (1991) The Fisherian time preference and the evolution of capital ownership patterns in a global economy, NIRA research output 4, no. 1, 86–97

Giovannini A (1985) Saving and the real interest rate in LDCs. J Dev Econ 18:195–217

Hall R (1978) Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. J Polit Econ 86:971–987

Haque NU, Montiel P (1989) Consumption in developing countries: tests for liquidity constraints and finite horizons. Rev Econ Stat 71:408–415

Hayashi F (1982) The permanent income hypothesis: estimation and testing by instrumental variables. J Polit Econ 90:895–916

Hong W (1988) Time preference in dynamic trade models: an empirical critique. Econ Dev Cult Chang 36:741–751

Lawrance EC (1991) Poverty and the rate of time preference: evidence from panel data. J Polit Econ 99:54–77

Ogawa K (1990) Cyclical variations in liquidity-constrained consumers: evidence from macro data in Japan. J Jpn Int Econ 4:173–193

Ogawa K (1991) Income risk and precautionary saving. Keizai Kenkyu 42:139–152 (in Japanese)

Ogawa K (1993) Economic development and time preference schedule: the case of Japan and East Asian NICs. J Dev Econ 42:175–195

Ogawa K (2010) Consumption function. In: Minotani T, Maki A (eds) Handbook of applied econometrics. Asakura Shoten, Japan, pp 106–138 (in Japanese)

Uzawa H (1968) Time preference, the consumption function, and optimum asset holdings. In: Wolfe JN (ed) Value, capital and growth: papers in honour of Sir John Hicks. University of Edinburgh Press, Edinburgh, pp 485–504

White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48:817–838

Chapter 10
Luxury and Wealth

Shinsuke Ikeda

Countries which have sumptuary laws, are generally poor.
—(Sir Dudley North, *Discourses Upon Trade*, 1691, p. 14)

Abstract I develop a dynamic theory of luxury consumption, particularly emphasizing the causal effect that pursuit of luxury goods has on wealth accumulation. A quasi-luxury is defined as a good whose marginal rate of substitution is increasing in a utility index. Under certain conditions, it is indeed a luxury good. When current wealth holding falls short of (exceeds) long-run needs, luxury consumption is postponed more (less) easily than necessity consumption, due to a lower (higher) time preference for luxury and/or a higher intertemporal elasticity of substitution thereof. Preferences for quasi-luxuries lead to a higher steady-state value of wealth or capital.

Keywords Luxury • Weakly non-separable preferences • Time preference • Wealth • Intertemporal elasticity of substitution

1 Introduction

By definition, the wealthier allocate higher proportions of their expenditures to luxuries than the poorer do. The standard price theory would describe this by saying that the richer consume more luxuries because they have more wealth. Using this static argument which takes wealth holding as given, the pursuit of luxuries has often been condemned for enhancing the propensity to spend and decreasing saving,

---

The original article first appeared in International Economic Review 47: 495–526, 2006. A newly written addendum has been added to this book chapter.

S. Ikeda (✉)
Institute of Social and Economic Research, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan
e-mail: ikeda@iser.osaka-u.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_10

thereby harming capital formation and/or worsening the balance of payments. Partly for this reason, many countries, especially developing ones, commonly levy high luxury taxes and/or high luxury import tariffs.[1] From a dynamic viewpoint, however, how much wealth consumers accumulate is a part of their lifetime utility-maximization problem, as is how much of each good they consume in each period. If optimal saving behavior itself depends on the preference for luxuries, the effects of luxury taxation might well differ from what has been commonly believed. A comprehensive understanding of luxury consumption and its policy implications requires dynamic analysis.

This chapter develops a dynamic theory of luxury consumption by using a recursive preference model, focusing on the bilateral relationship between luxury and wealth accumulation. Particularly emphasized is the causal effect that pursuit of luxury goods has on wealth accumulation. The topic of how preferences for luxury goods affect the total amount of national wealth is so old that it goes back to at least the seventeenth century. For example, in his famous, pioneering work, Mandeville (1714), who was an immigrant from The Netherlands in a Golden Age, stressed that the pursuit of luxury, which had been long criticized for corrupting people, should be evaluated positively as a means to greater national wealth.[2] List (1841) emphasized that the pursuit of luxuries encourages saving, thereby stimulating economic growth. Sombart (1912) examined how strong preferences for luxury had contributed to the development of capitalistic economies and capital accumulation.[3] The chapter is a tentative response to the long controversy using modern consumption theory.

Although there are strong empirical evidences against preference homotheticity (e.g., Attanasio and Weber 1995; Blundell et al. 1994; Parker 1999), few attempts have been made to analyze luxury consumption from the viewpoint of intertemporal utility maximization.[4] Browning and Crossley (2000), in an important exception using a two-good, two-period model of the time-additive utility function, show that luxury goods have higher intertemporal elasticities of substitution (IES)and therefore characterize them as "easier to postpone" than necessity goods. Their

[1]In the recent past, for example, Thailand levies luxury taxes on the entertainment industry which could be as much as 25 %. Algeria imposes a 150 % tax on caviar. Indonesia's luxury tax and import tariff on passenger cars of more than 3,000 cc engines amount to 50 and 80 %, respectively.

[2]For the old controversy on luxuries, see Brinkmann (1957), Mason (1998), and Brewer (1998).

[3]The focus of this chapter is somewhat different from what those old social scientists stressed. They emphasized the effects of luxury on capital accumulation through market expansions and/or creations (see Brinkmann 1957; Mason 1998), whereas I focus on the saving-promoting effect of luxury. However, these two are not so different since capital accumulation could not be sustained without saving when international capital markets are imperfect.

[4]Besley (1989) proposes a new definition for luxury which is useful in "dynamic"applications. However, dynamic optimization is not discussed there. Baland and Ray (1991) examine the effect of capital accumulation on unemployment by using a model in which luxury and basic goods compete for the use of the scarce resources. However, the analysis is essentially static, assuming that capital accumulation is exogenous.

characterization with the usual time-additive preferences, however, does not imply any particular effect of luxuries on wealth accumulation.

The novelty of this chapter is to characterize luxury in the recursive preference framework, instead. Based on the procedures used by Shi (1994), weakly non-separable preferences are specified in a simple two-good model of recursive preferences, such that the intratemporal marginal rate of substitution (MRS) between two goods depends on the future consumption streams through current welfare.[5] Relative preferences for the two goods, measured by the MRS, then depend on current welfare. With the resultant non-homothetic preference structure, a good whose MRS is increasing (decreasing) in welfare is called a *quasi-luxury* (*quasi-necessity*). Under a certain condition, quasi-luxuries are identical to luxury goods. The purposes of the present chapter are (i) to characterize (quasi-) luxury goods from the viewpoint of intertemporal resource allocation; (ii) to show that wealth accumulation and distribution depend on consumers' *intratemporal* preferences for these luxurious goods; and (iii) to examine the effects of taxation on the luxurious goods.

Weakly non-separable preferences produce two properties. First, in contrast to Browning and Crossley (2000), luxury and necessity goods are characterized in terms of both good-specific time preferences and good-specific IESs. By using the general characterization, it is shown that a higher IES is neither necessary nor sufficient for a good to be a luxury. For example, even when a good x has a lower IES than the other good c, x can be a luxury good if, under wealth accumulation (decumulation), the rate of time preference with respect to x is lower (higher) than that with respect to c. My more general model helps to better understand consistently the often-observed consumption patterns of luxuries (e.g., nice restaurant dinners) and necessities (e.g., ordinary dinners at home). When current wealth holding falls short of the long-run required level, consumers usually accumulate wealth by holding down spending on luxuries more than on necessities. This tendency is explained by a lower time preference for luxury consumption and/or a higher IES thereof. When wealth exceeds long-run needs, in contrast, consumers are likely to increase spending more on luxury goods relative to the long-run level than on necessities *today* to decumulate wealth. This is attributed to a higher time preference for luxury consumption and/or a higher IES thereof.

Secondly, quasi-luxuries in weakly non-separable preference models induce a preference for wealth and thereby raise the long-run wealth level. Given constant market prices, strong preferences for quasi-luxury goods are shown to promote consumers' optimal wealth accumulation. The discussion is then extended to two general-equilibrium frameworks, a neoclassical production-economy model and a

[5]Weakly non-separable preferences under recursive preferences are analyzed by Lucas and Stokey (1984), Judd (1985), Epstein et al. (1988), Shi (1994), and Ikeda (2001, 2003). Shi conducts the most systematic analysis to discuss the intertemporal leisure-consumption choice under distortionary taxation on capital and labor.

two-country world economy model, thereby showing two results: (i) the stronger the quasi-luxury preference, the more capital is accumulated in steady state; and (ii) even if utility-discounting functions are both internationally (interpersonally) identical, a country with a stronger quasi-luxury preference holds more wealth in the long run than a country with a weaker preference. As a corollary of (i), luxury taxes are likely to decrease the long-run capital stock. Two corollaries obtain from (ii). First, a less-patient country with a stronger quasi-luxury preference can be wealthier than a more-patient one. Second, tariffs on luxury imports are likely to worsen the current account.[6]

There are several contributions to dynamic macroeconomics that can be reinterpreted in the luxury-necessity terms. In the status (wealth)-seeking literature such as Cole et al. (1992) and Corneo and Jeanne (1999), a high "status" could be regarded as a luxury good, so that the preference for luxury is growth promoting. Futagami and Shibata (1998) develop a status-preference model in which relative wealth generates utility, and thereby show that a less patient agent can possess larger wealth than a more patient one if the less patient agent is greedier than the other. In the status (wealth)-seeking models, however, they directly assume a preference for wealth. In my model, in contrast, a wealth preference is induced by preferences for certain marketed goods.

The rest of the chapter is organized as follows. Section 2 presents the basic framework to define (quasi-) luxury goods, characterizes the optimal consumer behavior, and analyzes the relation between the preference for (quasi-) luxury goods and consumers' optimal wealth accumulation. Section 3 extends the analysis to a simple neoclassical model and a two-country economy model. Section 4 concludes the chapter.

2 Consumer Behavior with Quasi-luxury Goods

I start with defining and characterizing quasi-luxury and luxury goods by considering intertemporal consumer choice, provided that good prices and the interest rate are given constant.

[6]In the dynamic macro literature, international (interpersonal) wealth distribution has been explained by referring to differences in four determinants: (i) the subjective discount rate (e.g., Devereux and Shi 1991; Ramsey 1928); (ii) productivity growth (e.g., Frenkel and Razin 1992; Obstfeld and Rogoff 1996); (iii) age structures (e.g., Blanchard 1985; Buiter 1981); and (iv) random income fluctuation (e.g., Becker and Zilcha 1997; Clarida 1990). To focus on implications of the luxury preference, these factors are not considered here.

2.1 Consumer Preferences and Quasi-luxury Goods

Consider an infinitely lived consumer. There are two distinct consumption goods c and x. Let $u(c,x)$ and $\delta(c,x)$ denote the instantaneous utility function and the subjective discount-rate function, respectively. I specify consumer preferences by the following lifetime utility function:

$$U(0) \equiv \int_0^\infty u(c(t), x(t))\Theta(t)\,\mathrm{d}t \tag{10.1}$$

where $\Theta(t)$ represents a discount factor,

$$\Theta(t) = \exp\left\{-\int_0^t \delta(c(s), x(s))\,\mathrm{d}s\right\} \tag{10.2}$$

The functions $u(c,x)$ and $\delta(c,x)$ are assumed to be twice continuously differentiable and satisfy the standard regularity conditions: (i) $u < 0$; (ii) u is strictly increasing and strictly concave; (iii) $-u$ is log-convex; (iv) $\delta > 0$; and (v) δ is concave. Provided appropriate Inada conditions are met, the regularity conditions enable me to describe interior optimal solutions by the familiar marginal equalities and the transversality conditions as shown later.[7] As in the literature (e.g., Uzawa 1968), I assume increasing marginal impatience: $\delta_c > 0$ and $\delta_x > 0$,[8] where $\delta_c = \partial\delta(c,x)/\partial c$, etc.

Let $\phi(t)$ represent a time-t utility index which equals the utility $U(t)$ obtained from the consumption stream after time t. The corresponding generating function g is given by

$$g(c,x,\phi) = u(c,x) - \phi\delta(c,x)$$

with which utility evolution is expressed as

$$\dot{\phi} = -g(c,x,\phi) \text{ s.t. } \lim_{t\to\infty}\phi(t)\Theta(t) = 0 \tag{10.3}$$

where a dot represents the time derivative, i.e., $\dot{\phi}(t) = \mathrm{d}\phi(t)/\mathrm{d}t$. The first-order partial derivatives $g_c(c(t), x(t), \phi(t))$ and $g_x(c(t), x(t), \phi(t))$ equal the current-value marginal utilities of $c(t)$ and $x(t)$ in the Voltera sense, respectively. Since utility index ϕ takes negative values from the regularity condition (i),

[7]For the detailed discussions on the regularity conditions, see Epstein (1987a, pp. 72–75) and Obstfeld (1990, pp. 49–50).

[8]This property is controversial especially from the empirical viewpoint. This assumption is necessary for the appealing stability property derived below. For justifications of the assumption, see Epstein (1987a,b). For unstable dynamics under decreasing marginal impatience, see Hirose and Ikeda (2001, 2008).

the regularity conditions (ii) and (v), together with the assumption of increasing marginal impatience, ensure that $g_c(c, x, \phi) > 0$; $g_x(c, x, \phi) > 0$; $g_{cc}(c, x, \phi) < 0$; and $g_{xx}(c, x, \phi) < 0$. For brevity, I assume the following:

Assumption 1

$$g_{cx}(c, x, \phi) = 0$$

The intratemporal MRS between c and x is given by $-\mathrm{d}c/\mathrm{d}x|_{U(0)=\text{const.}} = g_x/g_c(c, x, \phi)$. When the MRS indeed depends on the current value of the utility index ϕ, consumer preferences are not weakly separable since the choice of the time-t consumption basket depends on future consumption plans. With the preferences, which are called by Shi (1994) weakly non-separable preferences, I define *quasi-luxury goods* as follows:

Definition 1 A good is a quasi-luxury (-necessity) good or simply a quasi-luxury (-necessity) to a consumer if his or her relative preferences for it, measured by the MRS, are increasing (decreasing) in the utility index ϕ.[9]

It is convenient if the non-separability index ξ is defined as

$$\xi(c, x, \phi) \equiv \frac{1}{g_x/g_c}\frac{\partial(g_x/g_c)}{\partial\phi} = \frac{\delta_c}{g_c} - \frac{\delta_x}{g_x} \tag{10.4}$$

The index ξ captures how relative preferences for x depend on ϕ.[10] I assume the following:

Assumption 2 Good x is a quasi-luxury good:

$$\xi(c, x, \phi) > 0$$

Assumption 2 holds valid for all feasible (c, x, ϕ), thereby ruling out the possibilities that quasi-luxury x reverses to a quasi-necessity good at some (c, x, ϕ). I assume it for simplicity, although the main results of comparative dynamics below remain valid as long as ξ is locally positive around the steady-state point.

Example 1 Assumptions 1 and 2 are satisfied if $u(c, x)$ and $\delta(c, x)$ are given by $u(c, x) = q(c) + v(x)$ and $\delta(c, x) = \theta(c) + \varphi(x)$, respectively, where (i)

[9] A quasi-luxury good is identical to what Shi (1994) calls a "less welfare-stabilizing" good. I coin the different terminology so that we can easily see in which direction a wealth increase affects the relative preference for the good, thereby focusing on the issue of luxury consumption. See also Footnote 13.

[10] The index ξ would be zero if the discount rate δ were a function of felicity u, as in Uzawa (1968) and Obstfeld (1982), or if δ were constant, as in the case of time-additive preferences, or if u were constant, as in Epstein and Hynes (1983).

functions q, v, θ, and φ satisfy the required regularity conditions (e.g., $q < 0, v < 0$); (ii) there exist $\inf q(c)$ and $\inf v(x)$; and (iii) functions θ and φ are given by

$$\theta(c) = \eta^c \{q(c) - \inf q(c)\} \text{ and} \varphi(x) = \eta^x \{v(x) - \inf v(x)\}$$

with η^c and η^x being constants satisfying $\eta^c > \eta^x > 0$. In this case, the non-separability index ξ satisfies

$$\xi(c, x, \phi) = \frac{\eta^c - \eta^x}{(1 - \phi\eta^c)(1 - \phi\eta^x)} > 0$$

as in Assumption 2.

2.2 *Optimal Consumer Behavior*

2.2.1 Optimization

Taking the good c as numeraire, let p and r denote the relative price of quasi-luxury x and the interest rate, respectively. In this section, I assume that p and r are given constant. All economic models studies in this chapter will exhibit the property that p is constant over time. Let a be the consumer's *total* wealth, which is the sum of financial wealth and human wealth. He or she maximizes lifetime utility (10.1) subject to the flow budget constraint,

$$\dot{a} = ra - c - px \tag{10.5}$$

and the initial condition, $a(0) =$ given. Letting λ denote the current-value shadow price of savings, the optimality conditions are given in the current-value terms as:

$$\begin{aligned} g_c(c, \phi)\,(\equiv u_c(c, x) - \phi\delta_c(c, x)) &= \lambda \\ g_x(x, \phi)/g_c(c, \phi) &= p \\ \dot{\lambda} &= (\delta(c, x) - r)\,\lambda \end{aligned} \tag{10.6}$$

together with (10.3) and the transversality condition for a.

Following Obstfeld (1990), define the rate of time preference ρ^c with respect to c as

$$\rho^c \equiv -\left.\frac{\mathrm{d}\ln\{g_c(c(t), \phi(t))\,\Theta(t)\}}{\mathrm{d}t}\right|_{\dot{c}=0}$$

where $g_c \Theta(t)$ represents the present-value marginal utility of c.[11] Differentiating the first equation of (10.6) by t and taking into account $g_{c\phi} = -\delta_c$ and (10.3), we obtain $\dot{\lambda} = g_{cc}\dot{c} + \delta_c g$. Substituting this and $g_c = \lambda$ into the third equation of (10.6) and rearranging yields

$$\frac{\dot{c}}{c} = \sigma^c (c, \phi) (r - \rho^c (c, x, \phi)) \tag{10.7}$$

where

$$\sigma^c (c, \phi) \equiv -\frac{g_c (c, \phi)}{c g_{cc} (c, \phi)} \tag{10.8}$$

$$\rho^c (c, x, \phi) = \delta (c, x) - \frac{\delta_c (c)}{g_c (c, \phi)} g (c, x, \phi) \tag{10.9}$$

Function $\sigma^c (c, \phi)$, defined by (10.8), represents the IES with respect to good c. In exactly the same way, the optimal dynamics of quasi-luxury consumption x can be expressed in terms of the IES and the time-preference rate with respect to x; i.e., $\sigma^x (x, \phi)$ and $\rho^x (c, x, \phi)$.[12] In sum, given the market prices (p, r) and the initial total wealth a_0, the optimal consumption plan $\{a(t), c(t), x(t), \phi(t)\}_{t=0}^{\infty}$ for the consumer is generated by (10.3), (10.5), (10.6), and (10.7) under the initial and transversality conditions.

By comparing (10.9) with the corresponding equation for good x, quasi-luxury goods can be characterized from the viewpoint of impatience, as in Shi (1994): from (10.3) and the definition of ξ, the difference between the two good-specific time preferences satisfies

$$\rho^c - \rho^x = \xi \dot{\phi} \tag{10.10}$$

which, under Assumption 2, implies $\rho^x \lesseqqgtr \rho^c$ as $\dot{\phi} \gtreqqless 0$.[13] When ρ^x is lower (or higher) than ρ^c, I say that consumers are more (or less) patient with respect to good x than with respect to c. With the terminology, Eq. (10.10) characterizes quasi-luxuries as follows:

[11] In discrete-time settings, the rate of time preference is defined as the MRS of today's consumption to tomorrow's, evaluated at a flat consumption path. The ρ^c is the continuous-time limit of the rate.

[12] In the general case in which relative price p varies over time, the rate of time preference with respect to x depends on $\dot{p}/p$: $\rho^x = \delta - g\delta_x / g_x + \dot{p}/p$. Throughout the chapter, however, price p does not change over time.

[13] From (10.9) and the definition of ξ, $\rho_\phi^x = \rho_\phi^c - \xi r$ holds around the steady state. A quasi-luxury good ($\xi > 0$) can thus be characterized around the steady state by $\rho_\phi^x < \rho_\phi^c$. Since increasing impatience stabilizes consumption dynamics, this implies that a quasi-luxury good is "less welfare-stabilizing" (Shi 1994) than the other good and vice versa.

Proposition 1 *When the utility index is increasing (decreasing) over time, consumers are more (less) patient with respect to quasi-luxury consumption than with respect to quasi-necessity consumption.*

Let $V(a)$ denote the value function, i.e., the maximized value of utility expressed as an increasing function of wealth holding, which is time-variant as in the standard recursive preference model (e.g., Obstfeld 1990). By construction, the maximized $\phi(t)$ equals $V(a(t))$, implying that optimal ϕ comoves positively with total wealth a. From Proposition 1, therefore, under wealth accumulation (decumulation), consumers are more (less) patient with respect to quasi-luxuries than with respect to quasi-necessities.

2.2.2 Steady State

The steady-state consumption basket $(\bar{c}, \bar{x})$ is determined by

$$\delta(\bar{c}, \bar{x}) = r \tag{10.11}$$

$$\frac{g_x(\bar{x}, u(\bar{c}, \bar{x})/\delta(\bar{c}, \bar{x}))}{g_c(\bar{c}, u(\bar{c}, \bar{x})/\delta(\bar{c}, \bar{x}))} = p \tag{10.12}$$

Steady-state wealth holding and welfare are then given by

$$r\bar{a} = \bar{c} + p\bar{x} \tag{10.13}$$

$$\bar{\phi} = u(\bar{c}, \bar{x})/r$$

respectively.

By using Fig. 10.1, the determination of the steady-state consumption plan can be illustrated as follows. Schedule RR' represents (10.11), depicting the locus of $(\bar{c}, \bar{x})$ that equalizes the steady-state rate of time preference to the interest rate. It could be referred to as the steady-state time preference curve. It is downward-sloping since the gradient is given by $-\delta_c/\delta_x < 0$. Schedule FF' represents (10.12), along which the steady-state MRS between the two goods equals the corresponding relative price. I call it the steady-state contract curve. From (10.12) and the definition of ξ, its gradient is given by

$$\left.\frac{\mathrm{d}\bar{x}}{\mathrm{d}\bar{c}}\right|_{FF'} = \frac{\xi - \delta g_{cc}/g_c^2}{p\left(-\delta g_{xx}/g_x^2 - \xi\right)}$$

implying that schedule FF' is upward- or downward-sloping as ξ is smaller or larger than $-\delta g_{xx}/g_x^2$. To ease the local analysis below, I follow Shi (1994) in assuming:

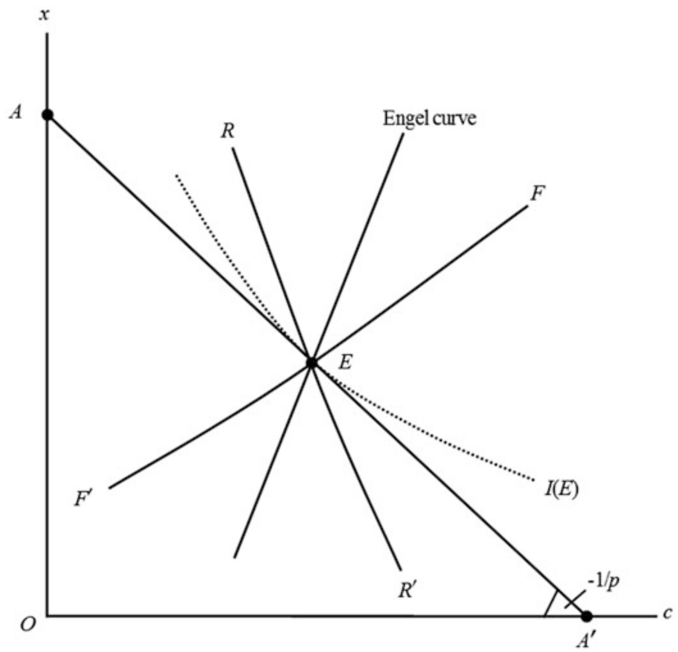


Fig. 10.1 The determination of the steady-state consumption plan

Assumption 3 The steady-state contract curve FF' is upward-sloping around the steady-state point:

$$\xi < -\frac{\delta g_{xx}}{g_x^2}$$

As shown later, this assumption assures that the optimal consumption dynamics are locally stable; and that c and x are both normal goods around the steady-state point.[14,15]

The steady-state consumption basket $(\bar{c}, \bar{x})$ is determined uniquely at the intersection point E of schedules RR' and FF'. Given the consumption basket, in turn, the no-saving condition (10.13) is depicted by schedule AA', which goes through point E with slope $-1/p$. Its horizontal intercept gives the steady-state interest income $r\bar{a}$. Define steady-state indifference curve $I\ (E)$ as

[14]Shi (1994, appendix) proves that this assumption ensures the local concavity of the preferences.

[15]Even without Assumption 3, the main results of the chapter remain valid if the steady-state point is locally stable, $\Psi > 0$ (see (10.16) below).

$$I\left(E\right)=\left\{\left(\bar{c},\bar{x}\right)\mid\frac{u\left(\bar{c},\bar{x}\right)}{\delta\left(\bar{c},\bar{x}\right)}=\text{utilityat}E\right\}\tag{10.14}$$

The indifference curve, whose gradient equals the MRS for the quasi-luxury $\bar{c}$, is tangent to budget schedule AA' at point E. Figure 10.1 is similar to the usual map for static consumption choice, except that the location of budget schedule AA' is endogenously determined to attain the consumption basket $(\bar{c}, \bar{x})$ that is determined by schedules RR' and FF'. The resulting steady-state wealth holding may well reflect *intratemporal* preferences for the two goods as well as time preference.

Note from (10.4) and Assumption 2 that the MRS for the quasi-necessity $\bar{c}$, g_c/g_x, is smaller than that along the discount rate, δ_c/δ_x, at point E:

$$-\left.\frac{\mathrm{d}\bar{x}}{\mathrm{d}\bar{c}}\right|_{\bar{\phi}=\text{utilityat}E}<-\left.\frac{\mathrm{d}\bar{x}}{\mathrm{d}\bar{c}}\right|_{\delta=r}\tag{10.15}$$

This implies that an increase in quasi-luxury consumption x in exchange for quasi-necessity consumption c, keeping the discount rate equal to r, would necessarily enlarge utility index $\bar{\phi}$ and hence the required amount of wealth $\bar{a}$. This property plays an important role for consumers' quasi-luxury preferences to affect wealth accumulation.

2.2.3 Local Dynamics

To characterize dynamic properties of the optimal solution, let us examine the local dynamics around the steady-state point. As shown by Appendix A.1, the optimal dynamic system of (10.3), (10.5), (10.6), and (10.7) can be reduced to an autonomous system with respect to $l \equiv (c, x, a)$. Linearizing the system yields $\dot{l}\left(t\right)=A\hat{l}\left(t\right)$, where the coefficient matrix A is given in Appendix A.1; and where the hat above l represents deviations from the steady-state value of the variable. It can be verified that the linear system has two positive roots and one negative root ω:

$$\omega\equiv\frac{r-\sqrt{r^2+\frac{4g_c^2g_x^2}{g_{cc}g_{xx}}\Psi}}{2}$$

where

$$\Psi\equiv\frac{\delta_x}{g_x}\left(\xi-\frac{rg_{cc}}{g_c^2}\right)+\frac{\delta_c}{g_c}\left(-\frac{rg_{xx}}{g_x^2}-\xi\right)\tag{10.16}$$

which is positive under Assumption 3. Any other paths than the saddle path governed by ω cannot satisfy the transversality conditions. The optimal consumption plan is uniquely determined on the saddle arm, which can be derived from the eigen vector associated with ω as:

$$\hat{c}(t) = -\frac{g_x^2}{(g_{xx} + p^2 g_{cc})} \left\{ \frac{g_{xx}\omega}{g_x^2} + \left(-\frac{r g_{xx}}{g_x^2} - \xi \right) \right\} \hat{a}(t)$$

$$\hat{x}(t) = \frac{g_x g_c}{(g_{xx} + p^2 g_{cc})} \left\{ \frac{g_{cc}(r-\omega)}{g_c^2} - \xi \right\} \hat{a}(t) \tag{10.17}$$

$$\hat{\phi}(t) = g_c \hat{a}(t)$$

where the state variable $a(t)$ evolves by $\dot{a}(t) = \omega \hat{a}(t)$ subject to $\hat{a}(0) = a_0 - \bar{a}$.

As total wealth holding monotonically approaches its steady-state quantity from a given a_0, the transitional paths for consumptions c and x and utility index ϕ are determined on stable arms (10.17). Under Assumption 3, the stable arms are all positively-sloping: consumptions c and x as well as utility index ϕ co-move positively with total wealth a. That is, the two goods are normal. By eliminating $\hat{a}$ from the first two equations in (10.17), a positively-sloping saddle trajectory in the (c, x)-space is obtained as

$$\hat{x}(t) = -\frac{g_c}{g_x} \left\{ \frac{g_{cc}(r-\omega)}{g_c^2} - \xi \right\} \left\{ \frac{g_{xx}\omega}{g_x^2} + \left(-\frac{r g_{xx}}{g_x^2} - \xi \right) \right\}^{-1} \hat{c}(t) \tag{10.18}$$

This schedule depicts positive co-movements of c and x generated by (endogenous) wealth variation. It can be regarded as the (linearized) *Engel curve* defined with respect to total wealth or permanent income. The Engel curve is also illustrated in Fig. 10.1.

The rates of time preference ρ^c and ρ^x play a key role in producing the optimal dynamics. Note that $\dot{c} = \omega \hat{c}$ and $\dot{x} = \omega \hat{x}$ on the saddle arm. Substituting the relations into (10.7) and the corresponding equation for x, and solving the results for the time preference rates yield

$$\rho^c(t) = r - \frac{g_x^2 g_{cc}\omega}{g_c(g_{xx} + p^2 g_{cc})} \left\{ \frac{g_{xx}\omega}{g_x^2} + \left(-\frac{r g_{xx}}{g_x^2} - \xi \right) \right\} \hat{a}(t) \tag{10.19}$$

$$\rho^x(t) = r + \frac{g_c g_{xx}\omega}{g_{xx} + p^2 g_{cc}} \left\{ \frac{g_{cc}(r-\omega)}{g_c^2} - \xi \right\} \hat{a}(t) \tag{10.20}$$

From Assumption 3, the time preference rates, ρ^c and ρ^x, are thus positively correlated to total wealth a. When $a(t)$is smaller than $\bar{a}$, ρ^c and ρ^x are both lower than r, causing consumers to accumulate wealth. As $a(t)$ accumulates, ρ^c and ρ^x rise gradually up to r.

2.3 Luxury Goods

Let us next relate quasi-luxury goods to luxury goods. By using a two-period, time-additive preference model, Browning and Crossley (2000) prove that luxury goods have higher IESs. My main interest is to give a more general characterization to luxury goods by using the present recursive-preference model.

In considering luxury goods in a dynamic setting, two points should be noted: first, relevant income in considering consumption baskets is not usual current income but permanent income or total wealth[16]; second, however, total wealth is an endogenous variable in the intertemporal consumption choice setting. Letting z represent total expenditure: $z = c + px$, it is proposed that luxury goods be defined along the optimal time-path of (c, x, a):

Definition 2 A good is a luxury (necessity) good or simply a luxury (necessity) to a consumer if, for given constant market prices (p, r) and initial total wealth a_0, the consumption share of the good in the consumer's total expenditure z is increasing (decreasing) in total wealth a along the optimal time-path of (c, x, a).[17]

Remark 1 Alternatively luxuries could defined by the property that an *exogenous* increase in the initial value of total wealth $a(0)$ (due to a wealth transfer, for instance) causes *initial* expenditures on the luxury $px(0)$ to increase more than proportionately than *initial* expenditures on the necessity $c(0)$, i.e.,

$$\frac{\mathrm{d}\left[px(0)/c(0)\right]}{\mathrm{d}a(0)} > 0$$

The alternative definition is equivalent to my definition since the optimal consumption policies in this model are given by time-invariant functions of total wealth holding, $c(t) = P^c(a(t))$ and $x(t) = P^x(a(t))$,[18] so that

$$\frac{\mathrm{d}\left[px(0)/c(0)\right]}{\mathrm{d}a(0)} > 0 \Leftrightarrow \frac{\mathrm{d}\left[pP^x(a(0))/P^c(a(0))\right]}{\mathrm{d}a(0)} > 0$$
$$\Leftrightarrow \frac{\mathrm{d}\left[pP^x(a(t))/P^c(a(t))\right]}{\mathrm{d}a(t)} > 0$$

[16]Hamermesh (1982) estimates the *permanent*-income elasticities of various consumptions.

[17]Browning and Crossley (2000) define a luxury as a good whose total-expenditure elasticity of the Marshallian demand is larger than unity. My definition is a natural extension of theirs to the dynamic recursive preference framework.

[18]For the linearized case, solution (10.17) gives the optimal policy functions. More generally, solve the first and second conditions in (10.6) for c and x, and denote the results by $c = C(\phi, \lambda)$ and $x = X(\phi, \lambda)$, respectively. Since the optimal ϕ and λ can be expressed in terms of the value function $V(a)$ as $\phi = V(a)$ and $\lambda = V_a(a)$, the optimal consumption can be expressed as time-invariant functions of a: $c = C(V(a), V_a(a)) \equiv P^c(a)$ and $x = X(V(a), V_a(a)) \equiv P^x(a)$.

$$\Leftrightarrow \frac{\mathrm{d}\left[px(t)/c(t)\right]}{\mathrm{d}a(t)} > 0$$

The optimal value of utility index $\phi(t)$, which equals $V(a(t))$, is increasing in wealth holding $a(t)$. Thus an increase in wealth increases the optimal value of the utility index, enhances the relative preference for quasi-luxury x to quasi-necessity c, and is likely to cause *uneven* expansion of consumption in favor of x. To derive a necessary and sufficient condition for x to be a luxury good, consider the evolution of the optimal consumption ratio px/c when a accumulates. From (10.7) and the corresponding Euler equation for x, it is given by:

$$\dot{\ln}(px/c) = \sigma^x(r-\rho^x) - \sigma^c(r-\rho^c). \tag{10.21}$$

By using (10.10), the right hand side of this equation can be rewritten as $(r-\rho^x)(\sigma^x-\sigma^c) + \xi\sigma^c\dot{\phi}$. The necessary and sufficient condition for good x to be a luxury is thus given as follows:

Proposition 2 *Good x is a luxury good if and only if*

$$\mathrm{sign}\left((r-\rho^x)(\sigma^x-\sigma^c)/\sigma^c + \xi\dot{\phi}\right) = \mathrm{sign}(\dot{a}) \tag{10.22}$$

From the Euler equation for x, the term $(r-\rho^x)$ in (10.22) is of the same sign as $\dot{\phi}$ and $\dot{a}$ if x is a normal good. The first term on the left hand side represents what Browning and Crossley (2000) found. In the case of weakly separable preferences (i.e., $\xi = 0$), Proposition 2 indeed reduces to their proposition that good x is a luxury if and only if it has a larger IES than c: $\sigma^x > \sigma^c$. In the weakly non-separable preference case (i.e., $\xi \neq 0$), however, a higher IES is neither necessary nor sufficient for a good to be a luxury. Even when x is not as easy to postpone as c when measured by the IESs (i.e., $\sigma^c > \sigma^x$), x can be a luxury if it is a quasi-luxury as in Assumption 2 (i.e., $\xi > 0$), that is if, under wealth accumulation (decumulation), consumers are more (less) patient with respect to x than with respect to c: $\rho^x < (>)\,\rho^c$. On the contrary, even when σ^x is higher than σ^c, x can be a necessity if it is a quasi-necessity. The second term $\xi\dot{\phi}$ in (10.22) captures this effect.

The present model describes consistently the often-observed consumption patterns of luxuries (e.g., nice restaurant dinners) and necessities (e.g., ordinary dinners at home) in terms of consumers' wealth-accumulating behavior.[19] When current wealth holding falls short of the long-run required level, consumers usually save by holding down spending on nice restaurant dinners more than on ordinary dinners at

[19]Hamermesh (1982) estimates the permanent-income elasticity of "food consumed at home" as 0.240 and that of "food consumed away from home" as 0.820, which implies that the both "foods" are necessity goods. "Nice restaurant dinners" in my example of luxuries represent meals at such fancy restaurants as treated by gastronomic guidebooks (e.g., Michelin).

home. In this sense, under wealth accumulation, luxuries are likely to be postponed more easily than necessities, as predicted by Browning and Crossley (2000). This tendency is attributed not only to a higher IES of luxury consumption but also to a lower time preference therefor.

In contrast, when wealth exceeds long-run needs, consumers are likely to increase spending more on restaurant dinners relative to the long-run level than on dinners at home *today* to decumulate wealth. In this sense, under wealth decumulation, luxuries are given priority over necessities *today*. From Propositions 1 and 2, this behavior can be explained by a higher time preference for luxury consumption and/or a higher IES thereof.[20]

Remark 2 Equation (10.22) is not based on a linear approximation. By using the linearized Engel curve obtained in the previous section, I can simplify the condition for luxury goods. Along the linearized Engel curve, both z and px can be expressed as functions of a by substituting (10.17) into $z = \hat{c} + p\hat{x} + \bar{z}$ and $px = p\hat{x} + p\bar{x}$. Differentiating px/z with respect to a yields:

$$\frac{\mathrm{d}\,(px/z)}{\mathrm{d}a} > 0 \Leftrightarrow \sigma^x - \sigma^c + \frac{\xi\sigma^c\sigma^x g_c\bar{z}}{r-\omega} > 0$$

where the right hand sides are evaluated at the steady-state point.

2.4 The Preference for Quasi-luxuries and Optimal Wealth Accumulation

By using the optimal intertemporal consumption plan derived, let us examine the implication which the preferences for quasi-luxuries have for steady-state wealth holding. The relative preferences for quasi-luxury goods are parameterized by α, with which the instantaneous utility function is re-specified as $u\,(c, x; \alpha)$:

$$u_\alpha \geq 0, u_{c\alpha} \leq 0, u_{x\alpha} \geq 0 \tag{10.23}$$

where at least one inequality holds strictly.[21]

[20]When current a is larger than $\bar{a}$, $\dot{a} < 0$, $\dot{\phi} < 0$, and $r - \rho^x < 0$, irrespective of whether or not x is a quasi-luxury good. From (10.10) and (10.22), $\rho^x > \rho^c$ or $\sigma^x > \sigma^c$ should be valid for x to be a luxury good.

[21]The quasi-luxury preference α, defined by (10.23), can be introduced to Example 1 by setting, e.g., $u\,(c, x; \alpha) \equiv q\,(c)\,/\alpha + v\,(x)\,, \alpha > 0$ with the same discount rate function as in Example 1. In this case, the non-separability index ξ is given by

$$\xi\,(c, x, \phi) = \frac{\alpha\eta^c - \eta^x}{(1-\alpha\phi\eta^c)\,(1-\phi\eta^x)}.$$

Therefore, parameter α should satisfy $\alpha > \eta^x/\eta^c$ for Assumption 2 to be valid.

Parameter α is referred to as the degree of consumers' preferences for quasi-luxury goods x, or simply the quasi-luxury preference since the intratemporal MRS g_x/g_c in steady state, given by $\overline{\text{MRS}}\,(\bar{c}, \bar{x}; \alpha)$:

$$\overline{\text{MRS}}\,(\bar{c}, \bar{x}; \alpha) = \frac{u_x\,(\bar{x}; \alpha) - \frac{u(\bar{c},\bar{x};\alpha)}{\delta(\bar{c},\bar{x})}\delta_x\,(\bar{c}, \bar{x})}{u_c\,(\bar{x}; \alpha) - \frac{u(\bar{c},\bar{x};\alpha)}{\delta(\bar{c},\bar{x})}\delta_c\,(\bar{c}, \bar{x})} \tag{10.24}$$

satisfies

$$\overline{\text{MRS}}_\alpha = \left\{\frac{\xi u_\alpha}{\delta} + \left(\frac{u_{x\alpha}}{g_x} - \frac{u_{c\alpha}}{g_c}\right)\right\} \text{MRS} > 0 \tag{10.25}$$

under Assumption 2 and (10.23). An increase in α thus enhances the relative preferences for quasi-luxury good x to c.

By differentiating (10.11) through (10.13) with respect to α, the effect of an increase in the quasi-luxury preference on steady-state wealth holding can be computed as

$$\frac{\partial \bar{a}}{\partial \alpha} = \frac{\xi}{g_x \Psi}\overline{\text{MRS}}_\alpha \tag{10.26}$$

which is positive under Assumption 2. The result can be summarized as follows[22]:

Proposition 3 *An increase in the preference α for quasi-luxury goods, specified by (10.23), increases optimal steady-state wealth holding and hence promotes optimal wealth accumulation.*

Intuitively, the preference for quasi-luxury goods induces a preference for wealth. Figure 10.2 illustrates the property. An increase in α shifts the FF' schedule upward,[23] bringing the steady-state point from point E_0 to E_1. The steady-state consumption of quasi-luxury $\bar{x}$ thus increases and that of quasi-necessity $\bar{c}$ decreases (see Appendix A.1.2). Since the steady-state time preference curve RR' is steeper than the budget lines E_0A_0 and E_1A_1, the shift of the FF' schedule increases steady-state total wealth from OA_0 to OA_1.

Transition dynamics are as follows. An increase in quasi-luxury preference shifts the Engel curve upwards. Consumers instantly increase quasi-luxury consumption

[22]It might be somewhat controversial to consider an unanticipated α-shock in the context of a *perfect foresight* model, although it is conventional practice in the literature (e.g., Turnovsky 2000). In Remark 3 below, I shall give a cross-sectional reinterpretation of the analysis.

[23]From Eq. (10.24), the shift of the FF' schedule caused by an increase in α is given by

$$\left.\frac{d\bar{x}}{d\alpha}\right|_{FF'} = -\frac{\overline{\text{MRS}}_\alpha}{\overline{\text{MRS}}_x} = -\frac{\overline{\text{MRS}}_\alpha}{\left(\xi + \delta g_{xx}/g_x^2\right)\left(p g_x/\delta\right)},$$

which is positive under Assumption 3.

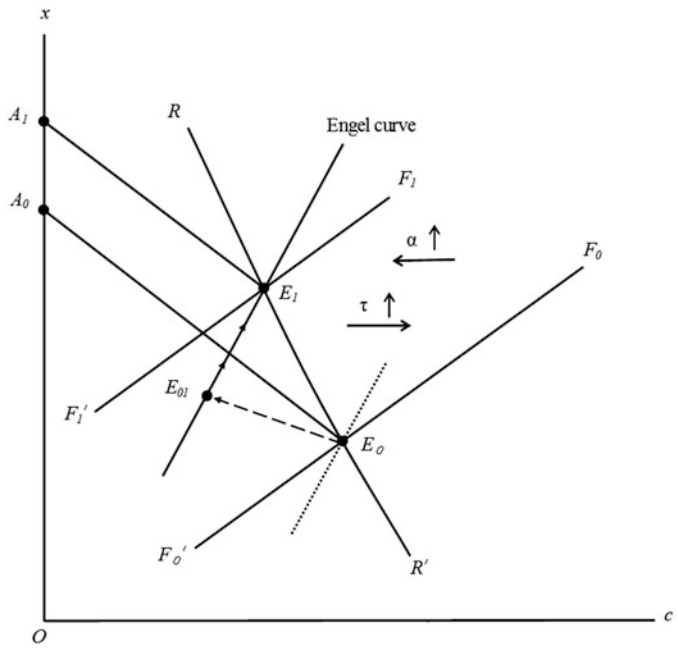


Fig. 10.2 Increases in the quasi-luxury preferences and the quasi-luxury tax

$x(0)$ and decrease quasi-necessity consumption $c(0)$. However, consumers become more patient instantly with respect to either good since (10.19) and (10.20) imply

$$\frac{\partial \rho^x(0)}{\partial \alpha} = -\frac{g_c g_{xx}\omega}{(g_{xx} + p^2 g_{cc})}\left\{\frac{g_{cc}(r-\omega)}{g_c^2} - \xi\right\}\frac{\partial \bar{a}}{\partial \alpha} < 0$$

$$\frac{\partial \rho^c(0)}{\partial \alpha} = -\frac{g_x^2 g_{cc}\omega}{g_c(g_{xx} + p^2 g_{cc})}\left\{\frac{g_{xx}(r-\omega)}{g_x^2} + \xi\right\}\frac{\partial \bar{a}}{\partial \alpha} < 0$$

so that, as illustrated by the discrete jump from point E_0 to E_{01} in Fig. 10.2, the instantaneous increase in $x(0)$ falls short of the long run increase in $\bar{x}$ whereas $c(0)$ decreases more than $\bar{c}$. In the interim run, as wealth is monotonically accumulated, both consumptions gradually increase up to the levels at point E_1.

Remark 3 The comparative dynamics could be taken as comparison between two consumers who differ with respect to α. Suppose that points E_0 and E_{01} represent optimal consumption baskets at time zero for two consumers with a low α and a high α, respectively. They hold the same initial wealth a_0. In the long run, the high-α agent consumes basket E_1 and holds larger wealth than the low-α consumer whose steady-state consumption is given by point E_0. At each point in time, the high-α consumer enjoys more quasi-luxury consumption and less quasi-necessity than the low-α consumer. Note that interpersonal difference in optimal quasi-luxury consumption is larger in the long run than in the short run, whereas the opposite is true for optimal quasi-necessity consumption. As the high-α consumer accumulates wealth, interpersonal difference in quasi-luxury consumption expands over time whereas that in quasi-necessity shrinks.

Remark 4 Even if consumer i is less patient than consumer j in that $\delta^i(c, x) > \delta^j(c, x)\ \forall\ (c, x)$, the optimal steady-state wealth holding for the less-patient consumer i is *larger* than that for the more-patient consumer j if consumer i's quasi-luxury preference is sufficiently larger than j's. To see this, introduce an impatience parameter β to specify the discount rate function as $\delta(c, x, \beta)$, where $\delta_\beta > 0, \delta_{c\beta} = \delta_{x\beta} = 0$. With increasing impatience, an increase in β, i.e., an exogenous increase in steady-state impatience, necessarily reduces steady-state wealth holding $\bar{a}$, as shown by Appendix A.1.2.[24] When two parameters α and β increase at the same time, the effect on $\bar{a}$ depends on the relative magnitudes of the positive effect of the α-increase and the negative one of the β-increase. If the α-increase is sufficiently large, $\bar{a}$ increases. As in Remark 3, take this as a comparison between two consumers with different (α, β) to see that: the less patient consumer with respect to the discount rate function can hold more optimal steady-state wealth than the more patient one if the less patient one has a sufficiently big α.[25]

3 Extensions to General Equilibrium Frameworks

The previous section has examined the relation between the quasi-luxury preference and the optimal wealth accumulation for given market variables r and p. In this section, the analysis is extended into two simple general equilibrium models, a neoclassical production-economy model and a two-country world economy model, where r and p are endogenously determined.

[24]An increase in β leads to a downward shift of the RR' schedule in Fig. 10.2 on one hand. On the other hand, since $u < 0$, the β-increase enlarges ceteris paribus the utility index, thereby shifting the FF' schedule upward. Although the two shifts affect $\bar{a}$ in the opposite directions, the downward shift of RR' is dominant in the sense that $\bar{a}$ is necessarily reduced.

[25]An increase in quasi-luxury preference α (i.e., an upward shift of the FF' schedule) and a decrease in steady-state impatience parameter β (i.e., an upward shift of the RR' schedule) commonly increase steady-state wealth and quasi-luxury consumption (see Appendix A.1.2). These two preference shifts could be distinguished by checking the effects on steady-state consumption of quasi-necessities: an increase in α decreases $\bar{c}$ whereas a decrease in β increases $\bar{c}$.

3.1 A Production Economy

Suppose that two goods are produced using capital and labor. The production functions in the two sectors are given by the usual linearly-homogeneous, concave functions $F^i\left(K^i, L^i\right)$ $(i = c, x)$, where K^i and L^i are the capital and labor employed by sector i, respectively. Labor is supplied inelastically. The total amount of labor L is constant. The total capital stock K accumulates from savings. To avoid complexities due to inter-sectoral differences in factor intensity, it is assumed that the two production functions $F^i\left(K^i, L^i\right)$ are similar in that the production functions satisfy: $F^c\left(K^c, L^c\right) = BF\left(K^c, L^c\right)$ and $F^x\left(K^x, L^x\right) = F\left(K^x, L^x\right)$, where F is a linearly homogeneous concave function. Letting $f\left(k^i\right)$ be $F\left(K^i, L^i\right)/L^i$ with k^i denoting K^i/L^i, profit maximization yields:

$$\frac{w}{r} = \frac{f\left(k^c\right) - k^c f'\left(k^c\right)}{f'\left(k^c\right)} = \frac{f\left(k^x\right) - k^x f'\left(k^x\right)}{f'\left(k^x\right)}$$

$$w = B\left\{f\left(k^c\right) - k^c f'\left(k^c\right)\right\} = p\left\{f\left(k^x\right) - k^x f'\left(k^x\right)\right\}$$

where w and r are the wage rate and the capital rent, respectively. The first equation implies

$$k^c = k^x = k$$

where $k \equiv K/L$ represents the aggregate capital-labor ratio. The above maximum profit condition thus implies

$$p = B \text{ and } r = Bf'(k) \tag{10.27}$$

The relative price p is fixed by the productivity factor B.

The demand side is the same as in the previous section except that total wealth a for the representative agent is specified explicitly as the sum of the per capita capital stock k as nonhuman wealth and the present value of the wage income flow as human wealth. The resultant optimality conditions are essentially the same.

The market-clearing conditions depend on which good is accumulated as capital goods. However, if investment $\dot{k}$ is always measured in units of the numeraire good c regardless of whether c or x is used for investment, capital accumulation is generated by the same aggregate equation in either case,

$$\dot{k} = Bf(k) - c - Bx \tag{10.28}$$

where (10.27) is substituted.[26]

[26] If good c is used for both investment and consumption, whereas good x is only used for consumption, then the equilibrium conditions for the two goods markets can be written as

The equilibrium dynamics are obtained by combining the supply side, represented by (10.27) and (10.28), the demand side, described by (10.3), (10.6), (10.7), and the transversality condition. As proven in Appendix A.2.1, the equilibrium dynamics are uniquely given by a saddle time-path governed by a negative root. Since the transition dynamics are monotonic, I focus on the steady-state effect of an increase in quasi-luxury preference α defined by (10.23). The steady-state equilibrium $\left(\bar{c}, \bar{x}, \bar{k}\right)$ is determined by

$$\delta\left(\bar{c}, \bar{x}\right) = Bf'(\bar{k}) \tag{10.29}$$

$$\frac{g_x\left(\bar{x}, u\left(\bar{c}, \bar{x}; \alpha\right) / \delta\left(\bar{c}, \bar{x}\right)\right)}{g_c\left(\bar{c}, u\left(\bar{c}, \bar{x}; \alpha\right) / \delta\left(\bar{c}, \bar{x}\right)\right)} = B \tag{10.30}$$

$$Bf\left(\bar{k}\right) = \bar{c} + B\bar{x} \tag{10.31}$$

Substituting (10.31) into (10.29) yields

$$\delta\left(\bar{c}, \bar{x}\right) = Bf'\left(f^{-1}\left(\frac{\bar{c} + B\bar{x}}{B}\right)\right) \tag{10.32}$$

The steady-state consumption basket $(\bar{c}, \bar{x})$ is determined by (10.30) and (10.32). Capital stock $\bar{k}$ is then given by (10.31). These relations are apparently the same as depicted in Fig. 10.1 if schedule RR' is read as representing (10.32),[27] the contract curve FF' as (10.30), and the long-run budget constraint AA' as (10.31). In fact, just as schedule RR' is steeper than AA' in Sect. 2, schedule (10.32) has a greater slope than (10.31) since

$$\begin{aligned} \left|\left.\frac{\mathrm{d}x}{\mathrm{d}c}\right|_{(10.32)}\right| &= \frac{1}{B} + \frac{\xi g_c}{\delta_x - Bf''/f'} \\ &> \frac{1}{B} \\ &= \left|\left.\frac{\mathrm{d}x}{\mathrm{d}c}\right|_{(10.31)}\right| \end{aligned} \tag{10.33}$$

$$Bl^c f\left(k\right) = c + \dot{k} \text{and} B\left(1 - l^c\right) f\left(k\right) = Bx,$$

where $l^c \equiv L^c/L$ represents the proportion of labor employed in sector c. When good x is used for investment purposes, the $\dot{k}$ term moves to the right-hand side of the second equation. It is obvious that both cases yield (10.28).

[27]To be precise, schedule (10.32) differs from the original RR' curve since schedule (10.32) gives the points of intersection between (10.29) and (10.31) that are obtained by changing $\bar{k}$ parametrically.

Therefore, from the same discussion as for Fig. 10.2, an increase in quasi-luxury preference α increases the steady-state capital stock $\bar{k}$: it shifts the contract curve FF' schedule upward, thereby increasing total expenditures $\bar{c} + B\bar{x}$ from A_0 to A_1 and raising the steady-state capital stock. Actually differentiating Eqs. (10.30)–(10.32) with respect to α yields

$$\frac{\partial \bar{k}}{\partial \alpha} = \frac{\xi g_c \overline{\mathrm{MRS}}_\alpha}{g_c g_x \ \Psi|_{r=Bf'} + Bf'' \left(\frac{g_{xx}}{g_x} + \frac{B g_{cc}}{g_c} \right)} > 0 \tag{10.34}$$

where, from Eq. (10.27), r's in $\overline{\mathrm{MRS}}_\alpha$ and Ψ are replaced by Bf'.

Proposition 4 *The greater is the preference α for quasi-luxury goods in (10.23), the more steady-state capital is accumulated.*

As an important policy implication of Proposition 4, taxation on quasi-luxury goods may well harm capital accumulation. Consider a tax τ on quasi-luxury consumption x, assuming that the tax revenue τx is paid back to households in a lump-sum manner. Then, in steady state, the first-order condition (10.30) becomes

$$\frac{g_x \left(\bar{x}, u \left(\bar{c}, \bar{x}; \alpha \right) / \delta \left(\bar{c}, \bar{x} \right) \right)}{g_c \left(\bar{c}, u \left(\bar{c}, \bar{x}; \alpha \right) / \delta \left(\bar{c}, \bar{x} \right) \right)} = (1 + \tau) \, B \tag{10.35}$$

Any other steady-state conditions are unchanged. Therefore, when the initial tax equals zero, an increase in quasi-luxury tax τ has qualitatively the same effect as a decrease in quasi-luxury preference α: it decreases the steady-state capital stock $\bar{k}$. Figure 10.2 can be reinterpreted as illustrating this. An increase in the quasi-luxury tax shifts downward the contract curve FF', defined by (10.30), thereby bringing the steady-state point from E_1 to E_0. Formally, from (10.31), (10.32), and (10.35), I obtain

$$\left. \frac{\partial \bar{k}}{\partial \tau} \right|_{\tau=0} = - \frac{\xi g_x}{g_c g_x \ \Psi|_{r=Bf'} + Bf'' \left(\frac{g_{xx}}{g_x} + \frac{B g_{cc}}{g_c} \right)} < 0 \tag{10.36}$$

Corollary 1 *Taxation on quasi-luxury goods from no initial distortion decreases the steady-state capital stock.*

Remark 5 From Corollary 1, introducing a tax on a luxury good necessarily decreases the steady-state capital stock as long as the good is a quasi-luxury. In this sense, luxury taxes are likely to harm capital formation, as Sir Dudley North pointed out more than three centuries ago (see the epigraph).

Remark 6 In the above discussion, there is no initial distortion. With a non-zero initial τ, I can derive from (10.31), (10.32), and (10.35)

$$\left.\frac{\partial \bar{k}}{\partial \tau}\right|_{\tau \neq 0} = -\frac{g_x\left(\xi - \tau\frac{\delta_x}{g_x}\right)}{\Omega}$$

$$\Omega = (1+\tau)^2\left\{g_c g_x \left.\Psi\right|_{r=Bf'} + Bf''\left(\frac{g_{xx}}{g_x} + \frac{Bg_{cc}}{g_c} + \tau\frac{\xi g_c}{f'}\right)\right\}$$

where $\Omega > 0$ as shown in Appendix A.2.2, implying that $\partial\bar{k}/\partial\tau\big|_{\tau\neq 0} \lesseqgtr 0$ as $\xi \gtreqless \tau\delta_x/g_x$. Since from (10.4) and (10.35) $\xi \gtreqless \tau\delta_x/g_x$ as $B \gtreqless \delta_x/\delta_c$, this implies that $\partial\bar{k}/\partial\tau\big|_{\tau\neq 0} > 0$ when $B < \delta_x/\delta_c$.[28] To understand this, note that the relative price of x, i.e., the private marginal rate of transformation (MRT) of x, as perceived by the representative consumer, equals $B(1+\tau)$, whereas, due to the lump-sum transfers resulting from the rebatement of the taxes, the economy's or social MRT is given by B. From the definition of quasi-luxury, the equilibrium private MRT of good x (i.e., $B(1+\tau)$), which equals MRS, is larger than δ_x/δ_c, although the social MRT of x (i.e., B) can be smaller than δ_x/δ_c due to distortionary taxation, in which case good x could be taken as *socially* a quasi-necessity good in equilibrium. Intuitively, when $B < \delta_x/\delta_c$, taxation on a quasi-luxury increases the steady-state capital stock since the good is *socially* a quasi-*necessity*. Taxation on a luxury good necessarily reduces the steady-state capital stock when the luxury is *socially* a quasi-*luxury*.

3.2 A Two-Country World Economy

Let us finally extend the discussion to a two-country world economy model. Consider a world economy composed of home and foreign countries, H and F. Foreign-country variables are denoted by asterisks. Any production activities are neglected for simplicity. The representative agents in both countries are endowed with $y^{c(*)}$ of quasi-necessity goods and $y^{x(*)}$ of quasi-luxury goods. The consumers can freely trade these two goods at the world price p, and bonds b at the world interest rate r. The flow budget constraint is given by

$$\dot{b}(t) = r(t)\,b(t) + y^c + p(t)\,y^x - c(t) - p(t)\,x(t) \tag{10.37}$$

It is equivalent to the flow budget constraint (10.5) in Sect. 2 since total wealth a is defined as the sum of nonhuman wealth and human wealth:

[28]Linear approximations of schedule (10.32), as assumed for schedule RR' in Fig. 10.2, are no longer sufficient to discuss the effects of taxation under a non-zero τ. It can be shown that $\partial\bar{k}/\partial\tau$ can only switch its sign if schedule (10.32) has a strictly convex segment so that the schedule becomes flatter than the long-run budget constraint AA' as c increases.

$$a(t) = b(t) + \int_t^{\infty} \{y^c + p(s)\, y^x\} \exp\left(-\int_t^s r(\tau)\, \mathrm{d}\tau\right) \mathrm{d}s$$

The market equilibrium conditions for the two goods and bonds are expressed as:

$$c(t) + c^*(t) = Y\left(\equiv y^c + y^{c*}\right)$$

$$x(t) + x^*(t) = X\left(\equiv y^x + y^{x*}\right)$$

$$b(t) + b^*(t) = 0$$

By the Walras law, when the two-good markets are in equilibrium for all $t \geq 0$ and when the international bond markets are in equilibrium at $t = 0$, then the international bond markets are in equilibrium for $t > 0$, too. Given the time paths of $p(t)$ and $r(t)$, the optimal consumption path $\{b(t), c(t), x(t), \phi(t)\}_{t=0}^{\infty}$ for the representative agent in country H is determined by (10.3), (10.6), (10.7), and (10.37), and that for country F by the corresponding equations with respect to $\{b^*(t), c^*(t), x^*(t), \phi^*(t)\}$. The equilibrium price path $\{p(t), r(t)\}_{t=0}^{\infty}$ is determined such that these optimal consumption paths satisfy the above market-equilibrium conditions.

I examine the effects of a permanent increase in country H's preference α for quasi-luxury on international wealth distribution. In so doing, consider an equilibrium when the felicity functions $u^{(*)}\left(c^{(*)}, x^{(*)}; \alpha^{(*)}\right)$ and the discounting functions $\delta^{(*)}\left(c^{(*)}, x^{(*)}\right)$ are internationally identical:

$$u(\bullet, \bullet; \bullet) = u^*(\bullet, \bullet; \bullet), \delta(\bullet, \bullet) = \delta^*(\bullet, \bullet), \alpha = \alpha^* \tag{10.38}$$

where quasi-luxury preferences $\alpha^{(*)}$ are introduced as in (10.23). The steady-state equilibrium is then determined from the following equations:

$$\delta(\bar{c}, \bar{x}) = \delta\left(\bar{c}^*, \bar{x}^*\right) = \bar{r} \tag{10.39}$$

$$\frac{g_x(\bar{x}, u(\bar{c}, \bar{x}; \alpha) / \delta(\bar{c}, \bar{x}))}{g_c(\bar{c}, u(\bar{c}, \bar{x}; \alpha) / \delta(\bar{c}, \bar{x}))} = \frac{g_x(\bar{x}^*, u(\bar{c}^*, \bar{x}^*; \alpha^*) / \delta(\bar{c}^*, \bar{x}^*))}{g_c(\bar{c}^*, u(\bar{c}^*, \bar{x}^*; \alpha^*) / \delta(\bar{c}^*, \bar{x}^*))} = \bar{p} \tag{10.40}$$

$$\bar{r}\bar{a}^{(*)} = \bar{r}\bar{b}^{(*)} + y^{c(*)} + \bar{p} y^{x(*)} = \bar{c}^{(*)} + \bar{p}\bar{x}^{(*)} \tag{10.41}$$

$$\bar{\phi}^{(*)} = \frac{u\left(\bar{c}^{(*)}, \bar{x}^{(*)}; \alpha^{(*)}\right)}{\delta\left(\bar{c}^{(*)}, \bar{x}^{(*)}\right)}$$

$$\bar{c} + \bar{c}^* = Y, \bar{x} + \bar{x}^* = X \tag{10.42}$$

Since the steady-state consumption baskets $\left(\bar{c}^{(*)}, \bar{x}^{(*)}\right)$ are determined from (10.39) and (10.40), the identical preference structure (10.38) produces identical consumption baskets $(\bar{c}, \bar{x}) = (\bar{c}^*, \bar{x}^*)$, and hence

$$\bar{c} = \bar{c}^* = Y/2, \bar{x} = \bar{x}^* = X/2$$

$$\bar{\phi} = \bar{\phi}^* = u\,(Y/2, X/2; \alpha)\,/\delta\,(Y/2, X/2)$$

$$\bar{r} = \delta\,(Y/2, X/2)\,, \bar{p} = \frac{g_x\,(X/2, u\,(Y/2, X/2; \alpha)\,/\delta\,(Y/2, X/2))}{g_c\,(Y/2, u\,(Y/2, X/2; \alpha)\,/\delta\,(Y/2, X/2))}$$

$$\bar{a} = \bar{a}^* = \frac{Y + \bar{p}X}{2\bar{r}}$$

Net foreign assets $\bar{b}^{(*)}$ are obtained by substituting these solutions to (10.41). Around the steady-state point, the equilibrium dynamics are uniquely given by a saddle time-path governed by a negative root (see Appendix A.3.1).

Starting from the steady-state equilibrium, suppose that country H's preference α for quasi-luxury goods increases. Then, as Appendix A.3.2 shows, it can be verified that

$$\frac{\partial \bar{a}}{\partial \alpha} = \left(\frac{\xi}{g_x \Psi} + \frac{X}{2\bar{r}}\right) \frac{\overline{\mathrm{MRS}}_\alpha}{2}\ (> 0) \tag{10.43}$$

$$\frac{\partial \bar{a}^*}{\partial \alpha} = \left(-\frac{\xi}{g_x \Psi} + \frac{X}{2\bar{r}}\right) \frac{\overline{\mathrm{MRS}}_\alpha}{2} \left(< \frac{\partial \bar{a}}{\partial \alpha}\right) \tag{10.44}$$

where the partial derivatives are evaluated at the symmetry steady-state equilibrium. Since $\bar{a}$ equals $\bar{a}^*$ in the initial steady-state equilibrium, these equations imply that $\bar{a}$ is larger than $\bar{a}^*$ after α increases.

Proposition 5 *Suppose that two countries have the same preferences except for quasi-luxury preferences, α and α^*. Then, the country with a stronger preference for quasi-luxury goods holds more total wealth in steady state than the other country with a weaker preference.*

The effects on net foreign assets $\bar{b}, \bar{b}^*$ are derived in Appendix A.3.2 from Eqs. (10.39) through (10.41) as

$$\frac{\partial \bar{b}}{\partial \alpha} = -\frac{\partial \bar{b}^*}{\partial \alpha} = \left\{\xi - \left(y_x - \frac{X}{2}\right)\frac{g_x \Psi}{\bar{r}}\right\} \frac{\overline{\mathrm{MRS}}_\alpha}{2 g_x \Psi} \tag{10.45}$$

$$\gtreqless 0 \Leftrightarrow \xi \gtreqless \left(y_x - \frac{X}{2}\right)\frac{g_x \Psi}{\bar{r}}$$

Whether or not an increase in α enlarges country H's net foreign assets depends on two effects represented by ξ and $(y_x - X/2)\, g_x \Psi / \bar{r}$. The term ξ represents the wealth-preference effect induced by the increase in the quasi-luxury preference, which has been focused on in this chapter. The second term captures the terms-of-trade effect: an increase in α raises p, thereby causing a terms-of-trade improvement or deterioration for country H as the country exports or imports x (i.e., as $y_x - X/2$ is positive or negative). Suppose that country H exports x: $y_x - X/2 > 0$. Then, an improvement in country H's terms of trade raises her real income. The country thus requires less external assets $\bar{b}$ than before to maintain the initial living standard. Only when the wealth-preference effect dominates the terms-of-trade effect, $\bar{b}$ increases and vice versa. When country H imports x (i.e., $y_x - X/2 < 0$), the preference shift is accompanied by a terms-of-trade deterioration, which, jointly with the wealth-preference effect, increases $\bar{b}$.[29]

A corollary of this result is that levying tariffs on quasi-luxuries necessarily *worsens* the current account. To show this, assume that country H imports quasi-luxury goods x (i.e., $y_x - X/2 < 0$). The country imposes an import tariff π from no initial distortion, transferring the tariff revenue to the residents in a lump-sum manner. Since the marginal condition for country H becomes $g_x/g_c = (1+\pi)\, p$, steady-state condition (10.40) is replaced by

$$\left(\frac{1}{1+\pi}\right) \frac{g_x\left(\bar{x}, u\left(\bar{c}, \bar{x}; \alpha\right)/\delta\left(\bar{c}, \bar{x}\right)\right)}{g_c\left(\bar{c}, u\left(\bar{c}, \bar{x}; \alpha\right)/\delta\left(\bar{c}, \bar{x}\right)\right)} = \frac{g_x\left(\bar{x}^*, u\left(\bar{c}^*, \bar{x}^*; \alpha^*\right)/\delta\left(\bar{c}^*, \bar{x}^*\right)\right)}{g_c\left(\bar{c}^*, u\left(\bar{c}^*, \bar{x}^*; \alpha^*\right)/\delta\left(\bar{c}^*, \bar{x}^*\right)\right)} = \bar{p}$$

with the other conditions being unchanged. As seen by differentiating totally the above equation, the effect of the import tariff on net foreign assets is obtained by replacing $\overline{\mathrm{MRS}}_\alpha$ on the right hand side of (10.45) by $-\bar{p}$ as

$$\left.\frac{\partial \bar{b}}{\partial \pi}\right|_{\pi=0} = -\left\{\xi - \left(y_x - \frac{X}{2}\right)\frac{g_x \Psi}{\bar{r}}\right\} \frac{\bar{p}}{2 g_x \Psi} < 0$$

Corollary 2 *Levying tariffs on quasi-luxury imports from no initial distortion decreases steady-state net foreign assets and hence worsens the current account in transition.*[30]

Remark 7 Corollary 2 implies that, in contrast to what is commonly believed, luxury import tariffs are likely to worsen the balance of payments.

---

[29]By using a recursive-preference model of the Uzawa type, Obstfeld (1982) shows that a terms-of-trade improvement worsens the current account. Ikeda (2001) gives a counter-example in a small country model of weakly nonseparable preferences. The above discussion is its two-country extension.

[30]When the import tariff rate is initially non-zero, the analysis of the two-country equilibrium is too complicated since the symmetric structure of the model is broken. See Ikeda (2003), which examines the effects of import tariffs under initial distortions by using a small country model with weakly nonseparable preferences.

As in Remark 4 in Sect. 2, by introducing impatience parameter β to the two-country model, it is easy to establish that, due to the stronger preference for quasi-luxury goods, the less patient country can hold more wealth than the more patient.[31]

Corollary 3 *Even when there is an international difference in the steady-state time preference (utility-discounting functions), the less-patient country can have more wealth in steady state than the more-patient one if the less-patient one has sufficiently stronger preferences for quasi-luxury.*

4 Conclusions

There is an old idea that certain goods are related to wealth accumulation. I have provided a model to formalize the idea by incorporating weakly non-separable preferences. The model helps to understand luxury consumption from a dynamic viewpoint. The phenomenon that wealthier agents consume more luxuries can be described by stating that the consumers are wealthier because they prefer luxuries (quasi-luxuries). A new insight is that wealth accumulation should reflect consumers' preferences for various kinds of goods as well as for time.

There are several ways in which the above analysis can be extended. First, it should be extended to the case of more than two goods. Quasi-luxury goods are defined here in terms of the MRS between two goods. With more than two goods, some other devices would be required. Second, empirical testing of our model should be conducted. The model could be hypothesized by the property that differences in commodity-specific time preferences between luxury and necessity goods depend on wealth accumulation.

Acknowledgements I am grateful to four anonymous referees for their useful comments. I especially thank one of the referees, who gave me detailed referee reports of totally more than 50 pages, for his or her beneficial comments. I also give thanks to D. Dasgupta, K. Futagami, C.Y. Horioka, K. Mino, D. Nahm, F. Ohtake, T. Ono, Y. Ono, K. Ryu, Y. Tsutsui, M. Yano and the participants at the 51th IAEC Meeting, Athens, 2001; the Canadian Economic Association 35th Annual Meeting, Montreal, 2001; the 2001 Far Eastern Meeting of the Econometric Society, Kobe; the Macro Research Seminar; and the seminars at Keio, Osaka, and the Seoul National Universities for helpful discussions on the earlier versions of the chapter. A part of this research is financially supported by The Matsushita International Foundation (No.00-049) and Grants-in-Aid for Scientific Research C (No.15530121).

[31]For the proof, see Appendix A.3.2, which analyzes the effects of shifts in quasi-luxury preferences (α, α^*) and impatience parameters (β, β^*) of the two countries.

Appendix

A.1 Properties of the Optimal Solution in Sect. 2.2

A.1.1 Dynamics

To obtain an autonomous dynamic system with respect to $l \equiv (c, x, a)$, solve (10.6) for ϕ in the form

$$\phi = h(c, x, p) \equiv \frac{pu_c - u_x}{p\delta_c - \delta_x} \tag{10.46}$$

and substitute it for ϕ in (10.7) and in the corresponding Euler equation for x. Then, from (10.8) and (10.9), the optimal dynamics can be expressed in the state-control space by

$$\dot{c} = -\frac{g_c(c, h(c, x, p))}{g_{cc}(c, h(c, x, p))}\left[r - \left(\delta(c, x) - \frac{\delta_c(c)}{g_c(c, h(c, x, p))} g(c, x, h(c, x, p))\right)\right]$$

$$\dot{x} = -\frac{g_x(x, h(c, x, p))}{g_{xx}(x, h(c, x, p))}\left[r - \left(\delta(c, x) - \frac{\delta_x(x)}{g_x(x, h(c, x, p))} g(c, x, h(c, x, p))\right)\right]$$

$$\dot{a} = ra - c - px$$

Linearizing the system around the steady-state point yields $\dot{l} = A\hat{l}$, where

$$A \equiv \begin{pmatrix} \frac{r\delta_c}{\xi g_c} & -\frac{g_c g_x}{g_{cc}}\left(\xi + \frac{r\delta_c g_{xx}}{\xi g_x^2 g_c}\right) & 0 \\ \frac{g_c g_x}{g_{xx}}\left(\xi + \frac{r\delta_x g_{cc}}{\xi g_c^2 g_x}\right) & -\frac{r\delta_x}{\xi g_x} & 0 \\ -1 & -\frac{g_x}{g_c} & r \end{pmatrix} \tag{10.47}$$

The trace and determinant of A are computed as

$$\mathrm{trace}A = 2r > 0$$

$$\det A = -\frac{r g_c^2 g_x^2}{g_{cc} g_{xx}} \Psi < 0$$

which imply that the steady-state point is saddle-point stable.

A.1.2 The Effects of Preference Shifts in Sect. 2.4

Introduce quasi-luxury preference α and impatience parameter β into the steady-state equilibrium conditions (10.11) through (10.13), where $\delta_\beta > 0, \delta_{c\beta} = \delta_{x\beta} = 0$ as in Remark 4. Differentiating totally the result yields

$$\delta_c \mathrm{d}\bar{c} + \delta_x \mathrm{d}\bar{x} = -\delta_\beta \mathrm{d}\beta \tag{10.48}$$

$$\overline{\mathrm{MRS}}_c \mathrm{d}\bar{c} + \overline{\mathrm{MRS}}_x \mathrm{d}\bar{x} = -\overline{\mathrm{MRS}}_\alpha \mathrm{d}\alpha - \overline{\mathrm{MRS}}_\beta \mathrm{d}\beta \tag{10.49}$$

$$\mathrm{d}\bar{a} = \frac{\mathrm{d}\bar{c} + p\mathrm{d}\bar{x}}{r} \tag{10.50}$$

where

$$\overline{\mathrm{MRS}}_c = p\left(-\frac{g_{cc}}{g_c} + \xi\frac{g_c}{r}\right) > 0; \overline{\mathrm{MRS}}_x = \frac{g_x^2}{rg_c}\left(\frac{rg_{xx}}{g_x^2} + \xi\right) < 0 \tag{10.51}$$

$$\overline{\mathrm{MRS}}_\beta = -\frac{p\xi\bar{\phi}\delta_\beta}{r} > 0 \tag{10.52}$$

(a) *An increase in α*. From (10.48) and (10.49), the effects of an increase in α on $\bar{c}$ and $\bar{x}$ are obtained as

$$\frac{\partial\bar{c}}{\partial\alpha} = -\frac{r\delta_x}{g_x^2\Psi}\overline{\mathrm{MRS}}_\alpha < 0$$

$$\frac{\partial\bar{x}}{\partial\alpha} = \frac{r\delta_c}{g_x^2\Psi}\overline{\mathrm{MRS}}_\alpha > 0$$

Substituting the above equations into (10.50) yields (10.26):

$$\frac{\partial\bar{a}}{\partial\alpha} = \frac{\xi}{g_x\Psi}\overline{\mathrm{MRS}}_\alpha > 0$$

By substituting (10.26) and the above results into (10.17), the impact effects on $c\,(0)$ and $x\,(0)$ are given by

$$\frac{\partial c\,(0)}{\partial\alpha} = \frac{g_x^2\Psi + \xi g_{xx}\omega}{(g_{xx} + p^2 g_{cc})\,\Psi}\overline{\mathrm{MRS}}_\alpha < 0$$

$$\frac{\partial x\,(0)}{\partial\alpha} = -\frac{g_c^2\Psi - \xi g_{cc}\omega}{(g_{xx} + p^2 g_{cc})\,g_c\Psi}\overline{\mathrm{MRS}}_\alpha$$

which can be shown to be positive.

(b) *An increase in β*. From (10.48) and (10.49), the effect of an increase in β on the steady-state consumption basket can be obtained as

$$\frac{\partial \bar{c}}{\partial \beta} = \frac{r\left(\overline{\mathrm{MRS}}_x \delta_\beta - \overline{\mathrm{MRS}}_\beta \delta_x\right)}{g_x^2 \Psi} < 0$$

$$\frac{\partial \bar{x}}{\partial \beta} = -\frac{r\left(\overline{\mathrm{MRS}}_c \delta_\beta - \overline{\mathrm{MRS}}_\beta \delta_c\right)}{g_x^2 \Psi} < 0.$$

The effect on $\bar{a}$ is thus derived from (10.50) as

$$\frac{\partial \bar{a}}{\partial \beta} = \frac{1}{r}\left(\frac{\partial \bar{c}}{\partial \beta} + p\frac{\partial \bar{x}}{\partial \beta}\right) < 0 \tag{10.53}$$

Consider increases in quasi-luxury preference α and impatience parameter β at the same time. The effect on $\bar{a}$ is given by combining (10.26) and (10.53) as

$$\mathrm{d}\bar{a} = \frac{\partial \bar{a}}{\partial \alpha}\mathrm{d}\alpha + \frac{\partial \bar{a}}{\partial \beta}\mathrm{d}\beta$$

where $\partial \bar{a}/\partial \alpha > 0, \partial \bar{a}/\partial \beta < 0$, implying that $\mathrm{d}\bar{a}$ is positive when $\mathrm{d}\alpha$ is sufficiently large relative to $\mathrm{d}\beta$. This result can be applied to a comparison between two consumers with different (α, β)'s in Remark 4.

A.2 Solutions to the Production-Economy Model in Sect. 3.1

A.2.1 Dynamics

In the production-economy model in Sect. 3.1, the equilibrium dynamics for (c, k, p, r, x, ϕ) are generated by (10.3), (10.6), (10.7), (10.27), (10.28) and the transversality condition. The system can be reduced to the autonomous system,

$$\dot{c} = -\frac{g_c(c, h(c, x, p))}{g_{cc}(c, h(c, x, p))} \left[Bf'(k) - \left(\delta(c, x) - \frac{\delta_c(c)}{g_c(c, h(c, x, p))} g(c, x, h(c, x, p))\right)\right]$$

$$\dot{x} = -\frac{g_x(x, h(c, x, p))}{g_{xx}(x, h(c, x, p))} \left[Bf'(k) - \left(\delta(c, x) - \frac{\delta_x(x)}{g_x(c, h(c, x, p))} g(c, x, h(c, x, p))\right)\right]$$

$$\dot{k} = Bf(k) - c - Bx$$

Let m denote $(c, \phi, k)'$. The autonomous system can be linearized as $\dot{m} = G\hat{m}$, where

$$G = \begin{pmatrix} \frac{\delta\delta_c}{\xi g_c} & -\frac{g_c g_x}{g_{cc}}\left(\xi + \frac{\delta\delta_c g_{xx}}{\xi g_x^2 g_c}\right) & -\frac{g_c}{g_{cc}}Bf'' \\ \frac{g_c g_x}{g_{xx}}\left(\xi + \frac{\delta\delta_x g_{cc}}{\xi g_c^2 g_x}\right) & -\frac{\delta\delta_x}{\xi g_x} & -\frac{g_x}{g_{xx}}Bf'' \\ -1 & -B & Bf' \end{pmatrix}$$

The trace of G equals $2Bf' > 0$. The determinant can be computed as

$$\det G = -\frac{g_c^2 g_x^2 Bf'}{g_{cc} g_{xx}}\Psi - \frac{B^2 g_c f' f''}{g_{cc}}\left(1 + B^2 \frac{g_{cc}}{g_{xx}}\right) < 0$$

implying that the equilibrium dynamics are saddle-point stable. It can be shown that the two positive roots are inconsistent with the transversality condition.

A.2.2 The Effects of Quasi-luxury Taxes

Take the total differential of (10.29), (10.35), and (10.31) to obtain

$$\begin{pmatrix} \frac{\delta_c}{\mathrm{MRS}_{\bar{c}}} & \frac{\delta_x}{\mathrm{MRS}_{\bar{x}}} & -Bf'' \\ 1 & B & -Bf' \end{pmatrix}$$

$$\begin{pmatrix} \frac{\delta_c}{\mathrm{MRS}_{\bar{c}}} & \frac{\delta_x}{\mathrm{MRS}_{\bar{x}}} & -Bf'' \\ & & 0 \\ 1 & B & -Bf' \end{pmatrix} \begin{pmatrix} \mathrm{d}\bar{c} \\ \mathrm{d}\bar{x} \\ \mathrm{d}\bar{k} \end{pmatrix} = \begin{pmatrix} 0 \\ B\,\mathrm{d}\tau - \mathrm{MRS}_\alpha \mathrm{d}\alpha \\ 0 \end{pmatrix}$$

After tedious computation, the determinant Λ of this matrix can be obtained as

$$\begin{aligned} \Lambda &= (1+\tau)B\left\{g_c g_x\ \Psi|_{r=Bf'} + Bf''\left(\frac{g_{xx}}{g_x} + \frac{Bg_{cc}}{g_c} + \tau\frac{\xi g_c}{f'}\right)\right\} \\ &= B\Omega/\,(1+\tau) \end{aligned}$$

where Ψ is given by (10.16); and Ω is defined in Remark 6. Ω and hence Λ are strictly positive since Ω can be rewritten as

$$\begin{aligned} \Omega = B(1+\tau)^2 \tfrac{g_c g_x}{\delta\delta_x} &\left\{\tfrac{\delta_x}{g_x}(\delta_x f' - Bf'')\left(\xi - \tfrac{\delta g_{cc}}{g_c^2}\right)\right. \\ &\left. + \tfrac{\delta_x}{g_c}(\delta_c f' - f'')\left(-\tfrac{\delta g_{xx}}{g_x^2} - \xi\right)\right\} \end{aligned}$$

which is strictly positive under Assumption 3.

Using these results, we obtain

$$\mathrm{d}\bar{c} = \frac{(1+\tau)\left(\delta_x f' - Bf''\right)\left(B\mathrm{d}\tau - \overline{\mathrm{MRS}}_\alpha \mathrm{d}\alpha\right)}{\Omega}$$

$$\mathrm{d}\bar{x} = \frac{(1+\tau)\left(\delta_c f' - f''\right)\left(B\mathrm{d}\tau - \overline{\mathrm{MRS}}_\alpha \mathrm{d}\alpha\right)}{\Omega}$$

$$\mathrm{d}\bar{k} = \frac{(1+\tau)\delta_x\left(\frac{\delta_c}{\delta_x} - \frac{1}{B}\right)\left(B\mathrm{d}\tau - \overline{\mathrm{MRS}}_\alpha \mathrm{d}\alpha\right)}{\Omega} \tag{10.54}$$

Equation (10.54) implies (10.34), (10.36), and the result in Remark 6.

A.3 The Two-Country Equilibrium in Sect. 3.2

A.3.1 Dynamics

Let us derive the equilibrium local dynamics around the steady-state point with the identical preference structure (10.38). Since Y is constant, market equilibrium requires $\dot{c} + \dot{c}^* = 0$. Substituting (10.7) and the corresponding equation for country F into this condition yields

$$c\sigma^c\left(r - \rho^c\right) + c^*\sigma^{c*}\left(r - \rho^{c*}\right) = 0$$

From (10.38), $\bar{c}\sigma^c\left(\bar{c}, \bar{\phi}\right)$ equals $\bar{c}^*\sigma^{c*}\left(\bar{c}^*, \bar{\phi}^*\right)$; function $\rho^c\left(\bullet, \bullet, \bullet\right)$ is identical to function $\rho^{c*}\left(\bullet, \bullet, \bullet\right)$; and $\bar{\rho}^c = \bar{\rho}^{c*} = \bar{r}$. Linearizing the above equation thus yields

$$\left(\hat{r} - \hat{\rho}^c\left(c, x, \phi\right)\right) + \left(\hat{r} - \hat{\rho}^c\left(c^*, x^*, \phi^*\right)\right) = 0$$

where $\hat{\rho}^c\left(c, x, \phi\right) = \hat{\rho}^c_c\hat{c} + \hat{\rho}^c_x\hat{x} + \hat{\rho}^c_\phi\hat{\phi}$, etc. Solve the equation for $\hat{r}$ and substitute (10.46) into the result to obtain

$$\hat{r} = \frac{\hat{\rho}^c\left(c, x, h\left(c, x, p\right)\right) + \hat{\rho}^c\left(Y - c, X - x, h\left(Y - c, X - x, p\right)\right)}{2} \tag{10.55}$$

where $h\left(\bullet, \bullet, \bullet\right) = h^*\left(\bullet, \bullet, \bullet\right)$ is substituted. By computing $\hat{\rho}^c\left(c, x, h\left(c, x, p\right)\right)$ and $\hat{\rho}^c\left(Y - c, X - x, h\left(Y - c, X - x, p\right)\right)$, this can be reduced to

$$\hat{r} = \frac{\bar{r}\delta_c}{\xi g_x}\hat{p} \tag{10.56}$$

Substituting (10.55) and (10.46) successively into the linear approximate of (10.7) yields

$$\dot{c} = -\frac{g_c}{2g_{cc}}\left[\hat{\rho}^c\left(Y-c, X-x, h\left(Y-c, X-x, p\right)\right) - \hat{\rho}^c\left(c, x, h\left(c, x, p\right)\right)\right]$$

which can be computed as

$$\dot{c} = \frac{\bar{r}\delta_c}{\xi g_c}\hat{c} - \frac{g_c g_x}{g_{cc}}\left(\xi + \frac{\bar{r}\delta_c g_{xx}}{\xi g_x^2 g_c}\right)\hat{x} \tag{10.57}$$

In the same way, the law of motion for x is given by

$$\dot{x} = \frac{g_c g_x}{g_{xx}}\left(\xi + \frac{\bar{r}\delta_x g_{cc}}{\xi g_c^2 g_x}\right)\hat{c} - \frac{\bar{r}\delta_x}{\xi g_x}\hat{x} \tag{10.58}$$

To obtain a dynamic equation for p, substitute (10.46) into (10.3), thereby obtaining $\dot{h} = -g\left(c, x, h\left(c, x, p\right)\right)$. It can be linearized as

$$h_c\dot{c} + h_x\dot{x} + h_p\dot{p} = -\left(g_c + g_\phi h_c\right)\hat{c} - \left(g_x + g_\phi h_x\right)\hat{x} - g_\phi h_p\hat{p}$$

Substitute (10.57) and (10.58) into the linearized equation. The resulting equation can be solved for $\dot{p}$ as $\dot{p} = \delta\hat{p}$, which implies

$$\hat{p} = 0 \tag{10.59}$$

for all $t > 0$: the relative price of quasi-luxury good x adjusts immediately after a permanent shock. Applying this to (10.56) in turn yields

$$\hat{r} = 0 \tag{10.60}$$

for all $t > 0$. The equilibrium interest rate always equals the steady-state value.

From (10.59) and (10.60), (10.37) is linearized as

$$\dot{b} = \bar{r}\hat{b} - \hat{c} - \bar{p}\hat{x} \tag{10.61}$$

Equations (10.57), (10.58) and (10.61) are combined as an autonomous dynamic system for $n \equiv (c, x, b)$,

$$\dot{n}(t) = A\hat{n}(t)$$

where A is matrix (10.47) describing the local dynamics for individuals' optimal consumption in Sect. 2. As seen in Sect. 2, A has one positive root and one negative root ω. The resultant equilibrium path is very similar to that obtained in (10.17):

$$\dot{b}(t) = \omega \hat{b}(t), \hat{b}(0) = b_0 - \bar{b}$$

$$\hat{c}(t) = -\hat{c}^*(t) = -\frac{g_x^2}{(g_{xx} + \bar{p}^2 g_{cc})} \left\{ \frac{g_{xx}\omega}{g_x^2} + \left(-\frac{\bar{r} g_{xx}}{g_x^2} - \xi \right) \right\} \hat{b}(t)$$

$$\hat{x}(t) = -\hat{x}^*(t) = \frac{g_x g_c}{(g_{xx} + \bar{p}^2 g_{cc})} \left\{ \frac{g_{cc}(\bar{r} - \omega)}{g_c^2} - \xi \right\} \hat{b}(t)$$

$$\hat{\phi}(t) = -\hat{\phi}^*(t) = g_c \hat{b}(t)$$

where $\hat{\phi}$ and $\hat{\phi}^*$ are obtained by substituting (10.59) and the solution for $(\hat{c}, \hat{x})$ into (10.46) and the corresponding equation for country F, $\phi^* = h(Y - c, X - x, p)$.

A.3.2 The Effects of Preference Shifts

With quasi-luxury preferences (α, α^*) and impatience parameters (β, β^*), where $\beta^{(*)}$ satisfies $\delta_\beta^{(*)} > 0, \delta_{c\beta}^{(*)} = \delta_{x\beta}^{(*)} = 0$, as in Remark 4 in the text, the steady-state equilibrium conditions (10.39) through (10.42) reduce to

$$\delta(\bar{c}, \bar{x}; \beta) = \delta\left(Y - \bar{c}, X - \bar{x}; \beta^*\right) = \bar{r}$$

$$\overline{\text{MRS}}(\bar{c}, \bar{x}; \alpha, \beta) = \overline{\text{MRS}}^*\left(Y - \bar{c}, X - \bar{x}; \alpha^*, \beta^*\right) = \bar{p}$$

$$\bar{a} - \bar{a}^* = \frac{2\bar{c} - Y + \bar{p}(2\bar{x} - X)}{\bar{r}}$$

Differentiate totally these equations. Under assumption (10.38), the result can be arranged as

$$\delta_c \mathrm{d}\bar{c} + \delta_x \mathrm{d}\bar{x} = -\delta_\beta \left(\frac{\mathrm{d}\beta - \mathrm{d}\beta^*}{2} \right) \tag{10.62}$$

$$\overline{\text{MRS}}_c \mathrm{d}\bar{c} + \overline{\text{MRS}}_x \mathrm{d}\bar{x} = -\overline{\text{MRS}}_\alpha \left(\frac{\mathrm{d}\alpha - \mathrm{d}\alpha^*}{2} \right) - \overline{\text{MRS}}_\beta \left(\frac{\mathrm{d}\beta - \mathrm{d}\beta^*}{2} \right)$$

$$\mathrm{d}\bar{a} - \mathrm{d}\bar{a}^* = \frac{2}{\bar{r}} (\mathrm{d}\bar{c} + \bar{p}\mathrm{d}\bar{x}) \tag{10.63}$$

$$\mathrm{d}\bar{r} = \delta_\beta \left(\frac{\mathrm{d}\beta + \mathrm{d}\beta^*}{2} \right) \tag{10.64}$$

$$d\bar{p} = \overline{\text{MRS}}_\alpha \left(\frac{d\alpha + d\alpha^*}{2}\right) + \overline{\text{MRS}}_\beta \left(\frac{d\beta + d\beta^*}{2}\right)$$

where the partial derivatives of $\overline{\text{MRS}}$ are given by (10.25), (10.51), and (10.52).

From (10.62) through (10.63), the effects on $(\bar{c}, \bar{x}, \bar{a} - \bar{a}^*)$ can be derived as follows:

$$d\bar{c} = -d\bar{c}^* = -\frac{\bar{r}\delta_x \overline{\text{MRS}}_\alpha}{g_x^2 \Psi} \left(\frac{d\alpha - d\alpha^*}{2}\right) + \frac{\bar{r}\left(\overline{\text{MRS}}_x \delta_\beta - \overline{\text{MRS}}_\beta \delta_x\right)}{g_x^2 \Psi} \left(\frac{d\beta - d\beta^*}{2}\right)$$

$$d\bar{x} = -d\bar{x}^* = \frac{\bar{r}\delta_c \overline{\text{MRS}}_\alpha}{g_x^2 \Psi} \left(\frac{d\alpha - d\alpha^*}{2}\right) - \frac{\bar{r}\left(\overline{\text{MRS}}_c \delta_\beta - \overline{\text{MRS}}_\beta \delta_c\right)}{g_x^2 \Psi} \left(\frac{d\beta - d\beta^*}{2}\right) \tag{10.65}$$

$$d\bar{a} - d\bar{a}^* = \frac{\xi \overline{\text{MRS}}_\alpha}{g_x \Psi} \left(d\alpha - d\alpha^*\right) - \frac{\delta_\beta}{\bar{r} g_c \Psi} \left(\bar{\phi}\xi^2 - \frac{\bar{r} g_{xx}}{g_x^2} - \frac{\bar{r} g_{cc}}{g_c^2}\right) \left(d\beta - d\beta^*\right) \tag{10.66}$$

The effect (10.45) of an increase in α on net foreign assets is obtained from (10.41) and (10.64) through (10.65) as

$$\begin{aligned} \frac{\partial \bar{b}}{\partial \alpha} &= \frac{1}{\bar{r}} \frac{\partial \bar{c}}{\partial \alpha} + \frac{\bar{p}}{\bar{r}} \frac{\partial \bar{x}}{\partial \alpha} + \frac{X/2 - y_x}{\bar{r}} \frac{\partial \bar{p}}{\partial \alpha} - \frac{\bar{b}}{\bar{r}} \frac{\partial \bar{r}}{\partial \alpha} \\ &= \left\{ \xi - \left(y_x - \frac{X}{2} \right) \frac{g_x \Psi}{\bar{r}} \right\} \frac{\overline{\text{MRS}}_\alpha}{2 g_x \Psi} \end{aligned}$$

Corollary 3 follows from (10.66): Since $\left(\bar{\phi}\xi^2 - \frac{\bar{r} g_{xx}}{g_x^2} - \frac{\bar{r} g_{cc}}{g_c^2}\right) > 0$, the equation implies that the optimal $\bar{a}$ of a higher-β country can be larger than that of a lower-β country if the higher-β country has a sufficiently big α, compared with the lower-β country's.

Addendum: The Case of DMI[32]

In the text of this chapter, I have followed the literature (e.g., Uzawa 1968) in assuming that consumers exhibit increasing marginal impatience (hereafter IMI): $\delta_c > 0$ and $\delta_x > 0$. As shown in Hirose and Ikeda (2008), however, the assumption is empirically controversial. In this addendum, I discuss briefly on how the results would be (or not be) changed when assuming decreasing marginal impatience (hereafter DMI):

[32] This addendum has been newly written for this book chapter.

$$\delta_c < 0 \text{ and } \delta_x < 0$$

As pointed out by Hirose and Ikeda (2008), under DMI, the wealthier people are more patience and, ceteris paribus, become even wealthier over time. The property DMI is thus, by nature, destabilizing. Therefore, in the models of Sects. 2 and 3.2, where the interest rate is, exogenously or endogenously, constant over time, the dynamics could not satisfy the saddle-point stability under DMI. Indeed, since Ψ in (10.16) can be rewritten as

$$\Psi = -\left\{\xi^2 + \frac{Bf'\delta_c\delta_x}{g_c g_x}\left(\frac{g_{cc}}{\delta_c g_c} + \frac{g_{xx}}{\delta_x g_x}\right)\right\} \tag{10.67}$$

it is necessarily negative in the case of DMI, and hence the local dynamics do not satisfy the saddle-point stability condition in the models of Sects. 2 and 3.2.

In the production economy model of Sect. 3.1, in contrast, the equilibruim dynamics could be saddle-point stable under DMI. By assuming that there exists a non-satiated steady-state equilibrium such that $g_c > 0$ and $g_x > 0$, I focus on the local dynamics around the non-satiated steady-state point.[33] From the discussions in Appendix A.2.1, I can show easily that the dynamics are saddle-point stable under DMI if and only if

$$\Psi > -\frac{Bf''}{g_c g_x^2}\left(g_{xx} + B^2 g_{cc}\right) \tag{10.68}$$

where Ψ is now negative as seen from (10.67).

With (10.68), in turn, the signs of (10.34) and (10.36) are kept unchanged, and hence the results of comparative statics in Proposition 4 and Corollary 1 hold valid under DMI. That is, the greater preference for quasi-luxury goods leads to more steady-state capital stock. Taxation on quasi-luxury goods from no initial distortion decreases the steady-state capital stock.

References

Attanasio O, Weber G (1995) Is consumption growth consistent with intertemporal optimization? Evidence from consumer expenditure survey. J Polit Econ 103:1121–1157

Baland JM, Ray D (1991) Why does asset inequality affect unemployment? A study of the demand composition problem. J Dev Econ 35:69–92

Becker RA, Zilcha I (1997) Stationary Ramsey equilibria under uncertainty. J Econ Theory 75:122–140

Besley T (1989) A definition of luxury and necessity for cardinal utility functions. Econ J 99:844–849

Blanchard OJ (1985) Debt, deficits and finite horizons. J Polit Econ 93:223–247

[33] See Hirose and Ikeda (2008) for satiation under DMI, in which $g_c = 0$ and $g_x = 0$.

Blundell R, Browning M, Meghir C (1994) Consumer demand and the life-cycle allocation of household expenditures. Rev Econ Stud 61:57–80
Brewer A (1998) Luxury and economic development: David Hume and Adam Smith. Scott J Polit Econ 45:78–98
Brinkmann C (1957) Luxury. In: Seligman ERA (ed) Encyclopaedia of the social sciences. Macmillan, New York, pp 634–638
Browning M, Crossley TF (2000) Luxuries are easier to postpone: a proof. J Polit Econ 108:1022–1026
Buiter WH (1981) Time preference and international lending and borrowing in an overlapping-generations model. J Polit Econ 89:769–797
Clarida RH (1990) International lending and borrowing in a stochastic stationary equilibrium. Int Econ Rev 31:543–558
Cole HL, Mailath GJ, Postlewaite A (1992) Social norms, saving behavior, and growth. J Polit Econ 100:1092–1125
Corneo G, Jeanne O (1999) Social organization in an endogenous growth model. Int Econ Rev 40:711–710
Devereux M, Shi S (1991) Capital accumulation and the current account in a two-country model. J Int Econ 30:1–25
Epstein LG (1987a) A simple dynamic general equilibrium model. J Econ Theory 41:68–95
Epstein LG (1987b) The global stability of efficient intertemporal allocations. Econometrica 55:329–355
Epstein LG, Ham J, Zin SE (1988) Consumption, labor supply and portfolio choice with time and state nonseparable utility. Mimeo, University of Toronto
Epstein LG, Hynes JA (1983) The rate of time preference and dynamic economic analysis. J Polit Econ 91:611–625
Frenkel JA, Razin A (1992) Fiscal policies and the world economy, 2nd edn. MIT, Cambridge
Futagami K, Shibata A (1998) Keeping one step ahead of the Joneses: status, the distribution of wealth, and long run growth. J Econ Behav Organ 36:109–126
Hamermesh DS (1982) Social insurance and consumption: an empirical inquiry. Am Econ Rev 72:101–113
Hirose K, Ikeda A (2001) On decreasing impatience. ISER discussion paper No. 536
Hirose K, Ikeda A (2008) On decreasing impatience. Jpn Econ Rev 59:259–274
Ikeda S (2001) Weakly non-separable preferences and the Harberger-Laursen-Metzler effect. Can J Econ 34:290–307
Ikeda S (2003) Tariffs, time preference, and the current account under weakly nonseparable preferences. Rev Int Econ 11:101–113
Ikeda S (2006) Luxury and wealth. Int Econ Rev 47:495–526
Judd KL (1985) Redistributive taxation in a simple perfect foresight model. J Public Econ 28:59–83
List F (1841) The national system of political economy. Garland Publishing Co., New York. Reprinted (1974)
Lucas R, Stokey N (1984) Optimal growth with many consumers. J Econ Theory 32:139–171
Mandeville B (1714) The Fable of the Bees: or, private vices, public benefits. Clarendon Press, Oxford. Reprinted in Kaye FB (ed) (1924) Mandeville's Fable of the Bees: or, private vices, public benefits. Clarendon Press, Oxford
Mason R (1998) The economics of conspicuous consumption. Edward Elgar, Cheltenham
Obstfeld M (1982) Aggregate spending and the terms of trade: is there a Laursen-Metzler Effect? Q J Econ 97:251–270
Obstfeld M (1990) Intertemporal dependence, impatience, and dynamics. J Monet Econ 26:45–76
Obstfeld M, Rogoff K (1996) Foundations of international macroeconomics. The MIT Press, Cambridge
Parker JA (1999) The reaction of household consumption to predictable changes in social security taxes. Am Econ Rev 89:959–973
Ramsey F (1928) A mathematical theory of saving. Econ J 38:543–559
Shi S (1994) Weakly nonseparable preference and distortionary taxes in a small open economy. Int Econ Rev 35:411–428

Sombart W (1912) Liebe, Luxus und Kapitalismus. Dunckner & Humblot, Munich (Translated by W.R. Dittmar as: luxury and capitalism. The University of Michigan Press, Michigan, 1967)

Turnovsky SJ (2000) Methods of macroeconomic dynamics, 2nd edn. The MIT Press, Cambridge

Uzawa H (1968) Time preference, the consumption function and optimal asset holdings. In: Wolfe LN (ed) Value capital and growth: papers in honour of Sir John Hicks. Aldine, Chicago

Chapter 11
On Decreasing Marginal Impatience

Ken-ichi Hirose and Shinsuke Ikeda

Abstract One of the most controversial assumptions in endogenous time preference theory is that the degree of impatience is marginally increasing in wealth. We examine the implications of an empirically more relevant specification whereby time preference exhibits decreasing marginal impatience (DMI). With DMI, there are multiple steady-state non-satiated and satiated equilibria. In a constant interest rate economy, the non-satiated steady-state point is necessarily unstable. In a capital economy with decreasing returns technology, both the non-satiated and satiated steady-state points can be saddlepoint stable. The model is used to examine policy implications for the effects of capital taxation and government spending.

Keywords Decreasing marginal impatience • Time preference • Satiation • Capital taxation • Fiscal spending

1 Introduction

When the degree of impatience, as measured by the rate of time preference, is decreasing in wealth, the wealthier are more patient and, ceteris paribus, become even wealthier over time. Decreasing marginal impatience (hereafter DMI) is thus, by nature, destabilizing. To avoid the resulting analytical difficulty, the theory of endogenous time preference has usually assumed increasing marginal impatience (hereafter IMI) (e.g., Epstein 1987a,b; Epstein and Hynes 1983; Lucas and Stokey

---

The original article first appeared in the Japanese Economic Review 59: 259–274, 2008. A newly written addendum has been added to this book chapter.

K. Hirose
Department of Economics, Otaru University of Commerce, Otaru, Japan
e-mail: hirose@res.otaru-uc.ac.jp

S. Ikeda (✉)
Institute of Social and Economic Research, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan
e-mail: ikeda@iser.osaka-u.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_11

1984; Uzawa 1968). However, irrespective of the large amount of research conducted on time preference, limited research has thus far been conducted on the theoretical or economic implications of DMI. However, Svensson and Razin (1983) proposed that unstable equilibrium time paths of economies cannot be ruled out *a priori*. In fact, existing empirical studies, using either actual economic data (e.g., Lawrance 1991; Samwick 1998) or experimental data (e.g., Harrison et al. 2002; Ikeda et al. 2005), commonly and strongly support the validity of DMI. It is then of critical importance to determine the implications of DMI.[1]

The purpose of the present chapter is: (i) to examine the workings of a dynamic economic system under DMI; and (ii) to consider the policy implications. To do so, the optimal consumer behavior is examined first using a constant interest rate economy model, and then a capital economy model with variable interest rates.

With standard regularity conditions, including concavity, we show that under DMI there are multiple steady-state non-satiated and satiated equilibria. The presence of the satiated steady-state equilibrium allows us to consider the unstable optimal dynamics around usual non-satiated steady-state points. In a constant interest rate economy, the non-satiated steady-state point is shown to be necessarily unstable, which leads to the rich accumulating wealth up to a satiated level, while the wealth of the poor shrinks toward the zero consumption level.

In the neoclassical production economy with decreasing returns to capital, the usual modified golden-rule steady-state point, as well as the satiated point, can be saddlepoint stable. The model is used to examine the effects of capital taxation and government spending on capital accumulation.

An increase in capital taxes raises the long-run after-tax interest rate, such that, and in contrast to Epstein and Hynes (1983), the resultant decreases in steady-state capital and consumption, and hence welfare, are larger than in the case of a constant time preference. When time preferences exhibit DMI, a decrease in the capital stock caused by capital taxation makes consumers less patient, which raises the long-run interest rate irrespective of the capital tax increase, and thereby causes a further reduction in the capital stock. An increase in government spending, i.e., a negative income shock, under DMI raises the long-run interest rate and leads to a reduction in the long-run capital stock. This is also in sharp contrast to the implausible result found under IMI that a negative income shock enlarges the capital stock in the long run.

Das (2003), in an important contribution to a similar issue, also shows that DMI is compatible with saddlepoint stability in the neoclassical exogenous growth model. Our contribution differs in several regards. First, unlike her, we impose standard regularity conditions on consumer preference, which enables us to discuss both the

[1]As for the literature of a different interest, Becker and Mulligan (1997) discuss DMI using a "future-oriented" capital model, whereby the wealthy invest more "future-oriented" capital to obtain a lower time preference. Using a hyperbolic discounting model, Barro (1999) derives the possibility that time preference is decreasing in consumption during transition. Although it could be regarded as an observationally equivalent phenomenon to ours, the model is time inconsistent without any commitment tools.

DMI and IMI cases in the same setting. Second, with the regularity conditions, we obtain satiated steady-state solutions, which allow us to analyze optimal dynamics, even if the usual non-satiated steady-state point is unstable. Third, this chapter examines the macroeconomic policy implications of DMI.

The remainder of the chapter is organized as follows. Section 2 considers the effects of DMI using a simple intertemporal utility-maximization problem with a constant interest rate. In Sect. 3, the analysis is extended to the neoclassical model with capital accumulation to examine the effects of capital taxation and government spending. Section 4 concludes the chapter.

2 The Effects of Decreasing Marginal Impatience

To examine the effects of DMI, consider an infinitely lived consumer who maximizes lifetime utility by choosing the time profiles of consumption $\{c(t)\}_{t=0}^{\infty}$ and total asset holdings $\{a\,(t)\}_{t=0}^{\infty}$. His or her problem is specified in a simple form as:

$$\max \int_0^{\infty} u\,(c\,(t)) \exp(-\Delta\,(t))\mathrm{d}t, \tag{11.1}$$

subject to:

$$\dot{a}\,(t) = ra\,(t) - c\,(t)\,, a\,(0) = a_0(\text{constant}), a\,(t) \geq 0, \tag{11.2}$$

and

$$\dot{\Delta}\,(t) = \delta\,(c\,(t))\,, \Delta\,(0) = 0, \tag{11.3}$$

where a dot represents the time derivative; $u\,(c)$ represents the felicity function; and Δ denotes a cumulative discount rate with the instantaneous discount rate given by $\delta\,(c)$: $\Delta\,(t) = \int_0^t \delta\,(c\,(\tau))\mathrm{d}\tau$. Functions u and δ are twice continuously differentiable. In the present case, the rate of interest r is assumed to be constant. [2]

As in the literature (e.g., Epstein 1987a; Obstfeld 1990), consumer preferences are assumed to satisfy the following standard regularity conditions:

Assumption 1 $u < 0$;

Assumption 2 *u and δ* are *concave; and*

Assumption 3 *$-u$ is* log-*convex.*

With these conditions, the optimal solution is sufficiently given by the standard first-order conditions.

[2]Asset holding a includes both financial wealth and human capital. We can rewrite the positivity condition for a as the no-Ponzi game condition to bond holdings.

As is well known (and shown later by Eq. (11.13)), the degree of impatience is marginally increasing or decreasing in wealth as δ is increasing or decreasing in consumption. Note that what is required to the discount function δ by regularity conditions Assumptions 1–3 is just concavity,[3] and they are not related to whether the subjective discount rate δ is increasing or decreasing. Specifically, it can be hump shaped, where impatience is increasing for smaller consumption whereas it is decreasing for larger quantities of consumption.[4]

Remark 1 Several studies (e.g., Fukao and Hamada 1991; Jafarey and Park 1998) consider U-shaped nonmonotonic discount-rate functions. Their discount-rate functions do not satisfy Assumptions 1–3, implying that the usual first-order conditions may not provide the optimality conditions.

Let us focus on the case of DMI by additionally assuming the following:

Assumption 4 Impatience is marginally decreasing: $\delta_c(c) < 0$; where $\delta_c(c)$ represents $d\delta(c)/dc$.

Assumption 5 The upper bound $\bar{\delta}$ ($\equiv \delta(0)$) of $\delta(c)$ satisfies $\bar{\delta} > r$.[5]

Assumption 5 is related to the existence of a steady-state solution. From Assumptions 2, 4 and 5, there uniquely exists $c^* > 0$, such that:

$$\delta\left(c^*\right) = r. \tag{11.4}$$

The bottom panel of Fig. 11.1 depicts the resulting typical $\delta(c)$-schedule.

With IMI, a steady-state optimal solution would be uniquely given by (11.4). In the case of DMI, there is another steady-state solution, which is characterized by satiation. This takes place because, owing to the negativity of utility (Assumption 1), an increase in consumption has a negative impact on the lifetime utility level by lowering the discount rate when the consumption level is large enough.

More explicitly, let c' be any consumption level such that $\delta(c') > 0$. From concavity (Assumption 2), δ satisfies that for $c > c' - \delta(c')/\delta_c(c')$:

$$\begin{aligned}\delta(c) &< \delta\left(c'\right) + \delta_c\left(c'\right)\left(c - c'\right)\\ &< 0,\end{aligned}$$

[3]The regularity conditions, including the concavity of δ, ensure the concavity of the lifetime utility function. Obstfeld (1990) discusses this point intuitively using a two-period model. For a detailed proof in the infinite horizon case, see Hirose and Ikeda (2004a).

[4]Das (2003) proposes another set of regularity conditions under which δ *should* be decreasing in c. She cannot deal with the mixed case in which impatience can be marginally increasing or decreasing depending on the consumption level.

[5]From Assumptions 2 and 4, $\delta(c)$ necessarily has the upper bound at $c = 0$.

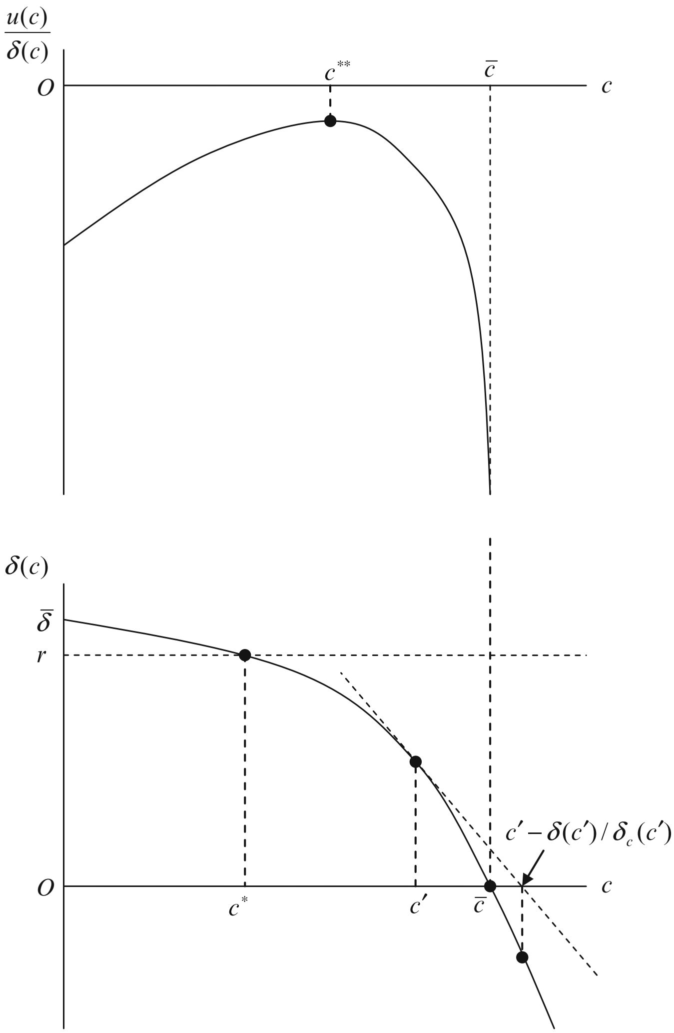


Fig. 11.1 Schedules of $u(c)/\delta(c)$ and $\delta(c)$

implying that, as illustrated by the bottom panel of Fig. 11.1, δ is negative for consumption levels that are higher than a finite critical value $\bar{c}$:

$$\delta(c) \gtreqless 0 \text{ as } c \lesseqgtr \bar{c}.$$

If the lifetime utility generated by a stationary consumption path $c(t) = c$ is denoted by $\mathcal{U}(c)$: $\mathcal{U}(c) \equiv \int_0^\infty u(c)\exp(-\delta(c)t)\mathrm{d}t$, it thus satisfies:

$$\mathcal{U}(c) = \begin{cases} u(c)/\delta(c), & \text{for } c < \bar{c}, \\ -\infty, & \text{for } c \geq \bar{c}, \end{cases} \tag{11.5}$$

implying that, as illustrated by the $\mathcal{U}(c)$-schedule in the top panel of Fig. 11.1, there exists a satiated consumption level $c^{**} \in [0, \bar{c})$ such that:

$$c^{**} = \arg\max_{c \geq 0} \mathcal{U}(c). \tag{11.6}$$

As for the satiated consumption level c^{**}, we assume the following:

Assumption 6 c^{**} is unique;

Assumption 7 For $c < c^{**}$, $\mathcal{U}(c)$ is strictly increasing; and

Assumption 8 $c^{**} > c^{*}$, where c^{*} is given in Assumption 4.

Assumptions 6 and 7 are put simply for analytical simplicity. If Assumption 8 were not satisfied, the steady-state solution c^{*}, defined by Assumption 4, would not be optimal when c^{*} is strictly larger than the satiation level c^{**}. The top panel of Fig. 11.1 depicts the resulting typical $\mathcal{U}(c)$-schedule.

The regularity conditions Assumptions 1–3 enable us to apply the usual maximization procedures. Letting λ and ϕ represent the current-value shadow prices for savings and the discount factor $\Omega = \exp(-\Delta)$, respectively, the optimal conditions to maximize the lifetime utility function, defined by Eq. (11.1), are given by

$$\xi_c(c, \phi) = \lambda, \tag{11.7}$$

$$\dot{\lambda} = (\delta(c) - r)\lambda, \tag{11.8}$$

$$\dot{\phi} = -\xi(c, \phi), \tag{11.9}$$

$$\lim_{t \to \infty} \exp(-\Delta(t))\lambda(t)a(t) = 0, \tag{11.10}$$

and

$$\lim_{t \to \infty} \exp(-\Delta(t))\phi(t)\Delta(t) = 0, \tag{11.11}$$

where ξ represents the generating function[6]:

$$\xi(c, \phi) = u(c) - \phi\delta(c).$$

[6]For the proof, see Hirose and Ikeda (2004a).

Two points are noteworthy. First, the optimal $\phi(t)$ equals the lifetime utility obtained from the optimal consumption stream after time t, as can be seen by solving Eq. (11.9) under the transversality condition, Eq. (11.11), and thus $\phi < 0$ from Assumption 1. Second, $\xi_c(c,\phi)\,(= u_c(c) - \phi\delta_c(c))$ represents the current-value marginal utility of c. As $\mathrm{d}(u(c)/\delta(c))/\mathrm{d}c = (u_c - (u/\delta)\delta_c)/\delta = \xi_c(c, u/\delta)/\delta$, Assumption 7 implies:

$$\xi_c\left(c, \frac{u(c)}{\delta(c)}\right) \begin{cases} > 0 \text{ for } c < c^{**}, \\ = 0 \text{ for } c = c^{**}, \end{cases}$$

meaning that the marginal utility evaluated at a stationary consumption time path $c(t) = c$ is strictly positive only when the consumption quantity is less than the satiation level, at which point marginal utility falls to zero. From Assumption 1 and $\phi < 0$, we obtain $\xi_{cc}(c,\phi) < 0$.

By substituting successively Eqs. (11.7) and (11.9) into (11.8), we obtain the optimal consumption dynamics as:

$$\dot{c} = -\frac{\xi_c(c,\phi)}{\xi_{cc}(c,\phi)}(r - \delta(c)) - \frac{\xi(c,\phi)}{\xi_{cc}(c,\phi)}\delta_c(c). \tag{11.12}$$

When $\xi_c > 0$, the rate of time preference ρ can be defined as $\rho = -\mathrm{d}\ln\Gamma(t)/\mathrm{d}t|_{\dot{c}=0}$, where $\Gamma \equiv \xi_c(c,\phi)\exp(-\Delta)$ represents the present value marginal utility of c, i.e.:

$$\rho(c,\phi) = \delta(c) - \frac{\xi(c,\phi)}{\xi_c(c,\phi)}\delta_c(c), \tag{11.13}$$

with which Eq. (11.12) can be rewritten in the usual form as:

$$\dot{c} = -\frac{\xi_c(c,\phi)}{\xi_{cc}(c,\phi)}(r - \rho(c,\phi)). \tag{11.14}$$

The optimal time path for (c,ϕ,a) must be jointly generated by Eqs. (11.2), (11.9), and (11.12) (or (11.14)). There are two steady-state optimal solutions: the (non-satiated) steady-state solution E^* and the satiated steady-state solution E^{**}, which are defined, respectively, by the following:

(a) A non-satiated steady-state solution E^*: (c^*,ϕ^*,a^*);

$$\delta(c^*) = r, \phi^* = \frac{u(c^*)}{\delta(c^*)}, ra^* = c^*, \text{ and } \xi_c(c^*,\phi^*) > 0. \tag{11.15}$$

(b) A satiated steady-state solution E^{**}: $(c^{**},\phi^{**},a^{**})$;

$$\xi_c(c^{**},\phi^{**})(= u_c(c^{**}) - \phi^{**}\delta_c(c^{**})) = 0, \phi^{**} = \frac{u(c^{**})}{\delta(c^{**})}, \text{ and } ra^{**} = c^{**}. \tag{11.16}$$

Solution E^* is characterized by the consumption level c^* in Eq. (11.4) and a strictly positive ξ_c, whereas the satiated steady-state solution E^{**} by the satiated consumption level c^{**} and $\xi_c = 0$. From Assumptions 4–6, both steady-state solutions E^* and E^{**} uniquely exist.[7,8]

Example 1 For a simple example, specify functions $u(c)$ and $\delta(c)$ in quadratic form as:

$$u(c) = -\left(\frac{100}{121}\right)(c-11)^2, \tag{11.17}$$

and

$$\delta(c) = -\frac{0.1}{100}c^2 + 0.1, \tag{11.18}$$

respectively. These functions can be shown easily to satisfy the regularity conditions Assumptions 1–5 for $c \in (0, 11)$.[9] The critical consumption level $\bar{c}$ that satisfies $\delta(\bar{c}) = 0$ equals 10. The resulting $\mathcal{U}(c)$-function (Eq. (11.5)) satisfies Assumptions 6 and 7, where the satiated consumption level c^{**} is uniquely given by 9.091. The non-satiated steady-state equilibrium E^* uniquely exists if $r \in [0.018, 0.1)$.[10] For example, when $r = 0.05$, the steady-state consumption c^* amounts to 7.071, which is smaller than satiated level 9.091, as required by Assumption 8.

By analyzing the local dynamics around the two steady-state points, we can show that non-satiated steady-state point E^* is unstable whereas the satiated point E^{**} is saddlepoint stable (see Appendix "Stability of Points E^* and E^{**} in Proposition 1").

Proposition 1 *Under Assumptions 1–8, the optimization problem with DMI, given by Eqs. (11.1)–(11.3), has two steady-state optimal points: a non-satiated point E^*, which is unstable, and a satiated point E^{**}, which is saddlepoint stable.*

The resulting optimal consumption dynamics are depicted in Fig. 11.2, where the top panel illustrates the (c, a)-dynamics and the bottom depicts (c, ϕ). In the bottom panel, we illustrate two $\dot{c} = 0$ schedules. One is defined by $\rho(c, \phi) = r$ and is locally flat at E^*. The other is upward sloping at E^{**}. Schedule $\dot{\phi} = 0$ illustrates the relation $\xi(c, \phi) = 0$. As a zero ξ means $\phi = u(c)/\delta(c)$, the schedule is exactly the same as depicted in the top panel of Fig. 11.1. Steady-state points $E^*(c^*, \phi^*)$

[7] This shows that, even if the felicity function is strictly increasing, satiation can arise under intertemporally nonseparable preference. Ryder and Heal (1973) show that habit formation can produce satiated steady-state optimal solutions.

[8] Satiation has also been reported in happiness studies; see, e.g., Leu et al. (1997) and Tsutsui et al. (2005).

[9] Function (11.17) can be respecified to satisfy the regularity conditions for *all* $c \geq 0$ by arbitrarily modifying the graph for $c \geq 11$.

[10] The exact value of the lower bound is $0.017355\cdots$.

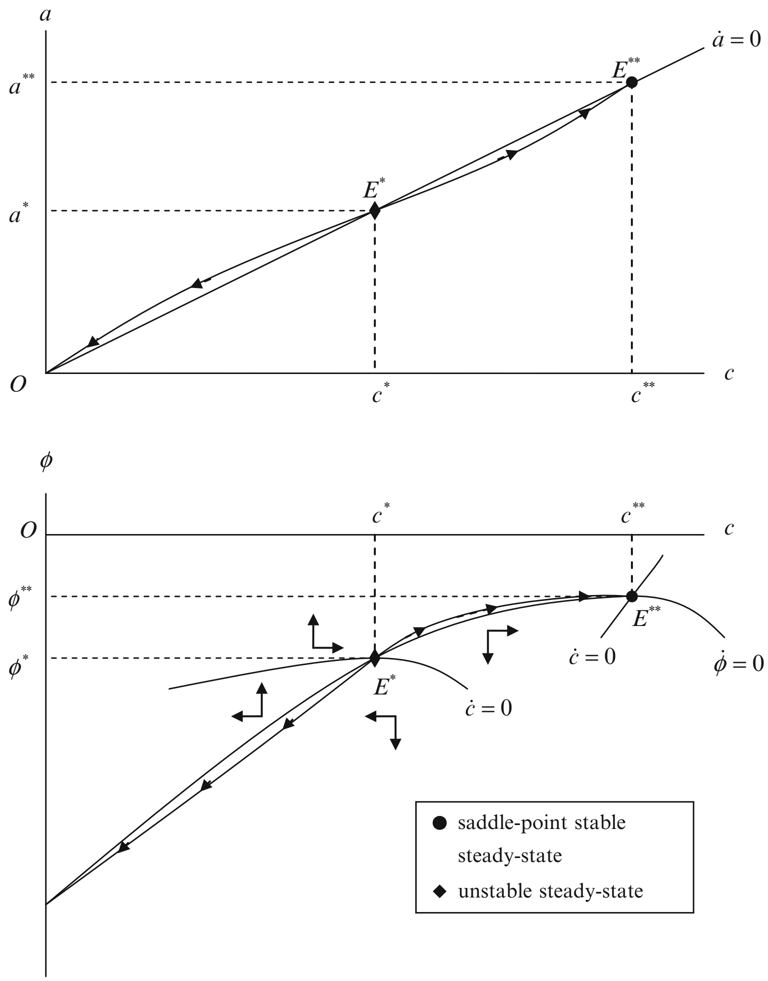


Fig. 11.2 Optimal consumption dynamics under decreasing marginal impatience with a constant interest rate

and E^{**} (c^{**}, ϕ^{**}) are determined at the intersections of the $\dot{\phi} = 0$ schedule and the two $\dot{c} = 0$ schedules. By assumption, the $\dot{\phi} = 0$ schedule has a unique peak at E^{**}. The optimal consumption dynamics are indicated by arrows. Point E^{*} is depicted as unstable whereas E^{**} is saddlepoint stable. Given an initial value a_0 of wealth holding, optimal consumption $c(0)$ is determined on the optimal trajectory in the (c, a)-plane. The (c, ϕ)-dynamics are then generated on the arrowed trajectory in the (c, ϕ)-plane.

The optimal consumption dynamics depend crucially on a_0. Suppose first that $0 < a_0 < a^{*}$. Then, $c(0)$ is determined as smaller than c^{*}, and thereafter $c(t)$ and hence $a(t)$ implode over time. Consider, instead, the case that $a^{*} < a_0 < a^{**}$.

In this case, $c(0)$ exceeds c^* but falls short of c^{**}. In the interim run, $c(t)$ grows gradually toward the satiation level c^{**}. Finally, when $a_0 > a^{**}$, a larger $c(t)$ than c^{**} would generate negative marginal utility. The optimal solution is thus not to choose greater consumption levels than c^{**}, even though they are feasible, but to keep $c(t)$ equal to the satiation level c^{**}.

As discussed, DMI in a constant interest rate economy leads poor consumers to decumulate wealth toward zero and rich consumers to accumulate wealth up to the satiated level. Note that the attained long-run consumption levels (zero or the satiation level) are insensitive to any income shocks, e.g., shocks in a_0 and/or r. With constant interest rates, the DMI model may thus not be suitable for analyzing the long-run effects of policy changes.

This, however, does not imply that we cannot analyze any DMI models. We can consider well-behaved models with DMI by introducing some stabilizing decreasing return properties into the production technology and/or consumer preference. We next incorporate capital accumulation with the usual decreasing returns technology.

3 Decreasing Marginal Impatience and Capital Accumulation

3.1 The Neoclassical Model

Consider a stylized neoclassical model with two production factors, labor and capital, a single multipurpose commodity produced using constant-to-scale technology F, and competitive firms. Consumers inelastically supply one unit of labor in each instant. Their preferences are specified as in the previous section. In particular, we assume DMI, $\delta_c(c) < 0$. The government spends g by levying capital taxes τ and lump-sum taxes. Letting k represent the capital–labor ratio and f a per capita production function satisfying $f_k > 0$, $f_{kk} < 0$, and the Inada conditions, we can easily obtain a reduced dynamic system as follows:

$$\begin{aligned} \dot{c} &= -\frac{\xi_c(c,\phi)}{\xi_{cc}(c,\phi)}\left((1-\tau) f_k(k) - \delta(c)\right) - \frac{\xi(c,\phi)}{\xi_{cc}(c,\phi)}\delta_c(c), \\ \dot{\phi} &= -\xi(c,\phi), \qquad (11.19) \\ \dot{k} &= f(k) - c - g. \end{aligned}$$

The solution of this system that satisfies the transversality condition is sufficiently optimal. When $\xi_c(c,\phi) > 0$, the first equation can be rewritten as:

$$\dot{c} = -\frac{\xi_c(c,\phi)}{\xi_{cc}(c,\phi)}\left((1-\tau) f_k(k) - \rho(c,\phi)\right),$$

where the rate of time preference $\rho(c,\phi)$ is given by Eq. (11.13).

As in the previous section, we consider two steady-state equilibria: (i) the non-satiated, modified golden-rule steady-state equilibrium $E^* \left(c^*, k^*\right)$ such that

$$\delta\left(c^*\right) = (1-\tau)\, f_k\left(k^*\right), \tag{11.20}$$

and

$$f\left(k^*\right) = c^* + g, \tag{11.21}$$

and (ii) the satiated steady-state equilibrium $E^{**}\left(c^{**}, k^{**}\right)$, where c^{**} is given by (11.6); and k^{**} is given by $f\left(k^{**}\right) = c^{**} + g$ as in Eq. (11.21).

By analyzing the local dynamics around these two steady-state points, we can show that the non-satiated steady-state point E^*, as well as the satiated point E^{**}, can be saddlepoint stable.

Lemma 1 1. *Non-satiated steady-state point E^* is saddlepoint stable if and only if:*

$$f_k\left(k^*\right) < (1-\tau)\, f_{kk}\left(k^*\right) / \delta_c\left(c^*\right). \tag{11.22}$$

2. *Satiated steady-state point E^{**} is saddlepoint stable if and only if:*

$$\delta\left(c^{**}\right) < (1-\tau)\, f_k\left(k^{**}\right). \tag{11.23}$$

Proof See Appendix "Proof of Lemma 1". ■

Figure 11.3a, b demonstrate the determination and the stability properties of the steady-state points E^* and E^{**}, where Eqs. (11.20) and (11.21) are depicted by upward sloping schedules. With the satiated consumption level c^{**} being determined from Eq. (11.6), satiated steady-state point E^{**} is determined on the schedule of Eq. (11.21). Non-satiated steady-state point E^* is given at the intersection of the two schedules if it exists on the left side of E^{**}. For point E^* to be saddlepoint stable, Lemma 1 (1) requires that the gradient of the schedule of (11.21) at E^* (the left-hand side of Inequality (11.22)) be smaller than that of Eq. (11.20) (the right-hand side). From Lemma 1 (2), point E^{**} is saddlepoint stable if, and only if, it is located above the schedule of Eq. (11.20).

Note that $\delta(0) = \bar{\delta}\,(< \infty)$ whereas, from the Inada condition, we have $\lim_{k \to 0} f_k(k) = \infty$. The schedule of Eq. (11.20) thus intersects the horizontal axis at a positive k. As the schedule of Eq. (11.21) goes through the origin when g is assumed to equal zero for brevity, this implies that the relative magnitudes of $\delta\left(c^{**}\right)$ and $(1-\tau)\, f_k\left(k^{**}\right)$, i.e., whether satiated steady-state point E^{**} is located above or below the schedule, Eq. (11.20) has critical implications for the existence and stability of the non-satiated and satiated steady-state points, and can be summarized as follows.[11]

[11] In the case that $g > 0$, Proposition 2 remains valid as far as $(1-\tau) f_k(f^{-1}(g)) > \bar{\delta}$.

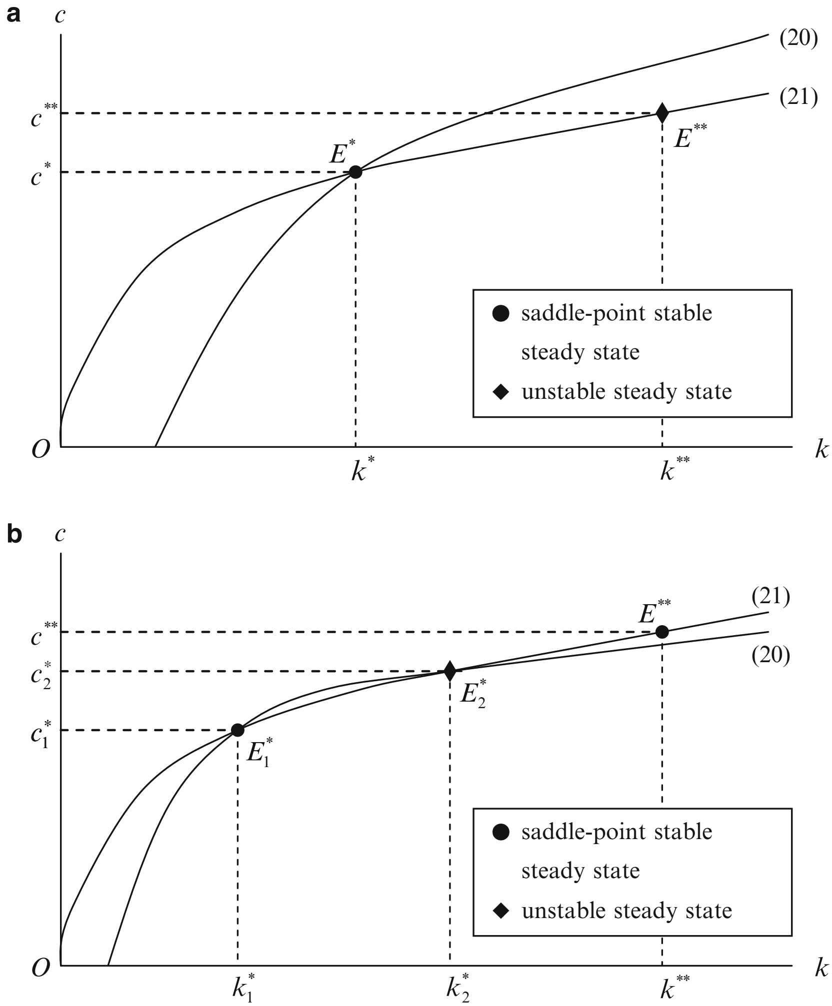


Fig. 11.3 Steady-state equilibria and stability. (**a**) $\delta(c^{**}) > (1-\tau)\, f_k(k^{**})$. (**b**) $\delta(c^{**}) < (1-\tau)\, f_k(k^{**})$

Proposition 2 1. *Suppose that* $\delta\,(c^{**}) > (1-\tau)\; f_k\,(k^{**})$. *Then, (i) there necessarily exist an odd number of non-satiated steady-state points, which are alternatively saddlepoint stable and unstable, i.e., the first is saddlepoint stable, the second is unstable,* $\cdots$, *and the last is saddlepoint stable; and (ii) the satiated steady-state point is unstable.*

2. *Suppose that* $\delta\,(c^{**}) < (1-\tau)\; f_k\,(k^{**})$. *Then, (i) if there exists a* $k < k^{**}$ *such that* $\delta\,(f\,(k)) \geq (1-\tau)\; f_k\,(k)$, *there exist an even number of steady-state points, which are alternatively saddlepoint stable and unstable, i.e., the first is*

saddlepoint stable, the second is unstable $\cdots$, and the last is unstable; otherwise there exists no non-nonsatiated steady-state point; and (ii) the satiated steady-state point is saddlepoint stable.

Figure 11.3a depicts the case that $\delta(c^{**}) > f_k(k^{**})$ by assuming that the non-satiated steady-state point E^* is unique. From Lemma 1 (1), point E^* is saddlepoint stable whereas, from Lemma 1 (2), satiated steady-state point E^{**} is unstable, as stated by Proposition 2 (1). Insofar as the initial capital stock k_0 lies below the satiated stock level k^{**}, the economy monotonically converges to the non-satiated steady-state point E^*. Figure 11.3b illustrates the case that $\delta(c^{**}) < f_k(k^{**})$, where two non-satiated steady-state points are assumed to exist. As implied by Proposition 2 (2), the first point E_1^* is saddlepoint stable whereas the second point E_2^* is unstable. Point E^{**} is saddlepoint stable. When k_0 lies below k_2^*, the economy gradually approaches non-satiated steady-state point E_1^*, whereas a higher k_0 than k_2^* is followed by monotonic convergence toward satiated steady-state point E^{**}.

Example 2 As in Example 1, we specify functions $u(c)$ and $\delta(c)$ in quadratic forms as Eqs. (11.17) and (11.18), respectively, and the production function as:

$$f(k) = Ak^{0.3}, A > 0.$$

Assume that $g = 0$ and $\tau = 0$. These functions satisfy all the regularity conditions. The satiated consumption level is obtained as $c^{**} = 9.091$. The existence of the non-satiated steady-state equilibrium E^* depends on the value of total factor productivity A. From Proposition 2, there exists a non-satiated equilibrium point if and only if for some $k < k^{**}$, $\delta(f(k)) \geq (1-\tau) f_k(k)$ and hence $\delta\left(Ak^{0.3}\right) \geq 0.3Ak^{-0.7}$. With a too large A, however, the capital productivity $0.3Ak^{-0.7}$ remains larger than $\delta\left(Ak^{0.3}\right)$ for all $k < k^{**}$. For the existence of the non-satiated steady-state equilibrium, A must thus be smaller than some critical value, which can be computed as 2.302 in the present example.[12] After tedious computation, we can show the following relations.

1. For $A \in (0, 1.993)$, the non-satiated steady-state equilibrium point E^* uniquely exists and is saddlepoint stable, as in Fig. 11.3a.[13] For example, when $A = 1.5$, (c^*, k^*) equals $(2.974, 9.786)$.
2. For $A \in (1.994, 2.301)$, there are two non-satiated steady-state equilibria, one saddlepoint stable and one unstable, as in Fig. 11.3b. For example, when $A = 2.1$, (c^*, k^*) is given by $(5.335, 22.373)$, which is saddlepoint stable, and $(8.826, 119.811)$, which is unstable.
3. For $A \geq 2.302$, there is a non-nonsatiated equilibrium.

[12] The exact critical value is $2.30160\cdots$.

[13] The exact value of the upper bound is $1.99397\cdots$.

With a decreasing return technology, a model of DMI can thus be well behaved in the sense that there exists a non-satiated steady-state point that is saddlepoint stable. This enables us to consider the policy implications of DMI by conducting usual comparable statics.

3.2 *The Effects of Capital Taxation*

Let us examine the implications of DMI on capital taxation. As shown by Chamley (1981), when the rate of time preference is constant, capital shifts away the entire burden of capital taxation in the long run because the long-run after-tax rate of return to capital must equal the constant time preference rate. The resulting reductions in the steady-state capital, consumption, and welfare are large. With endogenous time preference, Epstein and Hynes (1983) show that capital taxation reduces the steady-state capital stock, but not as much as it would under constant time preference, implying that the reductions in consumption and welfare are mitigated. However this depends crucially on the assumption of IMI. With DMI, the result is drastically changed, as we shall show below.

Assume that the economy initially stays at a non-satiated, saddlepoint stable steady-state point. From Eqs. (11.20) and (11.21), the steady-state capital stock k^* and the long-run after-tax rate of return to capital $r^* \equiv (1-\tau) f_k(k^*)$ are determined by:

$$(r^* =)\,(1-\tau)\, f_k\left(k^*\right) = \delta\left(f\left(k^*\right) - g\right). \qquad (11.24)$$

As shown in Fig. 11.4, $(1-\tau) f_k(k^*)$ on the left-hand side can be depicted as a downward-sloping schedule in the (r, k) plane. With DMI, the right-hand side, $\delta(f(k^*) - g)$, can also be expressed by a downward-sloping schedule. The steady-state capital stock k^* and the long-run after-tax rate of return to capital r^* are given at the intersection, say point E_0, of the two schedules. For the initial steady-state point E_0 to be saddlepoint stable, from Lemma 1, the $(1-\tau) f_k(k^*)$ schedule is steeper than the $\delta(f(k^*) - g)$ schedule at E_0.

Following Chamley (1981) and Epstein and Hynes (1983), suppose that the government raises capital tax τ and pays back the revenue to consumers in a lump-sum manner, by keeping fiscal spending g constant. It shifts the $(1-\tau) f_k(k^*)$ schedule downward, thereby bringing the steady-state point from point E_0 to E_1. Consequently, k^* decreases in response to the tax increase. Note that this reduction in k^* is larger than in the case of constant time preference: if δ were constant, the reduction would stop at k'. This property contrasts sharply with the results in Epstein and Hynes (1983) under IMI, in which case the reduction in k^* caused by capital taxation is smaller than in the case of constant time preference. With DMI, a decrease in k^* makes consumers less patient, which raises the long-run after-tax interest rate irrespective of the capital taxation, and thereby causes a further

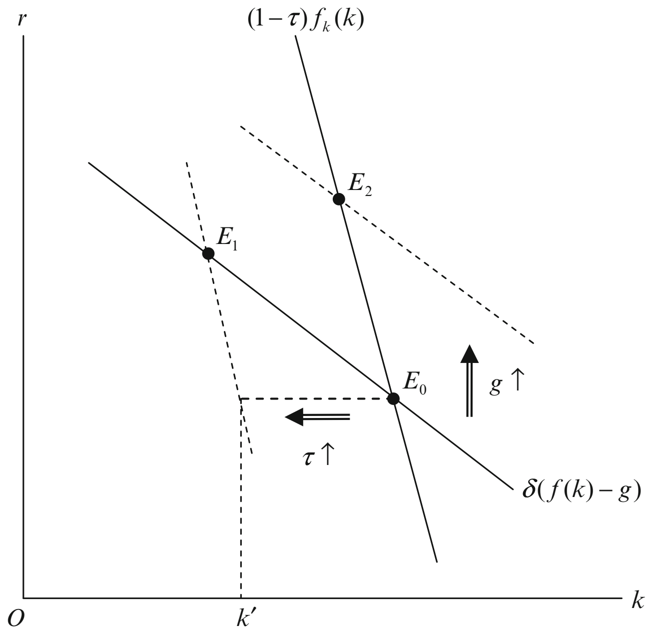


Fig. 11.4 Steady-state capital stock under decreasing marginal impatience

reduction in k^*. From Eq. (11.21), in turn, the resulting decreases in steady-state consumption and welfare $u(c^*)/\delta(c^*)$ are larger than they would have been under constant time preference.

Implication 1 *With DMI, an increase in capital taxes raises the long-run after-tax interest rate, so that, in contrast to the case of IMI, the resulting decrease in steady-state capital, consumption, and hence welfare are larger than they would have been under constant time preference.*

3.3 The Effects of Government Spending

By using Eq. (11.24) and hence Fig. 11.4, we can also consider the effect of an increase in government spending financed by lump-sum taxation. Suppose that the government increases its spending g permanently by raising the lump-sum tax while keeping the capital tax τ constant. As shown in Fig. 11.4, it shifts the $\delta(f(k)-g)$ schedule upward, thereby bringing the steady-state point from point E_0 to E_2 along the $(1-\tau) f_k(k)$ schedule. An increase in g thus raises r^* and reduces k^*. The fiscal policy makes consumers cut down consumption and raises the rate of time preference under DMI, which leads to a higher r^* and hence a smaller k^*. These properties differ from the result under constant time preference that either r^* or k^* is not affected by an increase in g, and the result under IMI that the g increase lowers r^* and thereby enlarges k^*.

Implication 2 *With DMI, an increase in government spending raises the long-run interest rate and harms capital accumulation.*

4 Conclusions

This chapter has demonstrated the implications of DMI for dynamic consumer behavior and macroeconomic policy. We have first shown that with DMI, there are multiple steady-state non-satiated and satiated equilibria. When the interest rate is constant, the non-satiated steady-state point is necessarily unstable, which leads the rich to accumulate wealth up to a satiated level, while the wealth of the poor shrinks toward zero consumption. In a capital economy with decreasing returns technology, both of the non-satiated and satiated steady-state points can be saddlepoint stable. Unlike the IMI case, the negative long-run effects of an increase in capital taxes on consumption, capital stocks and hence, welfare, are larger than they would have been under constant time preference. An increase in government spending reduces the long-run capital stock.

There are a number of interesting related issues. First, the model of constant interest rate economy can be applied straightforwardly to the analysis of various policy issues in small open economies. Second, in a two-country context, DMI leads to various interesting multiple steady-state equilibria, as discussed in Hirose and Ikeda (2004a). Third, the neoclassical model with DMI can be extended to incorporate money and thereby examine the effect of inflation.

Acknowledgements The authors would like to thank K. Futagami, W. Jodo, K. Koda and two anonymous referees for their helpful comments on an earlier version of this paper. A part of this research is financially supported by the Osaka University 21st Century Center of Excellence Program and Grants-in-Aid for Scientific Research C (No.18530137).

Appendix

Stability of Points E^* and E^{**} in Proposition 1

Local optimal dynamics around the non-satiated steady-state point (c^*, ϕ^*, a^*) is linearized as:

$$\begin{pmatrix} \dot{c} \\ \dot{\phi} \\ \dot{a} \end{pmatrix} = \begin{pmatrix} 0 & \frac{r\delta_c}{\xi_{cc}} & 0 \\ -\xi_c & r & 0 \\ -1 & 0 & r \end{pmatrix} \begin{pmatrix} c - c^* \\ \phi - \phi^* \\ a - a^* \end{pmatrix},$$

where the coefficient matrix is evaluated at (c^*, ϕ^*, a^*). This system has three positive characteristic roots:

$$\frac{1}{2}\left\{r + \left(r^2 - \frac{4r\delta_c\xi_c}{\xi_{cc}}\right)^{1/2}\right\}, \frac{1}{2}\left\{r - \left(r^2 - \frac{4r\delta_c\xi_c}{\xi_{cc}}\right)^{1/2}\right\}, \text{ and } r,$$

implying that the non-satiated steady-state point (c^*, ϕ^*, a^*) is unstable.

Local optimal dynamics around the satiated steady-state point $(c^{**}, \phi^{**}, a^{**})$ is linearized as:

$$\begin{pmatrix} \dot{c} \\ \dot{\phi} \\ \dot{a} \end{pmatrix} = \begin{pmatrix} \delta - r & \frac{r\delta_c}{\xi_{cc}} & 0 \\ 0 & \delta & 0 \\ -1 & 0 & r \end{pmatrix} \begin{pmatrix} c - c^{**} \\ \phi - \phi^{**} \\ a - a^{**} \end{pmatrix},$$

where the coefficient matrix is evaluated at $(c^{**}, \phi^{**}, a^{**})$. This system has characteristic roots r, $\delta(c^{**})$, and $\delta(c^{**}) - r$, which is negative as $\delta(c^{**}) < \delta(c^*) = r$ from Assumption 6. The satiated steady-state point $(c^{**}, \phi^{**}, a^{**})$ is thus saddlepoint stable.

Proof of Lemma 1

By linearizing system in Eq. (11.19) around the non-satiated steady-state point (c^*, ϕ^*, k^*), the local dynamic system can be obtained as:

$$\begin{pmatrix} \dot{c} \\ \dot{\phi} \\ \dot{k} \end{pmatrix} = \begin{pmatrix} 0 & \frac{\delta\delta_c}{\xi_{cc}} & -\frac{\xi_c(1-\tau) f_{kk}}{\xi_{cc}} \\ -\xi_c & \delta & 0 \\ -1 & 0 & f_k \end{pmatrix} \begin{pmatrix} c - c^* \\ \phi - \phi^* \\ a - a^* \end{pmatrix},$$

where the coefficient matrix is evaluated at (c^*, ϕ^*, k^*). For this coefficient matrix:

$$\text{trace} = \delta + f_k > 0, \text{det.} = \frac{\delta\xi_c}{\xi_{cc}} \left(\delta_c f_k - (1-\tau) f_{kk}\right).$$

The linear system thus has one negative and two positive roots if and only if $\delta_c(c^*) f_k(k^*) - (1-\tau) f_{kk}(k^*) > 0$, as stated as the first item in Lemma 1.

By linearizing the system equation (11.19) around the satiated steady-state point $(c^{**}, \phi^{**}, k^{**})$, the local dynamic system can be obtained as:

$$\begin{pmatrix} \dot{c} \\ \dot{\phi} \\ \dot{k} \end{pmatrix} = \begin{pmatrix} \delta - (1-\tau) f_k & \frac{\delta\delta_c}{\xi_{cc}} & 0 \\ 0 & \delta & 0 \\ -1 & 0 & f_k \end{pmatrix} \begin{pmatrix} c - c^{**} \\ \phi - \phi^{**} \\ a - a^{**} \end{pmatrix},$$

where the coefficient matrix is evaluated at $(c^{**}, \phi^{**}, k^{**})$. The characteristic roots are f_k, δ, and $\delta - (1-\tau)\, f_k$. The linear system thus has one negative and two positive roots if and only if $\delta\,(c^{**}) - (1-\tau)\, f_k\,(k^{**}) < 0$, as stated as the second statement in Lemma 1.

Addendum: Related Studies[14]

In the text article, we formalize DMI by specifying the subjective discount rate as a function of consumption (or instantaneous utility). Alternatively, it is possible to introduce DMI by assuming that the subjective discount rate is a decreasing function with respect to wealth (e.g., Schumacher 2009) or saving (e.g., Gootzeit et al. 2002). Becker and Mulligan (1997) deal with DMI in a "future-oriented capital" model, in which accumulation of the future-oriented capital leads to a lower discount rate, so that wealthier people become more patient. In either case, similar economic implications of DMI, as discussed in our article, could be obtained.

Literature on dynamic macroeconomic theory incorporates DMI for various purposes, such as to analyze growth dynamics in an overlapping generations model (e.g., Sarkar 2007) and in an AK model with borrowing constraints (e.g., Borissov 2013). It has helped investigate asset pricing in an overlapping generations model (e.g., Nath and Sarkar 2006) and to examine equilibrium indeterminacy in response to interest-rate rules (e.g., Chang et al. 2011). It also allows us to consider the effects of inflation on capital accumulation (e.g., Chen et al. 2008; Gong 2006; Hirose and Ikeda 2004b).

Based on the text article, we have been advancing further research on DMI. Hirose and Ikeda (2012a,b) investigate implications of DMI in a two-country world economy. If both countries exhibit DMI, the steady-state equilibrium is always unstable. For saddle-point stability, at least one country needs to exhibit IMI. Hirose and Ikeda (2012a) analyze the equilibrium dynamics in a one-good, two-country model where one country has DMI and the other has IMI.

Hirose and Ikeda (2012b) solve for two-good, two-country equilibrium dynamics with endogenous time preference, and re-examine the Harberger-Laursen-Metzler (hereafter HLM) effect, which states that a terms-of-trade deterioration would cause a reduction in national savings and a current-account deficit. Although the HLM effect is invalid for a small country with IMI preference (as shown in Obstfeld 1982), it can be rehabilitated in a two-country economy. The terms-of-trade deterioration affects the long-run accumulation of net foreign assets and hence the current account through the following three channels: (a) the income-compensating effect (which is always positive), (b) the welfare-supporting effect, and (c) the interest-income effect. In the case where both countries have IMI, the HLM effect can materialize if the negative welfare-supporting effect dominates the positive income-compensating

[14] This addendum has been newly written for this book chapter.

Table 11.1 Associations between time preference (discounting) and income

Study	Data	Sample regions	Impatience-income association	Budget variable
Hausman (1979)	Field survey (Midwest Research Institute study)	U.S.	Negative	Income
Lawrance (1991)	PSID	U.S.	Negative	Income
Ogawa (1993)	National macro data	Japan, Korea, Taiwan	U-shaped (Japan, Taiwan) const. (Korea)	National disposable income
Pender (1996)	A field household panel survey (ICRISAT)	India	Negative	Net wealth
Ogaki and Atkeson (1997)	A field household panel survey (ICRISAT)	India	No association	Income
Samwick (1998)	Survey of Consumer Finance	U.S.	Negative	Income
Coller and Williams (1999)	Experiment	U.S.	Positive	Income
Donkers and van Soest (1999)	Survey (CentERpanel)	Holland	No association	Income
Harrison et al. (2002)	Experiment	Denmark	Negative	Income
Kapteyn and Teppa (2003)	Survey (CentERpanel)	Holland	Negative	Income
Ventura (2003)	Survey (Bank of Italy Survey of Household Income and Wealth)	Italy	Negative	Income, wealth
Read and Read (2004)	Non-incentivized choice task	U.K.	Negative	Income
Anderson et al. (2004)	Field interview survey	Vietnam	No association	Income
Ikeda et al. (2005)	Experiment	Japan	Negative	Reward income in the experiments
Booij and van Praag (2009)	NIPO Post-Initial Schooling Survey	Holland	Negative	Income
Tanaka et al. (2010)	Experiment	Vietnam	Negative	Income
Wang et al. (2011)	An original international survey	45 countries	Negative	Income

effect. In the case where one country exhibits DMI and the other exhibits IMI, the HLM effect is necessarily invalid for the IMI country (since the welfare-supporting effect is positive) whereas it may be valid for the DMI country (due to the negative welfare-supporting effect).

As for empirical research, it is a matter of controversy as to how time preference and the discount rate relate to the decision maker's degree of affluence, measured by income and/or wealth. However, the majority of previous research reports that the degree of impatience, measured by time preference or personal discount rate, is negatively associated with income and/or wealth. Table 11.1 summarizes the previous literature. Although it covers only a part of the literature, 12 of the 17 studies listed indicate that richer people are more patient, as per the DMI model.

Note, however, that the detected associations do not capture any causality. In particular, since more patient people would have higher saving propensity and hence be wealthier, there could be an endogeneity problem when estimating how time preference relates to income and wealth. The previous studies in Table 11.1 do not cope with the problem.[15] It is an important research topic to tackle this problem and thereby detect causal relationship between time preference and income/wealth.

References

Anderson L, Dietz M, Gordon A, Klawitter M (2004) Discount rates in Vietnam. Econ Dev Cult Change 52:873–887

Barro RJ (1999) Ramsey meets Laibson in the neoclassical growth model. Q J Econ 114:1125–1152

Becker GS, Mulligan CB (1997) The endogenous determination of time preference. Q J Econ 112:729–758

Booij AS, van Praag BMS (2009) A simultaneous approach to the estimation of risk aversion and the subjective time discount rate. J Econ Behav Organ 70:374–388

Borissov K (2013) Growth and distribution in a model with endogenous time preferences and borrowing constraints. Math Soc Sci 66:117–128

Chamley C (1981) The welfare cost of capital income taxation in a growing economy. J Politi Econ 89:468–496

Chang W, Tsai H, Chang J (2011) Endogenous time preference, interest-rate rules, and indeterminacy. Jpn Econ Rev 62:348–364

Chen B, Hsu M, Lu C (2008) Inflation and growth: impatience and a qualitative equivalence. J Money Credit Bank 40:1309–1323

Coller M, Williams MB (1999) Eliciting individual discount rates. Exp Econ 2:107–127

Das M (2003) Optimal growth with decreasing marginal impatience. J Econ Dyn Control 27:1881–1898

Donkers B, van Soest A (1999) Subjective measures of household preferences and financial decisions. J Econ Psychol 20:613–642

Epstein LG (1987a) A simple dynamic general equilibrium model. J Econ Theory 41:68–95

Epstein LG (1987b) The global stability of efficient intertemporal allocations. Econometrica 55:329–355

Epstein LG, Hynes JA (1983) The rate of time preference and dynamic economic analysis. J Politi Econ 91:611–625

[15]Ikeda et al. (2005) is an exception. By using the instrumental variable method, they detect negative association between the discount rate and income that the subjects obtained in the risk-choice experiment conducted, before the time-discounting experiment.

Fukao K, Hamada K (1991) The Fisherian time preference and the long-run capital ownership pattern in a global economy. NIRA Res Output 4(1):86–97

Gong L (2006) Endogenous time preference, inflation, and capital accumulation. J Econ 87:241–255

Gootzeit M, Schneider J, Smith W (2002) Marshallian recursive preferences and growth. J Econ Behav Organ 49:381–404

Harrison GW, Lau MI, Williams MB (2002) Estimating individual discount rates in Denmark: a field experiment. Am Econ Rev 92:1606–1617

Hausman JA (1979) Individual discount rates and the purchase and utilization of energy-using durables. Bell J Econ 10:33–54

Hirose K, Ikeda S (2004a) On decreasing marginal impatience. Osaka University COE Discussion Paper No.44

Hirose K, Ikeda S (2004b) Decreasing marginal impatience in a monetary growth model. The Institute of Social and Economic Research, Osaka University, Discussion Paper No.622

Hirose K, Ikeda S (2008) On decreasing marginal impatience. Jpn Econ Rev 59:259–274

Hirose K, Ikeda S (2012a) Decreasing marginal impatience in a two-country economy. J Econ 105:247–262

Hirose K, Ikeda S (2012b) Decreasing and increasing marginal impatience and the terms of trade in an interdependent world economy. J Econ Dyn Control 36:1551–1565

Ikeda S, Ohtake F, Tsutsui Y (2005) Time discounting, hyperbolic discounting, and intertemporal choices: evidences from experiments and questionnaires. ISER Discussion Paper No. 638, Osaka University (in Japanese)

Jafarey S, Park H (1998) The dynamics of optimal wealth distributions with recursive utility. Econ Lett 61:149–158

Kapteyn A, Teppa F (2003) Hypothetical intertemporal consumption choices. Econ J 113:140–152

Lawrance EC (1991) Poverty and the rate of time preference: evidence from panel data. J Politi Econ 99:54–77

Leu RE, Buri S, Priester T (1997) Lebensqualität und Armut in der Schweiz. Haupt, Bern

Lucas R, Stokey N (1984) Optimal growth with many consumers. J Econ Theory 32:139–171

Nath HK, Sarkar J (2006) Diminishing marginal impatience: its promises for asset pricing. Appl Financ Econ Lett 2:61–64

Obstfeld M (1982) Aggregate spending and the terms of trade: is there a Laursen–Metzler effect? Q J Econ 97:251–270

Obstfeld M (1990) Intertemporal dependence, impatience, and dynamics. J Monet Econ 26:45–76

Ogaki M, Atkeson A (1997) Rate of time preference, intertemporal elasticity of substitution, and level of wealth. Rev Econ Stat 79:564–572

Ogawa K (1993) Economic development and time preference schedule the case of Japan and East Asian NICs. J Dev Econ 42:175–195

Pender JL (1996) Discount rates and credit markets: theory and evidence from rural India. J Dev Econ 50:257–296

Read D, Read NL (2004) Time discounting over the lifespan. Organ Behav Hum Decis Process 94:22–32

Ryder HE, Heal GM (1973) Optimal growth with intertemporally dependent preference. Rev Econ Stud 40:1–33

Samwick AA (1998) Discount rate heterogeneity and social security reform. J Dev Econ 57:117–146

Sarkar J (2007) Growth dynamics in a model of endogenous time preference. Int Rev Econ Financ 16:528–542

Schumacher I (2009) Endogenous discounting via wealth, twin-peaks and the role of technology. Econ Lett 103:78–80

Svensson LEO, Razin A (1983) The terms of trade and the current account: the Haberger–Laursen–Metzler effect. J Politi Econ 91:97–125

Tanaka T, Camerer CF, and Nguyen Q (2010) Risk and time preferences: linking experimental and household survey data from Vietnam. Am Econ Rev 100:557–571

Tsutsui Y, Ohtake F, Ikeda S (2005) The reason why you are unhappy. ISER Discussion Paper No. 630, Osaka University (in Japanese)
Uzawa H (1968) Time preference, the consumption function and optimum asset holdings. In: Wolfe JN (ed) Value capital and growth: papers in honour of Sir John Hicks. Aldine, Chicago
Ventura L (2003) Direct measures of time preference. Econ Soc Rev 34:293–310
Wang M, Rieger MO, Hens T (2011) How time preferences differ: evidence from 45 countries. Norwegian School of Economics Discussion Paper (ISSN: 1500-4066)

Part IV
Bubbles and Crash

Chapter 12
Why Did the Nikkei Crash? Expanding the Scope of Expectations Data Collection

Robert J. Shiller, Fumiko Kon-Ya, and Yoshiro Tsutsui

Abstract Why did the Japanese stock market lose most of its value between 1989 and 1992? To help us answer this and related questions, we have collected parallel time series data from market participants in both Japan and the United States 1989–1994 on their expectations, attitudes, and theories. Substantial variability within countries through time in these data and, notably, dramatic differences across countries in expectations were found. While no unambiguous explanation of the Japanese crash emerges from the results, we do find a clear relation of the crash to changes in Japanese price expectations and speculative strategies.

Keywords Bubble crash • Nikkei • Investor behavior

JEL Classification Codes G02

1 Introduction

The Nikkei stock price average in Japan, after rising dramatically through the 1980s, fell from 38915.9 on December 29, 1989 to 14309.4 on August 18, 1992, a decline of 63.2 % (see Fig. 12.1). In real terms, using the Japanese consumer price index

The original article first appeared in The Review of Economics and Statistics, 78(1): 156–164, 1996. A newly written addendum has been added to this book chapter.

R.J. Shiller
Sterling Professor of Economics, Yale University, 30 Hillhouse Avenue, New Haven, CT 06520, USA
e-mail: robert.shiller@yale.edu

F. Kon-Ya

Y. Tsutsui (✉)
Faculty of Economics, Konan University, 8-9-1 Okamoto, Hyogo, Kobe 658-8501, Japan
e-mail: tsutsui@center.konan-u.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_12

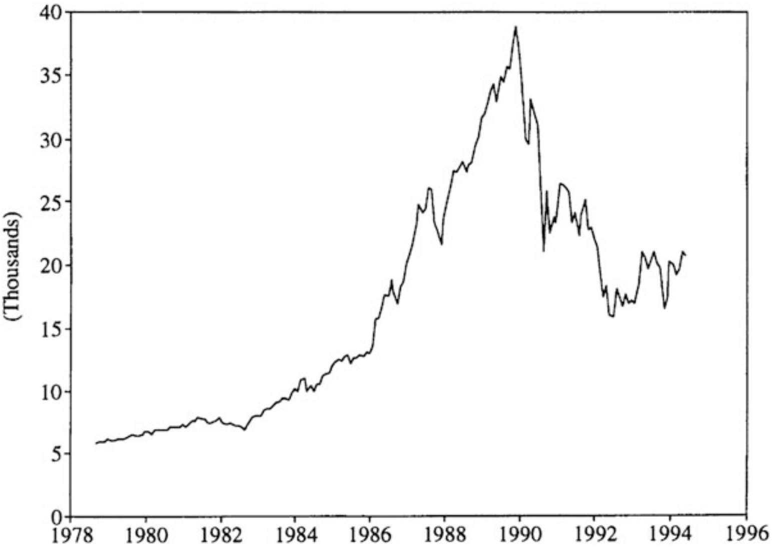


Fig. 12.1 Nikkei 225 stock price average, end of months, Sept. 1979 to June 1994 (Source: Nikkei Shinbun)

to correct for inflation, the decline between these two dates was 65.8 %. This stock market crash was not worldwide; in the United States over the same interval of time stock prices rose. Despite the magnitude and importance of the drop in the Nikkei, we know nothing solid about the origins of this event. Data about fundamentals of the Japanese economy provide no unambiguous reason for the crash. Thus, the Nikkei crash must have taken the form of a change in expectations or attitudes, about which there is little concrete to say beyond the fact that the Nikkei dropped.

The Nikkei crash is examined here as a study for the development of research methods that can give us a better understanding of such events. We report here on our collection of detailed time series data in Japan and the United States on expectations and understanding of speculative markets, before, during and after the crash of the Nikkei. We began our study before the crash partly because of a conjecture (expressed by some observers of the Tokyo market) that a crash might happen there. The questions for which we produced time series data on answers are unusual, and, we think, suggest some new methodology for studying financial markets. Some of our questions are intended to produce detailed accounts of expectations, over various horizons including long-term horizons. Other questions posed to our respondents in the surveys are of a rather more interpretive nature than are questions in most surveys, for example, questions about their speculative motives for holding stocks or their expectations about what would happen in the market *if* something else happened. All data are collected on a consistent basis about these expectations through time and across countries.

Time series data, data collected on a consistent basis at regular intervals for an extended period of time, are of fundamental importance to statistical analysis. Any such long systematic time series can be analyzed in connection with all other time series that are available over the same period. Experience with time series data, and a consensus on their meaning, develops gradually as the data series are extended.[1]

We do not expect to be able to offer a good understanding of the sources of the Nikkei crash from an analysis of the short (less than 5-year's span) time series we have produced for Japan and the United States. Our primary objective here is to establish that various expectations and attitudinal variables were changing over the time, and that the Japanese variables departed substantially from the corresponding variables measured in the United States, where the stock market behavior was quite different. We will also, however, offer some tentative interpretation of the Nikkei crash with the benefit of our data.

2 A Preliminary on Fundamentals in Japan

The crash in the Nikkei was followed by a sharp drop in the earnings of the constituent companies in Japan, so that the price-earnings ratio based on results rose, despite decline at the time of the crash in the Nikkei, in 1994 well above pre-crash levels: see Fig. 12.2. It is natural to hypothesize, then, that the crash in the Nikkei was due to new information about the outlook for earnings, information hitting the market before the actual drop in earnings. This simple hypothesis, however, may not be entirely satisfactory. The price-earnings ratio based on expected earnings (see also Fig. 12.2) declined about as much as the price-earnings ratio based on results between the peak and trough of the market.[2] There was virtually no decline between the end of 1989 and the end of 1990, a time interval during which most of the decline in the Nikkei occurred in 1-year-ahead forecasted earnings in Japan as compiled by I/B/E/S Inc.[3]

From publicly available data, we do not know whether market participants were reacting to information in 1990 about a less encouraging long-run outlook for earnings. We also do not know whether market participants were thinking in 1991 and 1992 that the decline in earnings since the crash is expected to be reversed, and that it was a temporary business-cycle-related decline that may not last more than a few years. If this was their expectation at the time, then the earnings decline would not appear adequate to explain a major crash in prices. Note that the sharp earnings declines reported in Japan near the end of our sample resulted in the sharp run up of price-earnings ratios in 1994, rather than yet another large drop in prices.

[1] In contrast, the post-event studies of stock market crashes that are typically conducted after the fact have relatively little power to discover what was changing importantly at the time of the crash.

[2] The Nikkei Shinbun price-earnings ratio based on expected earnings is an average across firms of price-earnings ratios, where the denominator of the ratio for each firm is expected earnings as reported by the firm itself. The horizon of these expectations differs across firms.

[3] See *Wall Street Journal*, March 17, 1994.

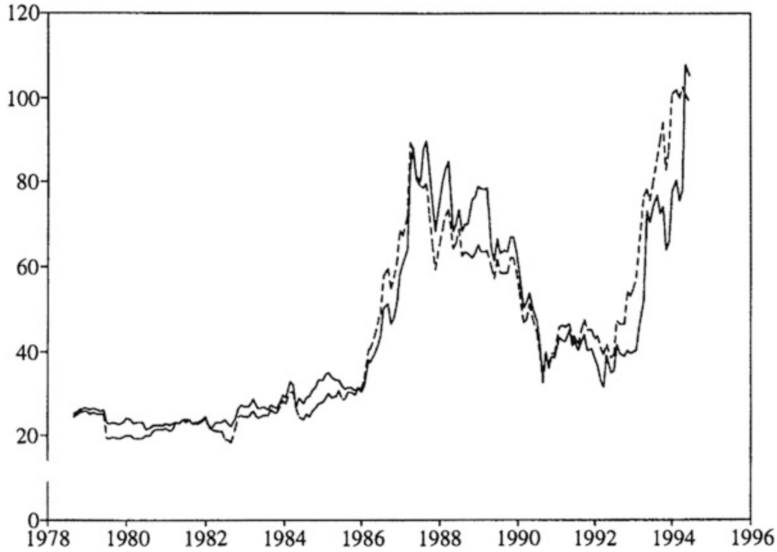


Fig. 12.2 Price-earnings ratio of Tokyo Stock Exchange 225 stocks, based on results (*solid line*) and based on expectations (*dashed line*), monthly, Sept. 1978 to June 1994 (Source: Nikkei Shinbun)

Movements in the stock markets of the world are not tightly related to earnings movements.

Of course, we do not deny that fundamentals play an important role in forming the level of the Nikkei. It is easy to count up facts that are consistent with the movement of the Nikkei for a limited period. It is hard, however, to find those which are consistent throughout a long period.

For example, the rise of Japanese long-term interest rates from July 1989 to September 1990 may be pointed out as a suspect in the crash. The rise is reflected in the consecutive increases in the discount rate from 2.5 % in May 1989 to 6 % at the end of August 1990. Thus, one might argue that the change in the attitude of the Bank of Japan toward a tight monetary policy is a cause of the crash.[4] However, the fact does not explain why the Nikkei continued rising sharply during 1989 despite the rapid rise of the interest rates, and why the crash began at the beginning of 1990. Historically, stock markets do not show any consistent behavior in response to sudden tightening of monetary policy; note for example, that the sudden tightening in monetary policy in the United States in 1994, roughly comparable in magnitude to the tightening in Japan in 1989–1990, produced no overall U.S. stock market decline.

[4]Ueda (1992) expresses this view.

3 Existing Time Series Data for the Japanese and United States Stock Markets

Few time series data are collected regarding stock market expectations. Governments are the main provider of high- quality time series data on an uninterrupted and inter-temporally consistent basis. Yet the Japanese and U.S. governments apparently collect no such series on expectations in the financial markets. In the industry, there are some attempts to collect time series data on stock market expectations, but none of these attempts matches the scope of our study.

In Japan, there appears to be only one published price expectations survey. The *Nikkei Financial Daily* reports every Saturday the results of a survey of five securities companies, three banks, seven institutional investors and three foreign companies, in which are given the number of respondents who expect that the markets will be more bullish, more bearish, or neutral compared with the current week. This is their only published expectations question, the number of respondents is quite small, and their time series goes back only to October, 1987. The Quick Research Corporation has been sending a questionnaire to about 300 securities companies and institutional investors in Japan every month since April 1994; they ask about 1-, 3- and 6-month ahead expectations for the Nikkei average. Their results are reported to subscribers by fax, but have not been published yet.

For the United States, there is the very long time series data, extending back to 1952, of Livingston, which is analyzed by De Bondt (1991). Livingston asked his panel of about 40 economists to forecast the Standard and Poor Index at horizons of 7 and 13 months. From the early 1980s and until its bankruptcy, Drexel, Burnham Lambert tabulated the results of a few expectations questions about the stock market under the direction of Richard Hoey. For the past 6 years, Money Market Services, Inc. of New York has collected 1-week and 1-month expectations for the Dow Jones Industrial Average and for the Standard and Poor Composite Index. All of these are surveys of experts only, not intended to be surveys of market participants. The American Association of Individual Investors has been sending out for the past few years weekly postcard questionnaires to their members, inquiring about their opinion as to the outlook for the market. As far as we have been able to determine, existing surveys ask only a few questions about the market, and do not try to devise batteries of questions that get at the reasons for market behavioral patterns.

4 Our Surveys

We tabulate here responses in both Japan and the United States in a number of mail surveys we conducted from 1989 to 1994. We created a biannual series of answers; questionnaires were mailed roughly every 6 months. For the Japanese sample, we mailed to almost all of the major Japanese financial institutions, which consist of 165 banks, 46 insurance companies, 113 securities companies, and 45 investment trust companies.[5] No non-financial corporations are included in the sample. The U.S. institutional investors were selected at random each time from the section "Investment Managers" from the *Money Market Directory of Pension Funds and their Investment Managers* (McGraw Hill). In each mailing, about 400 questionnaires were sent, yielding responses from about a third. Mailing dates in Japan were July 3, 1989 (1989-II), November 9, 1989 (end 1989), March 6, 1990 (1990-I), August 10, 1990, February 2, 1991, September 9, 1991, March 27, 1992, September 11, 1992, March 19, 1993, August 4, 1993 and February 28, 1994. First mailing dates in the United States were July 5, 1989, January 17, 1990, July 27, 1990, January 31, 1991, August 20, 1991, January 31, 1992, August 20, 1992, February 12, 1993, August 6, 1993, and February 28, 1994. In the United States, a second questionnaire and letter were sent out three weeks after the first mailing to those who had not responded yet.

In all but the 1989-II and 1990-I questionnaires the first portions of the questionnaires, which included the questions reported here, were nearly identical both through time and across the two countries, except, of course, for translation into English or Japanese. The responses thus enable us to make accurate comparison across countries and through time.

4.1 Questions About Expectations

We asked respondents to give forecasted changes in the Nikkei 225 (Nikkei Dow) and the Dow Jones Industrial Average for horizons of 3 months, 6 months, 12 months, and 10 years. The question on the questionnaires was

> I-1,2 "How much of a change in percentage terms do you expect in the following (use + before your number to indicate an expected increase, a - to indicate an expected decrease, leave blanks where you do not know): [FILL IN ONE NUMBER FOR EACH]"

After this question there were spaces to fill in the expectations for the various horizons and the two countries. The mean answers for the 1-year horizon are shown in Table 12.1; expectations in both countries for both countries are presented. The results confirm that the expectations do change through time both for the United States and Japan; the F-statistics (Table 12.1) for the null hypothesis of constancy through time of expectations are all highly significant.

We also see in the answers to the Table 12.1 questions confirmation that there are striking differences between U.S. and Japanese expectations, even for the same markets. The Japanese were uniformly more optimistic in their short-run expectations for the Japanese market than were the Americans. At a horizon of 1 year, there was usually a spread on the order of 20 % points between the Japanese and U.S. forecasts for the Japanese market; the spread was never less than 10 %

[5]These numbers vary slightly over time; the numbers given are for 1989-II and 1992-I surveys.

Table 12.1 Expectations questions

A. Expectations for Japanese economy				
		I–1	I–2	I–3
Date	Nikkei 225 index at time of survey	Japanese expected 1-year growth in Nikkei index (%)	U.S. expected 1-year growth in Nikkei index (%)	Japanese 10-year expected Japanese corporate earnings (annual rate) (%)
1989-II	33631	9.49	−7.67	5.02
1989 end	35894	13.02	–	–
1990-I	32616	10.84	−9.14	–
1990-II	26490	8.22	−8.76	5.01
1991-I	24935	19.33	0.94	4.68
1991-II	23332	18.36	−2.52	4.25
1992-I	18436	20.85	0.33	3.95
1992-II	18066	27.69	6.47	4.65
1993-I	19048	14.08	3.22	4.76
1993-II	20322	15.85	1.02	3.64
1994-I	20091	16.27	1.34	3.70
Test of time constancy:		$F(10,1237)=10.82$	$F(9,687)=9.19$	$F(8,1045)=6.19$
	$p=$	8.29×10^{-18}	1.06×10^{-13}	7.87×10^{-8}
B. Expectations for United States economy				
		I–1	I–2	I–3
Date	Dow Jones Industrial Average at Time of Survey (DJIA)	Japanese expected 1-year growth in DJIA (%)	U.S. expected 1-year growth in DJIA (%)	U.S. 10-year expected growth in U.S. corporate earnings (annual rate) (%)
1989-II	2554	8.48	3.49	5.57
1989 end	2553	12.57	−0.26	5.16
1990-I	2716	4.28	1.65	4.63
1990-II	2902	11.26	6.17	5.02
1991-I	3043	8.55	7.82	5.52
1991-II	3245	3.41	6.51	5.68
1992-I	3257	0.89	4.49	2.50
1992-II	3343	0.35	2.01	5.50
1993-I	3579	0.83	0.56	4.98
1993-II	3831	0.88	2.75	5.56
1994-I	2554	8.48	3.49	5.57
Test of time constancy:		$F(9,961)=14.53$	$F(9,1154)=4.65$	$F(9,1315)=13.36$
	$p=$	0.00	4.53×10^{-6}	1.19×10^{-20}

Note: Index values are for close of first market day 10 or more days after first mailing date for questionnaire. F-statistics test null hypothesis that values are constant through time points.[6] There is a strong correlation between the U.S. and Japanese forecasts for the Nikkei, the correlation coefficient between the average answers for questions I-1 and I-2 for the Nikkei as shown in Table 12.1 is 0.83. Respondents in both countries became relatively optimistic or pessimistic at about the same time, but there was always the enormous spread between their expectations.

What can we make of the stunning differences between the expectations in the two countries for the Nikkei? Investors on both sides of the Pacific Ocean have access to much of the same information, and they can talk to each other, they can listen to each others' pundits. Why should their expectations differ depending on which country is their home? Perhaps the difference has something to do with personal daily talk among investors or with some irrationality related to patriotism or wishful thinking; see Shiller (1995).

These remarkable differences in expectations between U.S. and Japanese respondents have some potential use in explaining other puzzles. Consider, for example, the puzzle posed by French and Poterba (1990), that there is very little cross-border stocks investment between the United States and Japan. Our results suggest a possibly simple explanation: investors in each country are relatively more optimistic about the stock market in their own country. For another example, consider the Feldstein-Horioka (1980) puzzle that aggregate investment in each country tends to be highly correlated with aggregate savings in that country; that people may be optimistic about their own country certainly must be relevant to understanding that puzzle. More research could be done to establish the potential validity of such notions, if longer time series become available.

We also asked for expected long-term earnings growth rates. The question was:

I-3 "What do you think the rate of growth of real (inflation adjusted) corporate earnings will be on average in the US over the next 10 years?
Annual percentage rate: ______%"

The 10-year horizon was chosen as a proxy for the kind of long-term expectations for earnings growth that are thought to influence price-earnings ratios. Asking directly for long-term expectations represents a significant new departure. In studying the reasons for high Japanese price-earnings ratios, French and Poterba (1991), lacking our data, used forecasted 10-year growth rates for Japanese gross national product provided by a single forecasting company; our survey data are a much more direct measure of the relevant expectations.

We see a fairly steady decline since 1989-II in these long-run expected growth rates in Japan (Table 12.1). Such a gradual decline, other things equal, might be expected to have produced a correspondingly gradual decline in price-earnings ratios in Japan.

[6] At a horizon of ten years, on the other hand, there was much less discrepancy between the Japanese and U.S. forecast for the Nikkei and in the most recent survey it was the U.S. respondents who were more optimistic about this long-run outlook for the Nikkei.

It should be noted that many researchers feel that the expectations data collected by surveys such as these are by necessity inferior to expectations inferred or derived from market prices. Consider, for example, the expectations for future stock price index changes that can be inferred from prices in the stock index options markets. It is possible to infer from options prices not only implied variances of price changes but also implied skewness of subjective distributions of price changes. There are thus, in market prices, implicit expectations of the probabilities of a market decline. Thus, for example, Bates (1991) was able to analyze whether the stock market crash of 1987 was expected. One might think that these probabilities or market expectations are inherently better than probabilities or expectations that people write down on survey forms. People who will go so far as to take a position in an options market are likely to think more carefully about the probability of a crash; their judgment is considered rather than hasty. Moreover, the sample size, the number of people whose expectations have an impact on the implied volatility, is enormously greater with the implied volatilities than with the survey data. When dealing with an entire options market, then, the results may in fact be considered not a sample at all, but the universe for that market.

In fact, however, these arguments that the implied volatilities or other market-derived expectations data are the final word on actual public expectations disregard the fundamental sociological fact that the expectations that are relevant for market behavior diffuse across different subpopulations of the investing public at different rates, and that attention of certain subpopulations shifts from one market to others. Surely, the prices in the options markets reflect the considered opinions of all people who are currently trading in these markets, but these people are hardly, by any stretch of the imagination, a random sample of all people who might sell stocks at the time of crash. Suppose we are interested in a theory of a crash wherein a small price drop acts as a trigger for a stock market crash, so that people, fearing a crash, thereby produce the very crash they feared. With such a theory, we would generally expect that most of these people may never have given careful consideration to the probability of a crash, are not closely involved with options markets and many may even have inconsistent or wrong theories of these markets. We will not know what they are thinking unless we ask, and the opportunity is lost forever if we wait beyond the length of people's short-term memories, or until after a major event that changes their patterns of thinking.

4.2 Qualitative and Scenario Questions

Our qualitative and scenario questions were questions aimed to be more in the mode of thinking of individual market participants, worded in everyday language. The hope was to pose questions in such a way that the questions represent categories of thought already in many respondents' thinking, not questions that would be difficult to answer. Katona (1975) argued, based on years of survey research, that most people do not have expectations for economic variables, and are forced to construct the expectations when surveyors ask for their expectations. Asking for their expectations may be a useful exercise, but it may sometimes fail to reveal people's concerns and understandings. We want now to know how our respondents interpret market phenomena, not to try to construct forecasts for us. We are applying here to economics the basic concepts of interpretative social science (Rabinow and Sullivan 1979), that stresses the importance in explaining human behavior of people's own interpretations of events.[7]

We asked, in questions II-1 and II-2, whether the market is overpriced, that is, high relative to fundamental value.

> II-1. "Stock prices in Japan, when compared with measures of true fundamental value or sensible investment value are: 1. Too low. 2. Too high. 3. About right. 4. Do not know."
>
> II-2. "Stock prices in the United States, when compared with measures of true fundamental value or sensible investment value, are: 1. Too low. 2. Too high. 3. About right. 4. Do not know."

These questions were included because we learned that the concept of an overpriced market was very much on people's minds at the time of the stock market crash of October 1987. At the time of this crash, when investors in the United States and Japan were asked in a questionnaire survey to explain the cause of the crash in their own words, and the responses coded, the most important theme in their answers was that the market was overpriced (Shiller 1989; Shiller et al. 1991).

Table 12.2 gives the proportion of respondents choosing answer 2 (too high) in each survey. We see here that the U.S. investors were consistently more likely to think that the market prices are too high, and were dramatically more likely to think this about the Japanese market. In 1989-II, 73.5 % of U.S. respondents thought the Japanese market was overpriced, while only 26.6 % of the Japanese did. Most Japanese became temporarily of the opinion that their market was too high right after the Japanese market had its spectacular 4.5 % drop on February 26, 1990: the 1990-I survey of Japanese investors (before most of the dramatic downturn in the Nikkei had occurred) shows that 61.1 % of them felt that the Japanese market was overpriced. But in 1990-II, a comparison of the United States and Japanese responses after most of the enormous decline in the Tokyo stock market and after the Iraqi oil crisis shows a return to nearly the same pattern as in 1989-II, with Americans strongly tending to think that the Japanese market is overpriced and the Japanese respondents again dramatically less likely to think so.

A common element in the popular notion of a speculative bubble is that during the expansion phase, or bull market, increasing numbers of investors are buying stocks because they think that prices will go up for a while longer, and hope to exit before the bubble bursts. Conversely, a bear market may be caused by increasing

[7] This is the first step that Sternberg (1987), in his proposed methodology for implicit theories research, called "behavioral listings." He, of course, expects his method to be applied to subjects in a psychology laboratory, not to the world financial markets; it is easier for psychologists to obtain large enough quantities of data to make a rapid transition to his second step of "prototypical analysis," where the popular theories and models are fleshed out.

Table 12.2 Qualitative and scenario questions

		II-1 (2)	II-2 (2)	II-3 (1)	II-4 (1)	II-5 (1)	II-6 (2)	II-7 (1)	II-8
Date	Stock price index	Stock prices too high Japan (%)	Stock prices too high U.S. (%)	Advise stocks now despite expected drop (%)	Advise against stocks despite expected rise (%)	See excitement about stocks (%)	Trend last 6 months was speculative (%)	If prices dropped 3 % would expect rise next day (%)	Probability of crash next 6 months (%)
A. Answers from Japanese respondents									
1989-II	33631	26.6	0.0	39.1	23.7	37.2	14.9	42.8	14.6
1989 end	35894	32.1	9.4	–	–	–	–	–	13.7
1990-I	32616	61.1	0.8	–	–	–	–	–	–
1990-II	26490	21.3	11.1	7.3	55.3	41.3	38.2	29.1	31.7
1991-I	24935	16.8	10.4	9.8	35.8	34.4	26.4	28.1	18.6
1991-II	23332	13.9	19.2	14.0	23.1	23.7	25.4	39.7	19.7
1992-I	18436	22.5	36.6	7.0	62.0	28.7	22.5	20.8	28.1
1992-II	18066	11.7	32.4	11.2	39.4	25.0	33.3	22.5	27.9
1993-I	19048	33.3	31.0	15.5	23.6	41.9	24.3	39.1	20.1
1993-II	20322	38.5	33.9	17.6	18.4	30.0	16.9	37.2	17.4
1994-I	20091	30.4	33.5	19.3	20.3	27.7	14.6	33.8	15.8
Test time constancy:		$\chi^2(10) = 118.2$	$\chi^2(10) = 167.8$	$\chi^2(8) = 73.5$	$\chi^2(8) = 112.8$	$\chi^2(8) = 21.7$	$\chi^2(8) = 40.13$	$\chi^2(8) = 26.19$	$F(9,1322) = 8.35$
$p =$		1.16×10^{-20}	7.75×10^{-13}	9.96×10^{-13}	1.01×10^{-20}	5.41×10^{-3}	3.02×10^{-6}	9.75×10^{-4}	3.38×10^{-12}

(continued)

Table 12.2 (continued)

Date	Stock price index	II-1 (2) Stock prices too high Japan (%)	II-2 (2) Stock prices too high U.S. (%)	II-3 (1) Advise stocks now despite expected drop (%)	II-4 (1) Advise against stocks despite expected rise (%)	II-5 (1) See excitement about stocks (%)	II-6 (2) Trend last 6 months was speculative (%)	II-7 (1) If prices dropped 3 % would expect rise next day (%)	II-8 Probability of crash next 6 months (%)
B. Answers from United States respondents									
1989-II	2554	73.5	18.7	34.4	24.6	55.5	19.1	33.3	14.9
1990-I	2553	81.0	37.9	16.0	86.3	41.1	41.2	34.8	22.0
1990-II	2716	82.6	39.2	11.1	53.7	43.5	36.9	18.6	23.7
1991-I	2902	67.2	35.4	26.4	34.7	54.8	36.9	22.9	17.3
1991-II	3043	71.0	47.1	17.6	38.4	44.1	21.1	36.2	14.4
1992-I	3245	65.9	46.6	19.2	32.3	48.3	14.8	37.9	19.6
1992-II	3257	54.8	44.4	12.3	44.9	45.9	18.1	31.4	19.7
1993-I	3343	55.7	42.1	27.5	32.8	54.1	18.8	29.5	20.3
1993-II	3579	55.2	42.5	30.7	22.0	45.2	13.2	37.0	20.8
1994-I	3831	55.9	42.4	19.2	42.7	50.8	21.0	33.6	16.2
1994-I	20091	30.4	33.5	19.3	20.3	27.7	14.6	33.8	15.8
Test time constancy:		$\chi^2(9) = 61.59$	$\chi^2(9) = 38.33$	$\chi^2(9) = 45.35$	$\chi^2(9) = 170.14$	$\chi^2(9) = 13.37$	$\chi^2(9) = 72.06$	$\chi^2(9) = 23.14$	$F(9,1393) = 3.38$
$p =$		6.61×10^{-10}	1.52×10^{-5}	7.95×10^{-7}	5.76×10^{-7}	0.15	6.00×10^{-12}	5.90×10^{-3}	4.14×10^{-4}

numbers of investors who think that the market will continue to go down for a while, and who are waiting for the recovery to enter the market. It is not obvious how to prove whether our respondents are thinking this way. The questions discussed in the preceding section about expectations at various horizons might reveal such thinking if the horizons asked about match-up with the dates at which the market is expected to turn, but we will probably not be so lucky as to choose the right horizons to ask about. We cannot ask for expectations at all horizons without exhausting respondents. Moreover, when asked to forecast the stock price index at a number of horizons, respondents may not even register their opinions about market dynamics: it may be too hard for them to translate their opinions into numbers. People may give us conventional or safe forecasts, even if they are themselves invested in thinking about market turns. People may have complicated vague impressions about the outlook for the market, even impressions that put them into two minds about the market, so that they may give different-sounding answers to similar questions that are posed differently.

A more interpretive method for deriving evidence on this speculative behavior can be had by asking whether respondents would advise staying in the market for the time being, even though they expect the market to drop, and conversely. Without specifying the horizon of the associated forecasts, we allow the respondent to reveal directly whether he or she is thinking in terms of short-term speculative advantage. Respondents were asked about their own countries, questions II-3 and II-4:

> II-3 "Although I expect a substantial drop in stock prices in [the US, Japan] ultimately, I advise being relatively heavily invested in stocks for the time being because I think that prices are likely to rise for a while. 1. True 2. False 3. No Opinion"

> II-4 "Although I expect a substantial rise in stock prices in [the US, Japan] ultimately, I advise being less invested in stocks for the time being because I think that prices are likely to drop for a while. 1. True 2. False 3. No Opinion"

These questions, in contrast to the expectations questions displayed above, are directly connected with investing strategy, and the stress on investing strategy in these questions may call forth a different type of expectation. These questions have been criticized as too long and too complicated; when a respondent answers "False" to II-3 we do not know whether a decline is not expected or whether a decline is expected but stocks are not thought likely to rise for a while. People who criticize our questions along these lines seem to be assuming that the question is designed to elicit well-defined expectations, while in fact the question is designed to discover whether respondents are familiar with a sort of popular theory. We worked a great deal on the wording of this question, but could not find a better way to ask respondents about their bubble-enforcing attitudes. (We did ask them too about the date of the presumed peak or trough in the market, to allow them more precision in answering.)

The proportions choosing answer 1 are shown in Table 12.2. It is striking that quite often most of both the U.S. and Japanese respondents answered "true" to one of questions II-3 or II-4. Thus, in a sense, most of our investors appear to be either relatively in the market hoping to get out before it drops or relatively out of the market hoping to get in before it rises, suggesting that the market is indeed a very "bubbly" place. The answers also reveal that strategies differed very much among investors; suggesting the importance of thinking about heterogeneity among investors. Of course, the tendency to answer "true" may be exaggerated by selection bias: those who have striking views about the outlook for the market may be more likely to fill out our questionnaire.

In the answers to these questions, we do see a change in the behavior of Japanese investors before and after the debacle in Japanese stock prices. Between 1989-II and 1990-II, when most of the Nikkei crash occurred, we see dramatic changes in the Japanese answers to these equations; there was substantially less evidence of a positive bubble mentality, as indicated by fewer "True" answers to II-3 later. This evidence is consistent with the notion that the Japanese stock market debacle might have been caused by changed short-run expectations for prices.

Question II-5 was directed at learning directly about a concomitant of the kinds of speculative booms that were widely reported about the booms preceding the 1929 crash and other booms: just that people seemed to be very excited about stock market investing:

> II-5 "Many people are showing a great deal of excitement and optimism about the prospects for the stock market in the [United States, Japan] and I must be careful not to be influenced by them. 1. True. 2. False. 3. No opinion."

That people were getting excited about investing is so much a part of the story people tell of these booms; if people are getting excited, one might think they would know it and could report it to us. The proportions of respondents who answered "True" about their own country are shown in Table 12.2. Time variation shows no clear relation in Japan to the Nikkei crash; moreover, our rejections of the null hypotheses that the proportions are constant through time are least significant for this question, when compared with all other questions we report here (see the χ^2 statistics in Table 12.2). Of course, the lack of relation of this answer to the Nikkei crash and lack of statistical significance may be because of the words "I must be careful not to be influenced by them." Some respondents may have answered "false" even when they agree with the former part of the question because they do not agree with the later part.

Question II-6 asked respondents whether the trend in stock prices over the past 6 months was due to fundamentals or to investor psychology:

> II-6 "What do you think is the cause of the trend of stock prices in [the United States, Japan] in the past six months? 1. It properly reflects the fundamentals of the U.S. economy and firms. 2. It is based on speculative thinking among investors or overreaction to current news. 3. Other 4. No opinion."

Respondents were asked about their own countries only. The proportions choosing response 2 in each country are given in Table 12.2. In Japan, the proportion selecting answer 2 was relatively high from 1990-II to 1993-I. This period corresponds approximately to the high proportion of the answers "too low" in question II-1 above in Japan. Thus, it is suggested that they think that the Nikkei became too low because of speculative thinking among the investing public in this period. In Japan, the percentage who chose, for II-6, answer 1 (fundamentals) was higher than the percentage who chose answer 2 (speculative thinking) at all times except for 1990-II, the time of the most rapid decline in the Nikkei shown in the tables. We should note that, based on our experience, investors seem to put much more importance on psychology when asked to explain big moves in short periods of time. Just after the biggest one-day stock market crash in history, October 19, 1987, 64 % of U.S. institutional investors (and 68 % of U.S. individual investors) (Shiller 1989) and 73 % of Japanese institutional investors (Shiller et al. 1991) thought that the crash was due to investor psychology. Just after the 6.9 % one-day drop in the Dow Jones Industrial Average on October 13, 1989, 77 % of U.S. investment professionals [8] and 83 % of Japanese institutional investors chose psychology as an explanation for the drop.

Question II-7 was phrased to get at a possibly time-varying parameter in a feedback mechanism that feeds past price movements into current changes in demand and hence into price movements, by asking how a past price change affects people's expectations for the future:

> II-7 "If the [Dow, Nikkei] dropped 3 % tomorrow, I would guess that the day after tomorrow the Dow would: 1. Increase. 2. Decrease. 3. Stay the same. 4. No opinion."

Table 12.2 shows the proportion in each country who chose "Increase;" respondents were asked about their own country only. We note the striking fact that the proportion expecting an increase was highest in Japan in 1989-II, right before the peak in the market.

Stock market crashes are often thought to be caused by a feedback mechanism, as initial price decreases engender pessimistic expectations and hence more price decreases, but if we hold such a theory we must explain why the feedback is not causing crashes every day. We would have an explanation if we understood how response patterns change through time. Changes in response patterns to price changes may be documented by changes in answers to this question. Our statistics show less significance in this sample than was the case with most of the other questions, but time variation in the proportion expecting to increase after an initial decrease was significant at conventional levels. This suggests that it may be useful to continue collecting such data. Of course, much more research is needed to know how to interpret such feedback mechanisms. Further survey work should inquire about other technical theories and trading rules (such as those concerning resistance levels, moving averages, etc.) to see how feedback might change through time.

Question II-8 asks respondents for their subjective probability of a stock market crash:

> II-8 "What do you think is the probability of a catastrophic stock market crash, like that of October 28, 1929 or October 19, 1987, in the next six months? (An answer of 0 % means that it cannot happen, an answer of 100 % means it is sure to happen.) Probability: ______%"

[8] See Robert Shiller and William Feltus, "Fear of a Crash Caused the Crash," *New York Times*, October 29, 1989.

Such subjective probabilities have obvious relevance to any theories that stock market crashes are caused by fears of crashes. Fear of a crash was at its highest (see Table 12.2) in Japan in our survey immediately after the most precipitous drop in the Nikkei, 1990-II. This fact seems to be consistent with the notion that the Japanese investors think the Nikkei became too low by speculative thinking in these periods, as argued above.

Time variation in the answers to all questions except II-5 is highly significant in both countries. There is even highly significant time variation in both countries in answers to question II-8 about the risk of a sudden crash in this sample period when there was no important one-day stock market crash.

5 Why Did the Nikkei Crash?

Our objective here was partly to illustrate a methodology that might allow us to understand events like the Nikkei crash, and to demonstrate the variability through time of the expectations and other parameters we assessed. Our surveys cannot be expected to provide a complete understanding of the causes of the crash in the Nikkei. A complete understanding cannot be obtained without first explaining such mysteries as the cause of the run-up of the Nikkei before 1989, or the Japanese tendency for very high (by world standards) price-earnings ratios; our surveys were not designed to elucidate such matters. Nor do our surveys enable us to evaluate the ultimate reasons why expectations and attitudes changed through time, or the role in these changes of all of the factors the media have stressed in connection with the crash, such things as expectations of the recession that depressed Japanese corporate earnings after the crash in the Nikkei, the increasing value of the yen, and policy actions of the Bank of Japan and the Ministry of Finance.

But our results do give us information about the kinds of changes in expectations that were associated with the crash in the Nikkei. We found that Japanese expectations for long-run earnings growth (question I-3, Table 12.1) in Japan became gradually less optimistic over the period 1989–1994. The earnings growth expectations did not surge up in response to the decline in actual Japanese earnings after 1990, which suggests that our respondents did not view the decline in earnings as temporary. We did not directly ask whether respondents viewed the decline in earnings as temporary, and so it is hard to say what they were thinking on this matter when answering a question about long-run earnings growth; they may not have given long-run earnings growth from the low current base of earnings.[9] Still,

[9]In our 1994-II Japanese survey, conducted after this chapter was written, we asked for 3-year expectations in addition to the 10-year expectations in question I-3. The average annual expected real earnings growth was 7.57 % over the next three years, versus 3.88 % over the next ten years. This suggests that part of the earnings decline was thought of as temporary, to be reversed in a relatively short period.

our results may be regarded as consistent with the notion that the overall drop in the Nikkei, the drop between the peak of the market at the end of 1989 and today, might well be viewed as nothing more than a response to the decline in earnings that was viewed as essentially permanent. The simplest story of the Nikkei crash is that it is just another example of a market's overreaction to earnings: it has been documented before for the United States that much of the volatility of stock prices has this form, as if people often fail to see that earnings movements may be transient, and do not expect them to be in any sense mean reverting (see Shiller 1989; Barsky and De Long 1993).

Still, the rough story of prices overreacting to earnings does not explain everything. The earnings expectations data do not help us to explain the relatively sudden initial crash of the Nikkei itself, the crash that occurred between the peak of the market in 1989 and the end of 1990. What changed rather suddenly and strikingly at the time of the crash were speculative attitudes, attitudes towards price movements, not earnings growth or expectations of earnings growth.

The initial crash in the Nikkei between 1989-II and 1990-II was accompanied by substantial changes in speculative factors as documented in our questions. Questions II-3 and II-4 (Table 12.2) show marked changes between 1989-II and 1990-II in opinions about whether it is advisable to buy for the short run. In 1989-II we saw the greatest proportion ever, 39.1 %, of Japanese who thought that this was a time when it was advisable to buy only for the short run; 1 year later this proportion had dropped to 7.3 %. Over the same interval, the proportion who advised against stocks in the short run despite an expected rise went up from 23.7 % to 55.3 %. These changes in response to questions about short run speculation are important evidence for a speculative element in the Nikkei crash.

Just before the crash of the Nikkei, in 1989-II, we see in answers to II-7 the highest proportion ever, 42.8 %, of Japanese who thought that if prices dropped 3 % in one day then the market would rise the next day. This impression of stability for the market may have encouraged the high prices that the Nikkei reached just before the crash. By early 1992, this proportion had fallen in half, to 20.8 %. The relative lack of confidence in the resiliency of the market would seem to encourage downward feedback loops, where price declines encourage further price declines, and such loops may well have been part of the decline in the market.[10]

There was a sudden, sharp, upspike in 1990-I, just before the biggest one-semester decline in the Nikkei in our sample, in the proportion of Japanese respondents who thought that the market was too high (question II-1, Table 12.2). In 1990-II, the date of the questionnaire immediately after the biggest 6-month decline in the Nikkei, the highest proportion ever reported that they thought the trend in the last six months was speculative (question II-6, Table 12.2).

These results paint a picture of a speculation-induced initial crash, from 1989 to 1990, in Japan. Still, the picture is not entirely clear. We do not know to what

---

[10]For a discussion of the theory of feedback loops in price changes, and the implication of such theory for the serial correlation properties of price changes, see Shiller (1990).

extent it was information of some sort about future earnings that stimulated the initial crash; the information may have prompted changes in expectations for the behavior of the market even though there were little changes in expected earnings growth. We also cannot yet understand why answers to certain of our questions showed little relation to the crash.

One fact that tempers our willingness to interpret the Japanese results in relation to the Nikkei crash is that when one looks at U.S. data for the same time period, there are sometimes important changes in answers to questions, even though the U.S. market did not crash. For example, responses to questions II-3 and II-4 showed just as dramatic movements in the U.S. as they did in Japan between 1989-II and 1990-II, even though the United States market experience was relatively uneventful. This result should help clarify why it is important to collect parallel time series in different countries.

On the other hand, it is in the comparisons with the United States that we see the most striking evidence that something crudely speculative was at work in driving the Nikkei. It is hard to imagine how we can reconcile the fact that those in Japan usually thought that the Nikkei would rise in the next year about 20 % more than those in the United States thought it would with any rational expectations model of the stock market. *Somebody* was exhibiting bad judgment if opinions differed so strikingly depending on where one sits.

Acknowledgments This research was supported by the Economics of Information and Risk Research Fund of Osaka University, the Japan Securities Research Institute, the Russell Sage Foundation and the U.S. National Science Foundation. This chapter is a revision of National Bureau of Economic Research Working Paper No. 3613. The authors wish to thank the respondents in the surveys for their participation, and Daniel Kahneman and Richard Thaler for their suggestions. Opinions expressed are those of the authors and not necessarily those of the supporting institutions.

Addendum: Was the Rise in American Stock Prices in 1990s a Bubble?[11]

In the text, we analyzed the crash and subsequent slump of the Nikkei in the early 1990s, utilizing the results of our survey until 1994. In this appendix, using longer results of the same survey we analyze whether the rapid rise in American stock prices in the late 1990s is a bubble.

In Fig. 12.3, we plot the quarterly data of the Dow Jones Industrial Average (DJIA) and Standard and Poor's 500 Index (SP500), normalizing their values as of the 1995Q1 to be 100. The indices rose gradually from 1990 to 1995, rose rapidly until 2000, and then declined until early 2002. The magnitude of the decline eventually reached about 30 % in the DJIA and about 45 % in the SP500, meaning

[11]This addendum has been newly written for this book chapter.

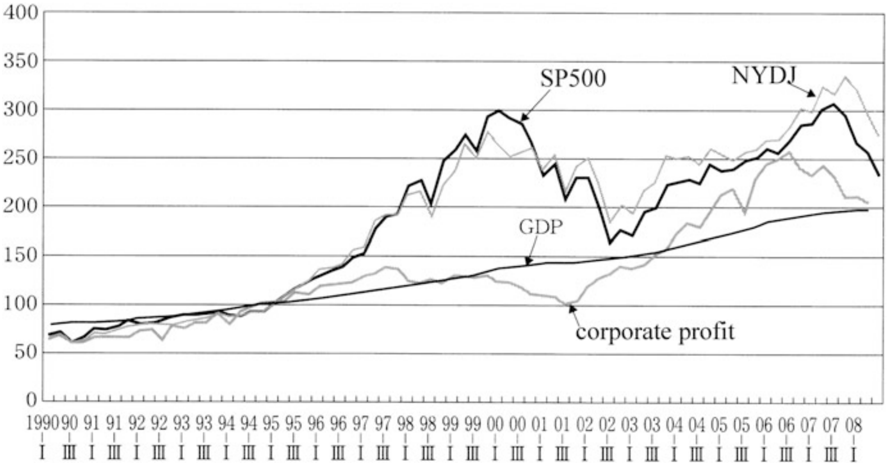


Fig. 12.3 Stock price indexes and fundamentals

that in 2 years they lost about half of the rapid gains they had made in the 5 years since 1995.

A bubble is often defined as the gap between an asset's price and its fundamental value. Thus, once we know the fundamental value, the size of the bubble is known. In Fig. 12.3, we also plot the GDP of the USA along with corporate profits. These generally kept pace with stock indices until 1995; a gap a gap then opened up between the two and grew until 2000, suggesting that the rapid rise in stock prices after 1995 may have been a bubble. Fundamental value, however, is the present value of the future earnings of a stock, not the current earnings. If investors have optimistic expectations for future earnings, the fundamental value is high even if the current earnings are low. Thus, we need information on investors' expectations of future earnings; our survey asks about these.

Our survey asks:

> What do you think the rate of growth of real (inflation adjusted) corporate earnings will be on average over the next 10 years?

We plot the result in Fig. 12.4. The average response from 1989 to 1994 was 5.35 %, and from 1995 to 1999 was 5.58 %, implying that expectations did not change much between the two periods. This suggests that the fundamental value did not change dramatically, so that the stock prices in the late 1990s contained a bubble.

Now, let us try to estimate the size of the bubble using some assumptions. Let's assume that the time discount rate r is constant, that stockholders are aware of all corporate earnings, and that stockholders expect that earnings will grow at a constant rate g. In this case, the fundamental value P is

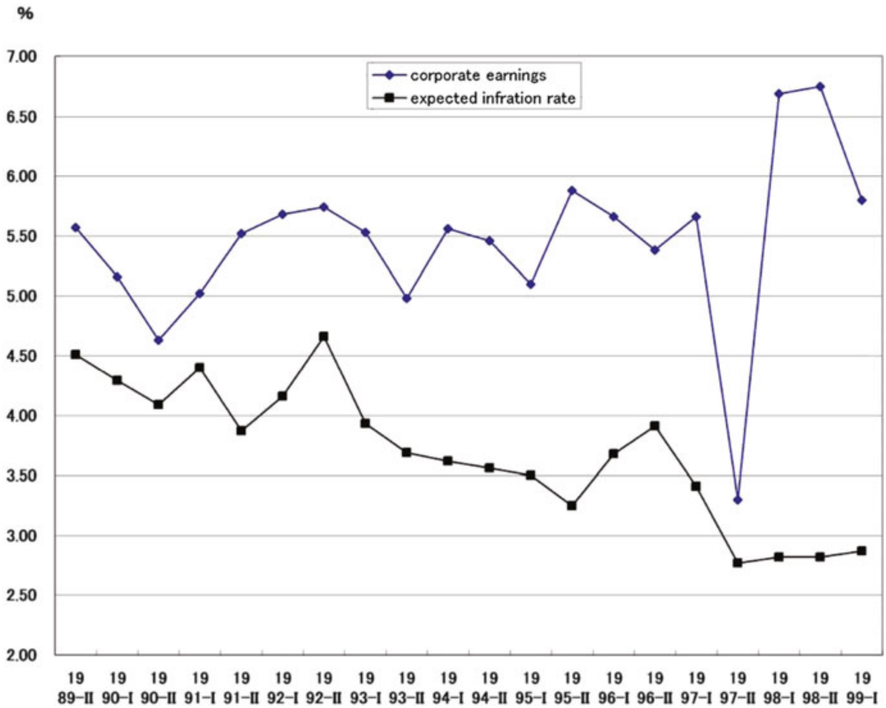


Fig. 12.4 Expectations of corporate earnings and inflation rate

$$\begin{aligned} P_t &= E_t\left[\frac{\pi_t\,(1+g)}{1+r} + \frac{\pi_t\,(1+g)}{(1+r)^2} + \cdots\right] \\ &= E_t\left[\frac{1+g}{r-g}\right]\pi_t \end{aligned} \tag{12.1}$$

Here, π_t represents corporate earnings (profits) known as of period t.

Denoting the expected inflation rate at t as f_t, the expected real growth rate of corporate earnings as $\hat{g}_t$, and the constant real discount rate as $\hat{r}$, (12.1) can be rewritten as

$$P_t = \frac{1+\hat{g}_t+f_t}{\hat{r}-\hat{g}_t}\pi_t \tag{12.2}$$

To calculate the fundamental value based on (12.2), we need data on the expected inflation rate, which is asked in our survey:

What do you think the inflation rate (rate of increase in the cost of living) in the US will be on average over the next 10 years?

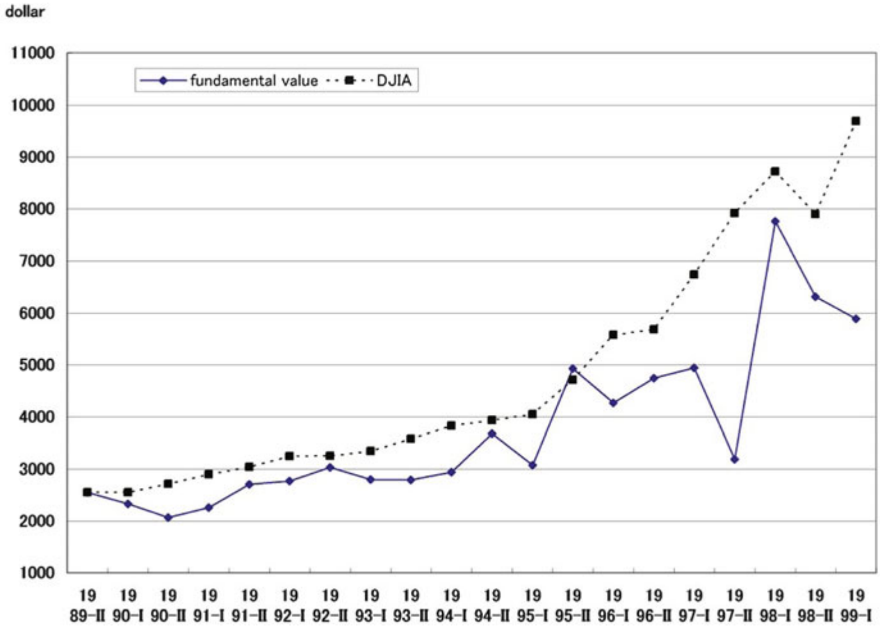


Fig. 12.5 DJIA and estimate fundamental value

The result is also plotted in Fig. 12.4, which shows a decline throughout the period, from 4.5 % in 1989 to 3 % in the late 1990s.

Substituting in the survey results for $\hat{g}_t$ and f_t and the actual value of corporate profits at t-1 into (12.2), and using the assumption that $r = 9$ %, we calculate the fundamental value P. We then adjust the value so the number at 1989-II equals the value of the DJIA at that time, which allows us to compare the estimated value from (12.2) with the historical DJIA. Specifically, we multiply the estimated fundamental value by DJIA and divide by the estimated value at 1989-II, which implies that the DJIA equaled the fundamental value in 1989-II.

The result is shown in Fig. 12.5. The figure reveals that DJIA was overpriced throughout the period. However, until 1995-II, the overpricing was temporary and was tended to disappear quickly. It was in 1996-I that the gap started to widen; the bubble reached \$4000 in 1999-I, whereas the fundamental value itself was \$6000. The successive rapid decline until 2002 precisely eliminated this bubble.

The estimated result depends on the assumption about the real discount rate. Lower assumed values result in smaller bubbles. If we assume a rate of 7.5 % or 8.0 %, the fundamental value exceeds the actual value of the DJIA at 1995-I and 1998-I and II, implying that the DJIA was underpriced in these periods. Still, the conclusion that a bubble existed in most of the periods, and that its size at 1999-I reached \$4000, is maintained under this different assumption.

We should be careful to note that the above estimation depends on various restrictive assumptions, so that the estimation is merely an exercise. In addition to the fact that (12.2) is based on restrictive assumptions, we did not estimate the value of the discount rate, but simply assumed its value. However, there is a possibility that we have underestimated the size of bubble. In the late 1990s, many argued that the US economy went into a new super-productive phase. This argument may have made people believe that future corporate earnings are high. If such a belief was wrong, and their expectation of future earnings was unreasonably high, we should say that 'fundamental value' itself, based on such an irrational belief, contained a bubble.

References

Barsky RB, Bradford De Long J (1993) Why does the stock market fluctuate? Q J Econ 108: 291–311

Bates DS (1991) The crash of 87: was it expected? The evidence from the options markets. J Financ 46(3):1009–1044

De Bondt WFM (1991) What do economists know about the stock market? J Portf Manag 84:84–91

Feldstein M, Horioka C (1980) Domestic saving and international capital flows. Econ J 90:314–329

French KR, Poterba JM (1990) Japanese and U.S. cross-border common stock investments. J Jpn Int Econ 4:476–493

French KR, Poterba JM (1991) Were Japanese stock prices too high? J Financ Econ 29:337–363

Katona G (1975) Psychological economics. Elsevier, North Holland, Amsterdam

Rabinow P, Sullivan WM (1979) Interpretive social science: a reader. University of California Press, London

Shiller RJ (1989) Market volatility. M.I.T. Press, Cambridge

Shiller RJ (1990) Market volatility and investor behavior. Am Econ Rev 80:58–62

Shiller RJ (1995) Conversation, information, and herd behavior. Am Econ Rev, 85:181–185

Shiller RJ, Kon-Ya F, Tsutsui Y (1991) Investor behavior in the October 1987 stock market crash: the case of Japan. J Jpn Int Econ 5:1–13

Shiller RJ, Kon-Ya F, Tsutsui Y (1996) Why did the Nikkei crash? Expanding the scope of expectations data collection. Rev Econ Stat 78(1):156–164

Sternberg RJ (1987) Implicit theories: an alternative to modeling cognition and its development. In: Bisanz J, Brainerd CJ, Kail R (eds) Formal methods in developmental psychology: progress in cognitive development research. Springer, New York, pp 155–192

Ueda K (1992) Monetary policy under disequilibrium in the balance of international payments. Toyo Keizai Shinpo Sha, Tokyo (in Japanese)

Chapter 13
Price Bubbles Sans Dividend Anchors: Evidence from Laboratory Stock Markets

Shinichi Hirota and Shyam Sunder

Abstract We experimentally explore how investor decision horizons influence the formation of stock prices. We find that in long-horizon sessions, where investors collect dividends till maturity, prices converge to the fundamental levels derived from dividends through backward induction. In short-horizon sessions, where investors exit the market by receiving the price (not dividends), prices levels and paths become indeterminate and lose dividend anchors; investors tend to form their expectations of future prices by forward, not backward, induction. These laboratory results suggest that investors' short horizons and the consequent difficulty of backward induction are important contributors to the emergence of price bubbles.

Keywords Stock price bubbles • Short-term investors • Backward induction • Market experiments

JEL Classification Codes G12, C91

1 Introduction

This chapter uses a laboratory experiment to explore how investors' decision horizons affect the formation of stock prices. It has long been argued that speculation by short-term investors induces price volatility. Speculators are concerned primarily with capital gains; the dividends paid during their short investment horizon are

---

The original article first appeared in Journal of Economic Dynamics and Control 31:1875–1909, 2007. A newly written addendum has been added to this book chapter.

S. Hirota (✉)
School of Commerce, Waseda University, 1-6-1 Nishiwaseda, Shinjuku,
Tokyo 169-8050, Japan
e-mail: shirota@waseda.jp

S. Sunder
Yale School of Management, 165 Whitney Avenue, New Haven, CT 06511, USA
e-mail: shyam.sunder@yale.edu

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_13

relatively insignificant. Expectations of capital gains depend on higher order expectations susceptible to cascading or mass psychology of the market. In markets populated by short-term investors, the argument goes, prices tend to lose their dividend anchors, can take any value depending on such expectations, and are therefore susceptible to price indeterminacy and bubbles.[1]

This conventional wisdom is not necessarily accepted in today's finance textbooks. We teach that the prices of securities are determined by their fundamental values—the sum of the discounted value of future dividends—irrespective of investors' time horizons. Even short-term investors are assumed to backward induct from future cash flows to arrive at the fundamental value of securities at the present time.

On the other hand, some theoretical research suggests that such backward induction may fail, and short-term speculative trading may give rise to bubbles. Rational bubble models (Blanchard and Watson 1982; Tirole 1985) consider indeterminacy of price levels of infinite maturity securities without terminal values. Short-term investors have no values from which they can backward induct. In addition, recent theoretical models argue that when investors have heterogeneous information and/or their rationality is not common knowledge, short-term investors may find it difficult to backward induct and security prices may diverge from their fundamentals (e.g., De Long et al. 1990a, b; Froot et al. 1992; Dow and Gorton 1994; Allen et al. 2006). Unlike psychological theories of mass hysteria or limited cognition, these models show that indeterminacy of security prices can arise because even rational investors may not have the knowledge, beliefs, and coordination devices necessary for prices to coincide with the fundamental values.

From these models, we conjecture that the difficulty of backward induction originating in investor short-horizons is a primary source of price bubbles. However, little empirical evidence exists to support this theoretical body of work. Since fundamental values of equities are rarely known, empirical studies of price bubbles using data from the field face the difficult challenge of separating bubbles from the possibility that the fundamental model is misspecified.[2]

Laboratory experiments can address this problem by letting the experimenter assign parameters to subjects to control the fundamental value. Smith et al. (1988) showed that bubbles can arise in simple laboratory asset markets and conjectured that investors may conduct speculative trades aiming to sell the security to others at higher prices. Lei et al. (2001) experiment, however, rejected this conjecture. It showed that bubbles arise even when investors cannot engage in speculative trades; bubbles arise from errors in investors' decisions themselves. In contrast to these works, the objective of our experiment is to explore how investors' decision horizons

[1] In UK, "short-termism" is a charge leveled at the expectations of financial institutions from the companies to which they provide capital. See Moore (1998) and Tonello (2006).

[2] See, Stiglitz (1990), and Fama (1991). LeRoy (2003) also states in a recent survey article that "One would like to see the development of empirical tests that could distinguish between bubbles and misspecification"(p. 25).

influence stock prices. To attain this aim, we control not only the fundamental value but also the investors' decision horizon relative to the maturity of the security. We report on the design and results of such an experiment.

The main treatment in the experiment is differentiated by long- and short-horizon investors. In the long-horizon sessions, the investors' decision horizon extends to the date of maturity of the security, at which time they receive an exogenously specified dividend. In the short-horizon sessions, investors' decision horizon ends well before the date of maturity, and they exit by receiving the prices endogenous to the session. This price is the average of the next period's predicted price; the predicted price is submitted not by investors, but by predictors who are a separate group of subjects who watch, but do not participate in, trading and who get paid based on the ex post accuracy of their predictions. This treatment is chosen to prevent the manipulation of the ending value by investors. It is important to note that this predicted price is not necessarily linked to the exogenously specified terminal dividend—the fundamental value—via backward induction. Following the above mentioned models, we predict that the security prices should deviate from the fundamentals in the short- but not in the long-horizon sessions. In addition to the main treatment, we examine the robustness of any effects of the main treatment with respect to several other variations described later.

We find that security prices tend to form bubbles in short-, but not in long-horizon markets. With short investor horizons, prices lose dividend anchors and their levels and paths become indeterminate. While parts of some paths are consistent with rational bubbles, others exhibit positive feedback loops (Shiller 2000). The results are consistent with the proposition that when they are unable to backward induct from dividend anchors, investors tend to form their expectations of future prices by forward induction using first-order adaptive or trend processes. In these markets, allocative efficiency is unpredictable, and the cross-sectional dispersion of wealth increases with the deviation of prices from fundamentals. In contrast, prices in markets populated by long-horizon investors tend to converge to the fundamentals.

These laboratory results support the proposition that the difficulty of backward induction by short-horizon investors is a critical factor in the generation of bubbles. They also suggest that bubbles are more likely to occur in markets for securities with longer duration or maturity, and more uncertain dividends. These laboratory findings are consistent with the stylized facts of the susceptibility of high-growth and new technology stocks to bubble formation.

The remainder of this chapter is organized as follows: Section 2 reviews the literature on linkages among investment horizon, backward induction, and the emergence of bubbles. Section 3 describes the experimental design and procedures. Section 4 reports our laboratory results, and Sect. 5 discusses the implications.

2 Investment Horizons and the Security Valuation

In standard theory, the security prices are, or tend toward the fundamental values—the sum of discounted present value of expected future dividends—irrespective of investors' decision horizons. This proposition is derived through backward induction from future dividends to present value. We argue that the backward induction may fail and the prices in markets populated by short-term investors deviate from the fundamentals and form bubbles. The discussion below helps guide the design of a critical laboratory experiment to examine the conditions that generate price bubbles or indeterminacy in stock markets.

Let us start with considering a security that matures at time $t + m$. For simplicity, the security pays only a terminal dividend D at time $t + m$. Assuming a zero discount rate and risk-neutral investors, the fundamental value of the security at time t is:

$$F_t = E_t(D) \tag{13.1}$$

where E_t (.) is investors' homogeneous expectation at time t.

2.1 Long-Term Investor's Valuation

We define a long-term investor as one whose investment horizon is longer than or equal to m. This investor holds the security until its maturity and receives the terminal dividend D at $t + m$. The value of the security to the investor at time t, V_t (and its price P_t in a market populated by such homogenous investors) is:

$$P_t = V_t = E_t(D) \tag{13.2}$$

This price is equal to the fundamental value F_t.

2.2 Short-Term Investor's Valuation

Next consider short-term investors with investment horizon $k < m$, who must sell the security before its date of maturity. The investor buys the security at time t, holds it for k periods, and sells it at $t + k$. The value of the security to this investor and its price P_t in a market populated by such homogenous investors is:

$$P_t = V_t = E_t\left(P_{t+k}\right) \tag{13.3}$$

where P_{t+k} is the stock price at $t + k$. Equation 13.3 indicates that price P_t depends on the investor's expectation of the future sales price, E_t (P_{t+k}). It opens the possibility that when investors' horizon is shorter than the maturity, the security price may not be equal to the fundamental value F_t.

In the standard backward induction treatment in finance, even in markets populated by short-term investors, P_t should be equal to F_t via backward induction. Let n be the number of successive generations of investors, each living for k periods, who populate the market between time t and $t+m$, and $(n\text{-}1)k < m < nk$. Then at time $t+(n\text{-}1)k$, the price $P_{t+(n\text{-}1)k}$ should be equal to the investor's expectation at $t+(n\text{-}1)k$ of the terminal dividend, $E_{t+(n\text{-}1)k}(D)$. If the investor at time $t+(n\text{-}2)k$ knows this, $P_{t+(n\text{-}2)k}$ should be equal to his expectation of $P_{t+(n\text{-}1)k}$, which is $E_{t+(n\text{-}2)k}(E_{t+(n\text{-}1)k}(D))$. If the investor at time $t+(n\text{-}3)k$ knows this, $P_{t+(n\text{-}3)k}$ should be equal to $E_{t+(n\text{-}3)k}(E_{t+(n\text{-}2)k}(E_{t+(n\text{-}1)k}(D)))$. Repeating this process back to t, we get

$$P_t = E_t(E_{t+k}(E_{t+2k}\left(\cdots\left(E_{t+(n-1)k}(D)\right)\cdots\right) \tag{13.4}$$

Assuming that all investors across generations are homogenous in information sets, we can use the law of iterated expectations and obtain $P_t = E_t\,(D) = F_t$.

2.3 Difficulty of Backward Induction

The possibility of the failure of the backward induction argument in markets with short-term investors as a source of price bubbles has been suggested in the rational bubbles literature (Blanchard and Watson 1982; Tirole 1982, 1985): when the maturity of the security extends indefinitely ($m \to \infty$), the investor cannot obtain the terminal value and backward induction becomes impossible. Prices become indeterminate and may deviate from fundamentals. It is also known that this indeterminacy may arise even when the maturity is finite, provided that there are an unlimited number of trading opportunities (Allen and Gorton 1993).

Second, when investors have heterogeneous beliefs, the law of iterated expectations is no longer applicable. In order to backward induct the future sales price in that case, investors must form higher-order expectations: If each generation of investors have a k-period investment horizon, the investors entering at t must decide on the basis of what they believe the investors at $t+k$ expect what investors at $t+2k$ expect ... and so on till $t+(n-1)k$. Froot et al. (1992) and Allen et al. (2006) consider the case where short-term investors have only limited information about the future investor's expectations. They show that the backward induction argument fails and that stock prices are affected by noisy or irrelevant public information. [3]

Third, the backward induction argument assumes common knowledge of investors' rationality: investors are not only rational but also know that other investors are rational as well. Recent theoretical research illustrates that when the common knowledge assumption of rationality does not hold, the backward induction argument fails and stock prices deviate from fundamentals. De Long et al.

[3]Allen et al. (1993) argue that even in markets with long-term investors the backward induction can fail and bubbles form when investors do not know others' expectations.

(1990a, b) show that when there are noise traders in the market, rational investors with short-term horizon expect future mis-pricing, and might not engage in arbitrage even if they know that the current price deviates from the fundamentals. Dow and Gorton (1994) argue that when there is uncertainty about the existence of informed rational arbitrageurs in the future, the current traders fail to backward induct the sales price from the future dividends.

We examine the empirical relevance of these theoretical ideas. Do short investment horizons, and consequent difficulty of backward induction when there is no reasonable basis to form common knowledge expectations of higher orders, give rise to price bubbles and indeterminacy? We test this hypothesis with the following laboratory experiment.

3 The Experimental Design

We created double auction markets for trading units of a security on a computer network in a laboratory. The security paid a single liquidating dividend to its holders at the end of its life, which was divided into many trading periods of 3 min each. Participating subjects were randomly assigned to one of two roles—investors and predictors. Each investor was endowed with 10 securities and 10,000 points of "cash" at the beginning of period 1, and could trade freely through the multiple periods without going short on securities or "cash." At the end of the session, the securities held by investors were liquidated by paying them either a dividend or a predicted price (as described later under the Main treatment). The investors could make money through trading and terminal liquidation of their securities.

The predictors studied all the instructions given to the investors. They did not get endowments of "cash" or securities, could not trade, and only knew the range of the traders' terminal dividends. At the end of each period they were asked to predict the average price of the security transactions for the following period. Their earnings depended on the accuracy of their predictions. In addition to these earnings, all subjects earned \$3 (in Sessions 1–6) or \$5 (in Sessions 7–11) if they arrived in the laboratory punctually.

3.1 Main Treatment: Long or Short Investment Horizons

In five sessions (numbered chronologically 3, 4, 5, 6, and 7 in Table 13.1), the investors were informed that the security would pay a terminal dividend (pre-written on their respective cards) at the end of period 15 and that the session would end at that time. This environment is designed to correspond to that of the long-term investors in Sect. 2.1: since the investors' investment horizon can extend to the security's maturity, we call these long-horizon sessions. In these sessions, if the investors buy (sell) securities depending on whether the price is lower (higher)

Table 13.1 Experimental sessions

M/D/Y	Session	Final payoff on stock	Dividends on cards given to investors	Range of dividend announced	Initial shares per subject	Initial cash	Number of investors	Number of predictors	Predictor fixed number[a] *(N)*	$/point conversion rate	Actual number of periods	Announced number of periods	Question and answer	Verification and correction
09/21/01	1	Predicted value	40 for 2 75 for 2	[10, 300]	10	10,000	4	2	150	0.025	12	≤30	No	No
09/29/01	2	Predicted value	70 for 3 130 for 3	[70, 130]	10	10,000	6	2	100	0.015	15	≤30	No	No
09/30/01	3	Dividend	80 for 3 150 for 3	[40, 300]	10	10,000	6	1	100	0.015	15	15	No	No
11/16/01	4	Dividend	120 for 2 205 for 3	[120, 205]	10	10,000	5	2	150	0.01	15	15	Yes	No
12/01/01	5	Dividend	105 for 2 175 for 3	[55, 350]	10	10,000	5	2	130	0.012	15	15	Yes	No
02/15/02	6	Dividend	150 for 5	[40, 300]	10	10,000	5	1	100	0.015	15	15	Yes	Yes
02/22/02	7	Dividend	105 for 3 175 for 3	[55, 350]	10	10,000	6	2	130	0.012	15	15	Yes	Yes

(continued)

Table 13.1 (continued)

M/D/Y	Session	Final payoff on stock	Dividends on cards given to investors	Range of dividend announced	Initial shares per subject	Initial cash	Number of investors	Number of predictors	Predictor fixed number[a] *(N)*	$/point conversion rate	Actual number of periods	Announced number of periods	Question and answer	Verification and correction
03/01/02	8	Predicted value	40 for 2 75 for 2	[10, 300]	10	10,000	4	2	150	[Points/average] × $25	17	≤30	Yes	Yes
04/26/02	9	Predicted value	70 for 2 130 for 3	[70, 130]	10	10,000	5	2	150	[Points/Average] × $25	15	≤30	Yes	Yes
07/11/02	10	Predicted value	150 for 8	150	10	10,000	8	2	100	[Points/average] × $25	15	≤30	Yes	Yes
07/12/02	11	Predicted value	75 for 6	75	10	10,000	6	2	200	[Points/average] × $25	15	≤30	Yes	Yes

[a]Predictor compensation = max (0, (*N* – absolute price prediction error))

than the terminal dividend, the price would converge to the fundamental value of the security, $P_t = F_t$. If some investors speculate seeking capital gains within the 15-period session, and push the prices away from the fundamental value, such deviations would also give the long-term investors an opportunity to make profits by arbitraging between such deviations and terminal dividends.

In six sessions (numbered 1, 2, 8, 9, 10 and 11 in Table 13.1), investors were informed that their securities would pay a terminal dividend at the end of period 30 if the session were to end in period 30. They were also informed that the session would end at a period written down inside a sealed envelope, and this period is very likely to be less than 30. Although they were not informed about the real number of periods in the session until it was actually terminated, they could have estimated that the length of time for which they had been recruited into the laboratory would end well before Period 30.[4] If the session ended earlier than period 30 (as it always did), investors would receive the average transaction price predicted (by the predictors) for the period immediately following the termination for each security they held. We call these short-horizon sessions because this treatment was designed to capture the environment of markets with the short-term investors described in Sect. 2.3: investors' horizon ends before the security matures, and they may find it difficult to use the future dividends to backward induct the sales price they should expect to get from exiting the market.

As we mentioned in Sect. 2.3, recent bubble literature illustrates several factors for breaking the link between investors' expectation of sale prices and the future dividends (fundamentals): Froot et al. (1992) and Allen et al. (2006) point out investors' limited information about the expectations of the subsequent generations of investors; De Long et al. (1990a, b) and Dow and Gorton (1994) suppose the existence of irrational future investors; the rational bubble models (Blanchard and Watson 1982; Tirole 1982, 1985) suggest that the expected sale prices may include the bubble term not linked to the future dividends at all.

Our short-horizon sessions try to realize this breach of the link in the laboratory without using either irrationality or a bubble in future dividends. Since the investors would receive the terminal dividend if their investment horizons were long enough to include period 30, this dividend can be considered the fundamental value. However, investors know that the session is very likely to end before Period 30 and when the session terminates they are paid off the prevailing market price (predicted price) for each security; there is no sensible way to expect this ending price by backward induction. In this manner, we intended to break, or substantially weaken the link between the expected future prices and the terminal dividends.[5] In such an

[4]When subjects were recruited, they were told to participate in an experiment in market decision-making for 2–3 h in total. At the beginning of the session in a laboratory, subjects knew that (i) they had already spent about an hour and half for instruction and trial sessions, and (ii) one period was three minutes long followed by the paper work for a minute or two. Thus they could predict that the session would end well before period 30.

[5]Other variations are also possible. An announcement that the terminal dividend would be paid at period 30 but the session will end earlier, say at period 15, for sure, would break the link

environment, the security prices would lose their dividend anchors, opening up the possibility that they would deviate from the fundamental values.

In this market, the experimenter provides the liquidity at the end of the session: investors' security holdings would be bought back by the experimenter at the prevailing market price, as proxied by the average of price predictions. In this environment, investors can exit the laboratory market without any impact on the price. We introduced this treatment because liquidity is well-documented feature of stock markets such as New York Stock Exchange. We use the average prediction for the period following the last period, instead of the actual market price in the last period, as the ending value of the securities. This helps prevent the manipulation of the ending value by investors with large security holdings.[6]

The short-horizon sessions are especially relevant to the markets for high growth stocks whose dividends may be paid in remote future beyond the investment horizons of the current investors. Investors' valuation of such securities depends mainly on the price expected to prevail in the market at the investment horizon. Even if the investors have their own respective estimates of the fundamental value, it would be difficult, if not impossible, for them to form an expectation of market price at their own investment horizon through backward induction. Such induction would require them to conjecture not only the dividend expectations of various generations of future investors but also the processes by which each generation carries out such backward induction. For high growth stocks we should expect only a weak link, at best, between the investors' valuation and the fundamental value, as in our short-horizon sessions.

3.2 Experimental Procedures

The experimental procedures common to all market sessions are as follows. We summarize information about the 11 sessions in Table 13.1. Each session consisted of some 12 to 17 periods, and each period consisted of 3 min of trading, followed by 1 or 2 min for paperwork. At the start of the session, each investor received 10 shares and "cash" of 10,000 points. The investors could buy securities if they had cash to

completely. Alternatively, the session could be terminated with a common knowledge probability distribution to retain a weak but well-specified link. For example, we could have announced that there will be a ten percent chance that the session will be terminated at the end of period 10, 11, 12, 13, or 14, with the predicted price payoff; if the session goes to period 15, it will be terminated with the pre-specified dividend payoff.

[6]One may wonder whether the average predicted price (liquidation value) may be a candidate for fundamental value in the short-horizon sessions. This predicted price, however, cannot be considered as the fundamental value because the prediction is, itself, endogenous to the market process that includes the behavior of investors and predictors. No concept of value that deserves the label of fundamental can properly be a function of such behavior because then the label itself becomes superfluous.

pay for them, and sell any shares they had. Short sales were prohibited. Securities and cash were carried over from one trading period to the next. The endowment of securities or cash was not replenished.

Before a session started, each investor drew a Dividend Card, which showed his/her terminal dividend per share. In the long-horizon sessions, this amount would be the actual terminal dividend received by the investor at the end of the last period (period 15). In the short-horizon sessions, the investor would receive this amount at the end of period 30 only if the session were to last for 30 periods. This personal dividend per share was each investor's private information (except that it was common knowledge in Sessions 10 and 11). They were told that the dividend might not be the same across the investors, and that the personal dividends of investors lay within the publicly announced range (see Table 13.1, column 5).

The session earnings of each investor were equal to the cash balance at the end of the final period's trading, plus the end-of-session payoff, minus the initial cash provided at the beginning of the session. In the long-horizon sessions, the end-of-session payoff was [his or her dividend per share on the Dividend Card × the number of shares he or she held at the end of the session]. In the short-horizon sessions, the end-of-session payoff was [average predicted price × the number of shares he or she held at the end of the session] if the session ended before period 30 (this always was the case); it was [his or her dividend per share on the Dividend Card × the number of shares he or she held at the end of the session] if the session lasted for 30 periods (this was never the case). With the exception of Sessions 8–11, which used relative performance evaluation, the investor's final earnings in all other sessions were converted from points into US dollars at a pre-announced rate, and paid in cash at the end of the experiment.

Trading was by continuous double auction, implemented with the Caplab™ software. Each investor was free to make bids (proposals to buy shares) and asks (proposals to sell shares) by entering the price and quantity through his/her mouse and keyboard during trading periods. The computer showed the number of shares he or she had, cash balance, market bid and ask price, and the price of the most recent transaction (see Instruction Set 2 for Trading Screen Operation in the Appendix).

All the sessions had predictors as well as investors. After the common instructions and training part of the session, each subject's role (investor or predictor) was determined by lots. The predictors had to estimate the prices at which the investors might trade securities. At the end of each period, they were asked to predict the average stock price of the following period by writing it down on their Price Prediction Sheet. The experimenter gathered this information before starting trading for the period. At the end of each period, the experimenter wrote the predicted price (averaged across all the predictors) on the board for all to see. The predictors' earnings for the period decreased with the magnitude of their prediction errors; they earned [Constant N – the absolute difference between the prediction and the actual average transaction price] points. If this value was negative, they earned zero points for the period. Constant N was the same for all the predictors in one session, but differed across sessions (see Table 13.1 for value of N). Their total earnings for all periods were converted from points into US dollars at a pre-announced rate (except in Sessions 8–11 that used relative performance evaluation).

The sequence of activities in a session was as follows: (1) Instruction sets (general instructions, investor instructions, predictor instructions, and trading screen instructions) were distributed and read out aloud. The subjects could ask questions at any time. (2) All the subjects participated in the trial session (2–3 rounds) until they got used to the trading screen operation using Caplab^{TM}. (3) Each subject drew a slip of paper from a bag that determined his or her role. (4) Each investor randomly picked a Dividend Card on which his or her dividend was written. (5) Trading period 1 of the session began and was followed by other periods.

3.3 Robustness Variations

As shown in the five sections of Table 13.2, the main treatment of long and short horizons was supplemented by five variations to examine the robustness of the main treatment to other plausible experimental conditions.

1. *Heterogeneity of terminal dividends:*

In Sessions 6, 10 and 11, the dividends were identical across all the traders. In contrast, in Sessions 1–5 and 7–9, the terminal dividends written on the cards given to the traders were not identical across traders (e.g., 40 for two traders and 75 for two traders in Session 1: see Table 13.1). This heterogeneous dividend setting creates opportunities to gain from trade and is often adopted in experimental asset market studies (see Sunder (1995) for a review).

2. *Potential inequality between the first and higher-order expectations;*

In Sessions 2, 4, 9, 10 and 11, there existed no gap existed between the actual range of dividends written on individual dividend cards and the maximum dividend range publicly announced to all traders and predictors. For example, in Session 2, three investors were given cards informing them that their own dividend was 70 points, while another three had 130 points as their dividend. It was publicly announced to all subjects that none of the investor dividends lay outside the 70–130 point range. In contrast, in Sessions 1, 3, 5, 6, 7 and 8, a gap existed between the actual range of private dividends and the publicly announced maximum dividend range. For example, as shown in Table 13.1, dividend cards distributed in Sessions 1 and 8 had a terminal dividend of 40 points for two investors and 75 points for the other two investors. It was publicly announced to all traders and predictors that none of the dividend numbers on the cards given to the investors lay outside the 10–300 point range. The information about this range had some chance of creating a non-zero subjective probability in the minds of investors that the other investors may have dividends as high as 300 points. If the investor's own expectation (first-order expectations) of dividends differs from his expectation of others' expectations (second or higher-order expectations), it is possible that even long-horizon investors participate in speculative trading (buy an asset hoping to sell it later to investors with

Table 13.2 Experimental design

		Main treatment: investment horizon	
Robustness variations		Long-horizon session	Short-horizon session
Subsidiary treatment 1: Heterogeneity of pre-determined dividends	Identical pre-written dividends	Session 6	Sessions 10 and 11
	Non-identical pre-written dividends	Sessions 3–5, and 7	Sessions 1, 2, 8, and 9
Subsidiary treatment 2: Potential inequality of the first and higher order beliefs about dividends	Equality between first and higher order beliefs	Session 4	Sessions 2, 9–11
	Potential for a gap between first and higher order beliefs	Sessions 3, 5–7	Sessions 1, 8
Subsidiary treatment 3: Common Knowledge of pre-determined dividends	Dividends common knowledge		Sessions 10 and 11
	Dividends not common knowledge	Sessions 3–7	Sessions 1, 2, 8, and 9
Subsidiary treatment 4: Verification of proper understanding of the instructions	Questionnaire, answer, no verification and correction	Sessions 4 and 5	
	Questionnaire, answer, verification and correction	Sessions 6 and 7	Sessions 8–11
	No questionnaire, answer, verification and correction	Session 3	Sessions 1 and 2
Subsidiary treatment 5: Subjects paid by absolute or relative performance	Payoff based on absolute performance	Sessions 3–7	Sessions 1 and 2
	Payoff based on relative performance		Sessions 8–11

higher private dividends); and such behavior may generate price bubbles.[7] We also check if this occurs in our laboratory.

3. *Common knowledge of predetermined dividends;*

In Sessions 10 and 11, the predetermined dividends written on the trader cards were made common knowledge through a public announcement. In all the other sessions, the predetermined dividends on the cards given to the traders were private knowledge.

[7] Biais and Bossaerts' (1998) model shows this possibility.

4. *Verification of understanding of instructions:*

In Sessions 4–11, the instructions were followed by a written questionnaire and an answer sheet to the questionnaire to help the subjects and the experimenter verify the former's understanding of the instructions and procedures. In addition, Sessions 6–11 included a review of each subject's answers by the experimenter, and an explanation of the relevant part of the procedures in case of any errors in the subject's answers.

5. *Subjects paid by absolute or relative performance:*

In Sessions 1–7, points earned by the subjects were converted into US dollars at a rate announced during the course of the instructions. In Sessions 8–11, the total dollar amount to be paid to the traders (and to the predictors) was announced at the outset. This amount was allocated to individuals in proportion to the number of points earned in the session.[8]

These five robustness variations, as well as the main treatment, are summarized in Table 13.2. This chapter reports on all 11 experimental sessions shown in Tables 13.1 and 13.2. The sessions were held at Yale University with undergraduate student subjects in the fall of 2001 through the summer of 2002. A fresh set of subjects were recruited for each of the 11 sessions, and none had participated in any previous research experiments with stock markets. The sessions lasted 2.5 h on average.

4 Experimental Results

Figures 13.1, 13.2, 13.3, 13.4, 13.5, 13.6, 13.7, 13.8, 13.9, 13.10 and 13.11 show the price and allocation data from the 11 laboratory sessions. Each figure shows the time series of transaction prices (diamond markers) with the average price for each period written at the top of the chart[9] The thick solid line indicates the market equilibrium price based on the fundamental value of the shares. The market equilibrium price is the higher of the two dividend values in the heterogeneous dividends sessions (e.g., 150 in Session 3) and the unique dividend value in the homogenous dividends sessions (e.g., 75 in Session 11). The thin solid line shows the upper limit of the publicly announced range of dividends (300 in session 3), which is also, presumably, the upper limit of the investors' and the predictors' second (or higher) order beliefs about dividends.

[8]We adopted relative performance based payment in Sessions 8–11 due to limitations of our budget. In the absolute performance based payment sessions, payment to subjects when a bubble arises could be considerable. For example, in session 2, we paid 138 dollars per subject on average for a 3 h session. We changed the payment policy from absolute to relative for the subsequent short-horizon sessions.

[9]The average prices were calculated by excluding transaction prices that result from order-of-magnitude typographical errors (8 transactions in total throughout 11 sessions).

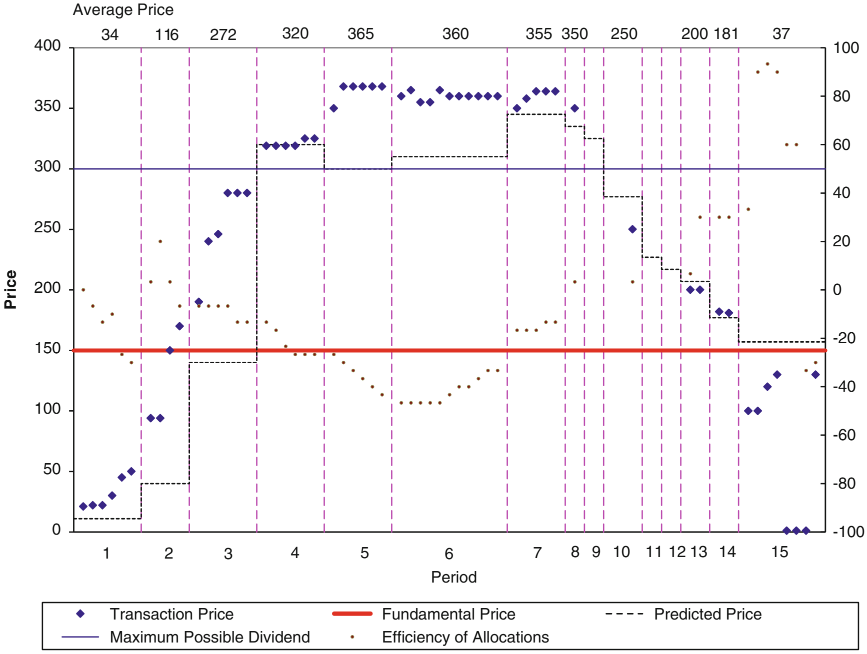


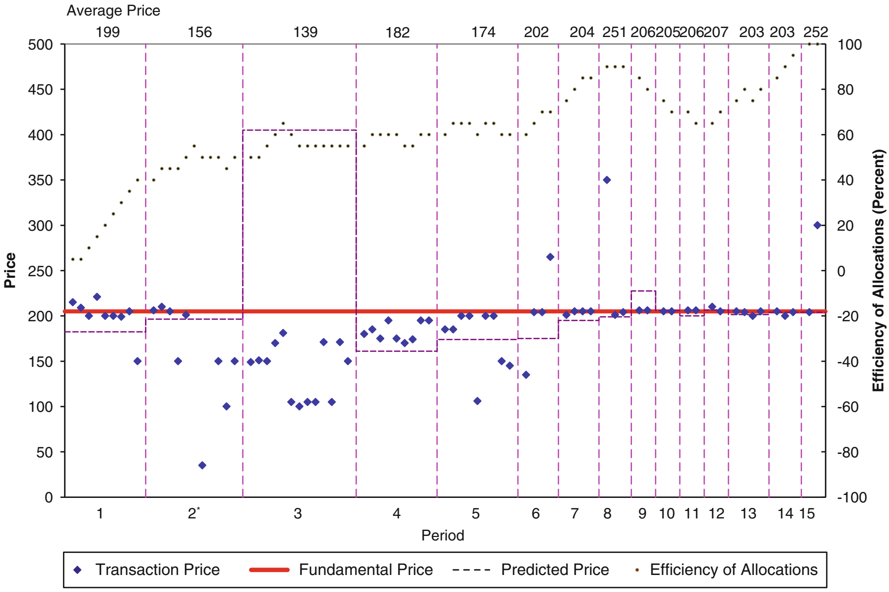
Fig. 13.1 Stock prices and efficiency of allocations for session 3 (long-horizon session)

The dotted line shows the average predicted price for the period. In the discussion below we use the predictions, submitted by the subjects who were assigned to play the role of predictors exclusively, as proxies for the expectations held by the investors in the experiment. It seems reasonable since the information sets of the predictors and the investors are essentially the same (except for any private dividends).[10]

The small dots plotted against the y-axis on the right hand scales track the allocation of securities relative to the initial endowment (0 %) after each transaction. If all the securities were to be transferred to the investors who had the higher dividend (fundamental value) on their cards, the allocative efficiency would be 100 %; if all the securities were to be transferred to the investors who had the lower dividend on their cards, the allocative efficiency would be a negative 100 %.

For example, in Session 1 (short-horizon treatment, Fig. 13.6) the transaction prices (diamonds) remained in the 80–85 range throughout the session, and stabilized at around 83, about 10 % above the fundamental value of 75 (thick solid line).

[10]We had different subjects play the two roles to avoid confounding the incentives of the investors and the predictors.



* Two transaction in period 2 occurred at 2,155 because the bidder said he inadvertly added 5s to the intended bids

Fig. 13.2 Stock prices and efficiency of allocations for session 4 (long-horizon session)

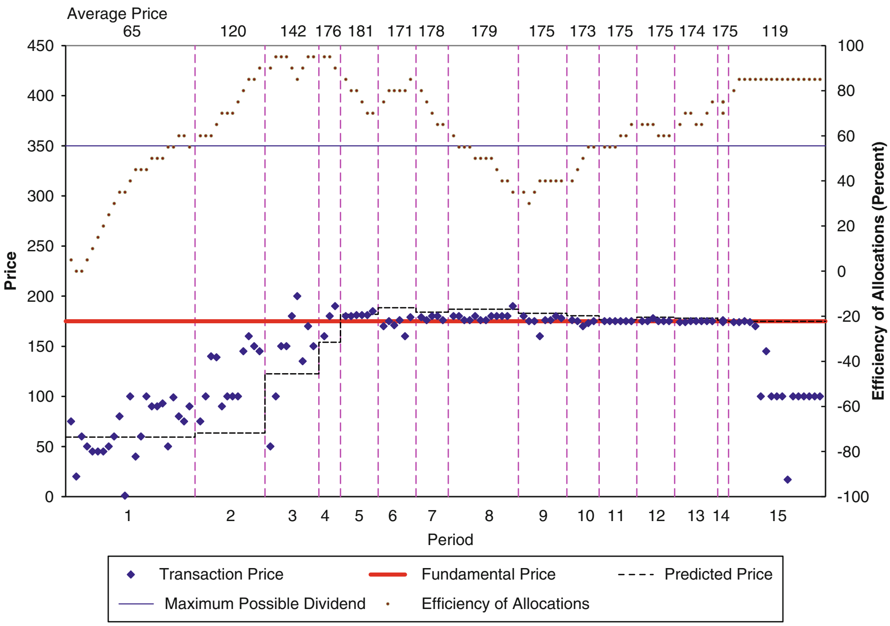
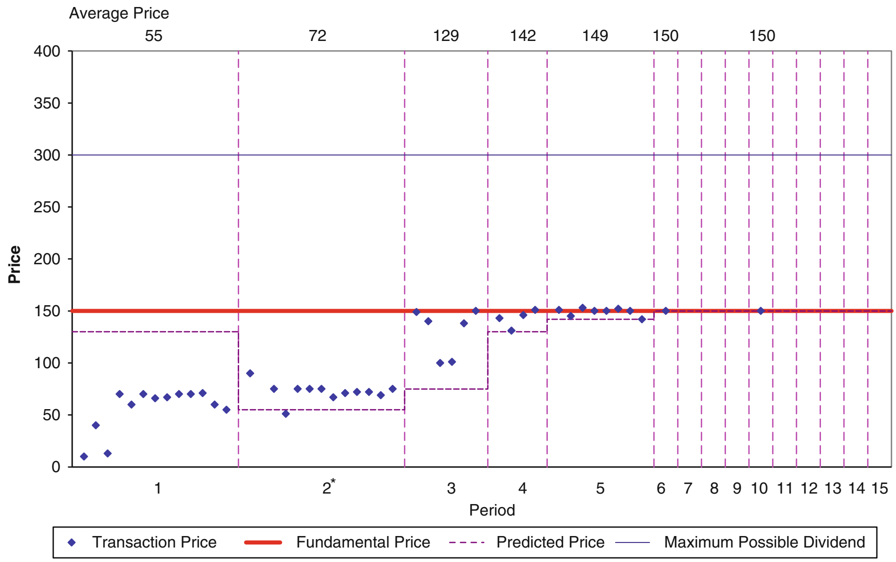
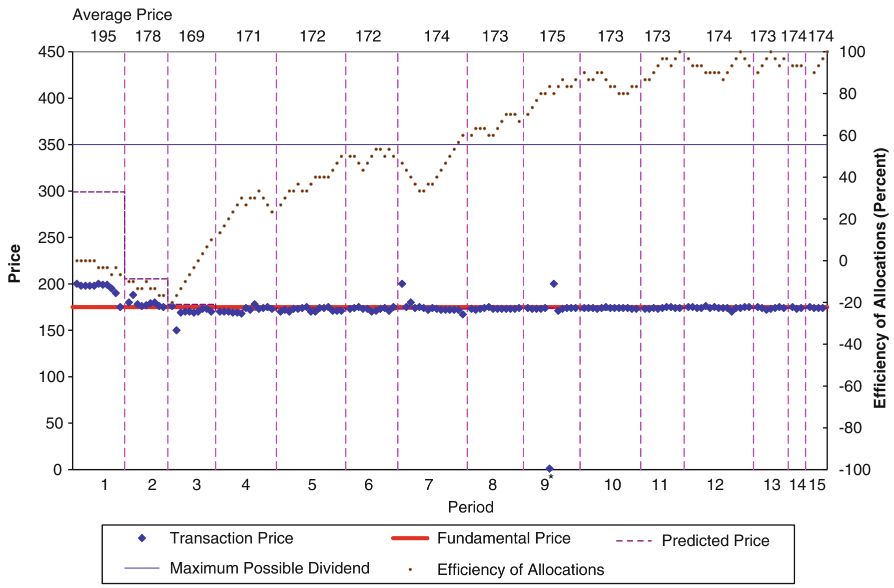


Fig. 13.3 Stock prices and efficiency of allocations for session 5 (long-horizon session)



* One transaction in period 2 occurred at 5,050 because the bidder said he inadvertently added 50 to the intended bids of 50. **Allocative efficiency of this market is undefined because all investors had identical dividends.

Fig. 13.4 Stock prices for session 6 (long-horizon session)



* One transaction in period 9 occurred at 1 because of mis-ask.

Fig. 13.5 Stock prices and efficiency of allocations for session 7 (long-horizon session)

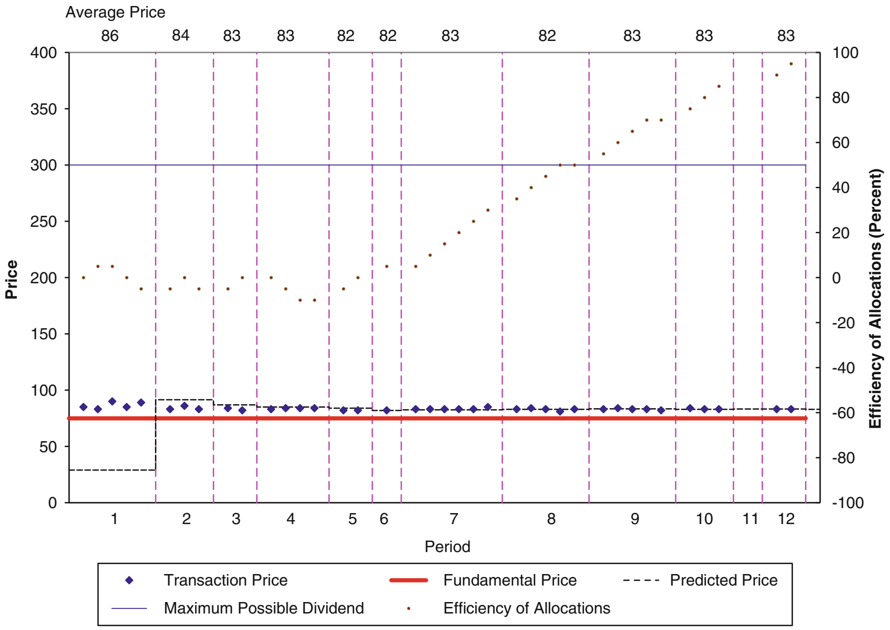


Fig. 13.6 Stock prices and efficiency of allocations for session 1 (short-horizon session)

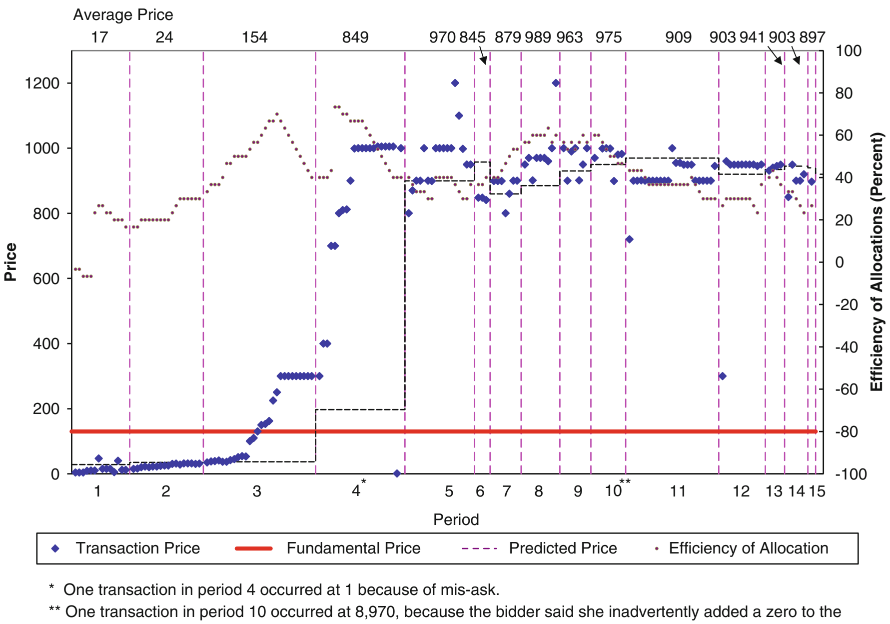


Fig. 13.7 Stock prices and efficiency of allocations for session 2 (short-horizon session)

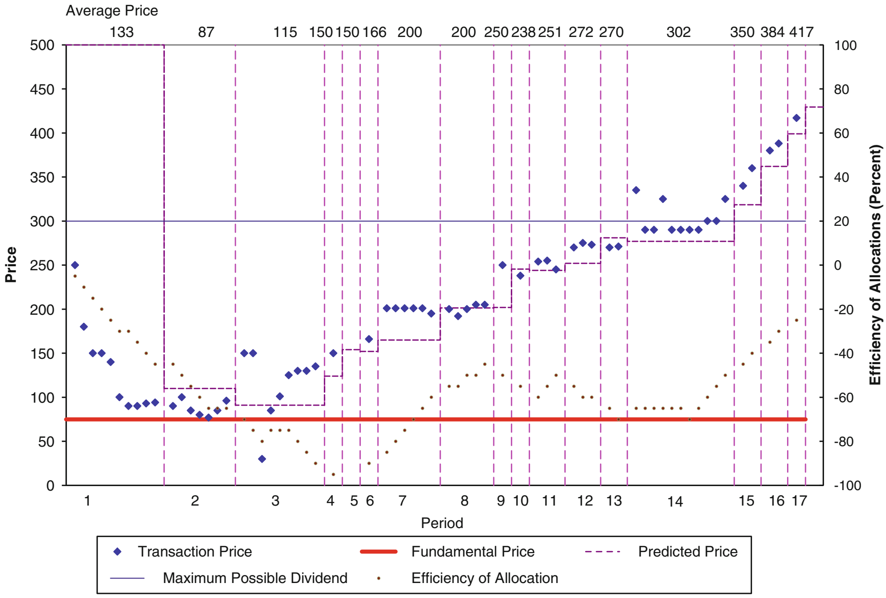


Fig. 13.8 Stock prices and efficiency of allocations for session 8 (short-horizon session)

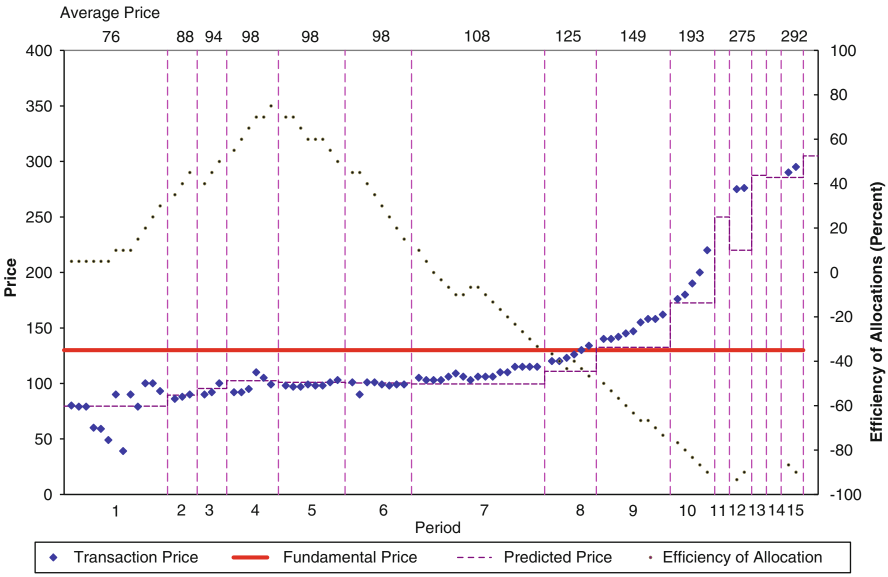
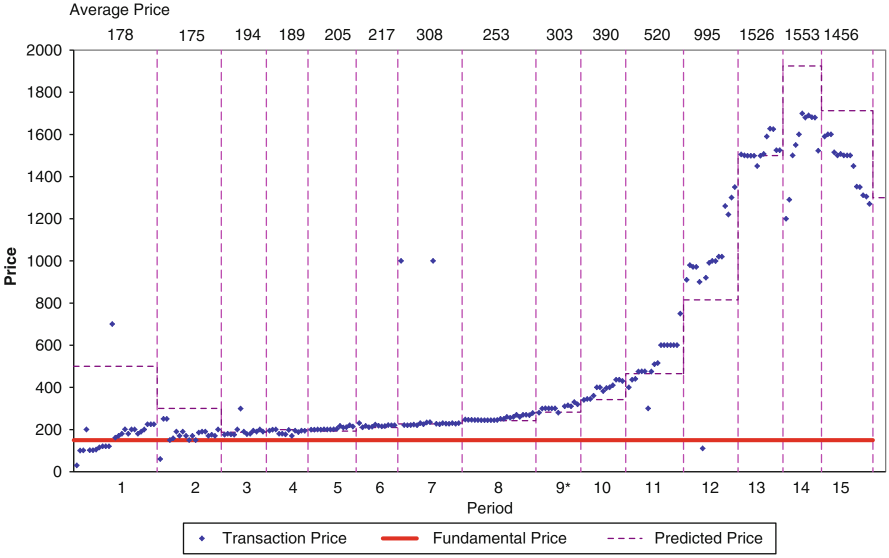
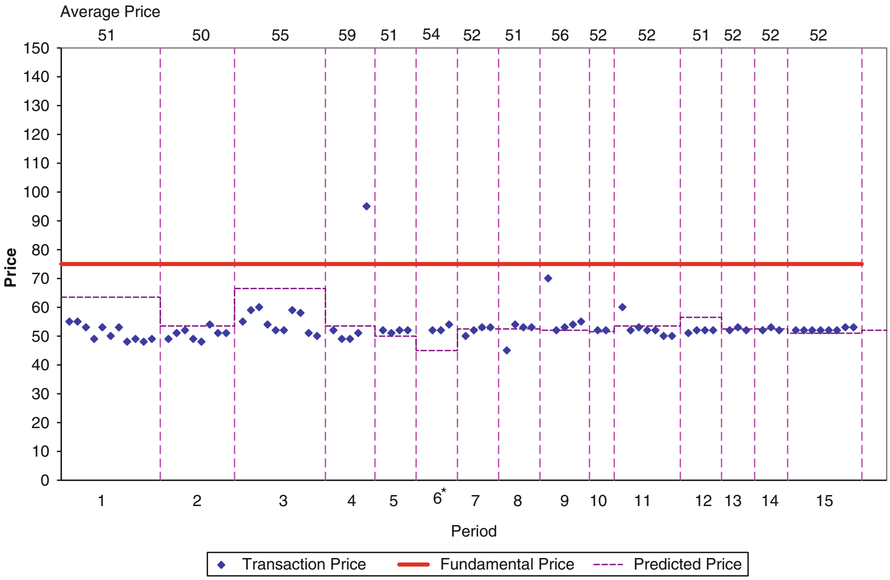


Fig. 13.9 Stock prices and efficiency of allocations for session 9 (short-horizon session)



* One transaction in period 9 occurred at 2,260 because the bidder said he inadvertently added 2 to the intended bids of 260. **Allocative efficiency of this market is undefined because all investors had identical dividends.

Fig. 13.10 Stock prices for session 10 (short-horizon session)



* One transaction in period 6 occurred at 200 because the bidder said he inadvertently added 0 to the intended bids of 20.

Fig. 13.11 Stock prices for session 11 (short-horizon session)

These prices were well below the upper limit of the potential second-order beliefs of investors (thin solid line at 300). The average predicted price remained close to 83 throughout (dotted line). The allocation of securities between the high and the low dividend investors hovered near the initial allocation for the first six periods, and then the higher dividend investors steadily bought all but 1 of the 40 securities by the end of the 12th period.

We organize the data around five results.

Result 1 *In the long-horizon sessions, the security prices converge to the equilibrium level derived from the fundamental values of individual investors.*

In the long-horizon sessions (Sessions 3, 4, 5, 6, and 7) all the traders knew that there would be 15 periods in the session and that a terminal dividend would be paid at the end of the session. In this zero-discount rate and ample liquidity environment, the fundamental (equilibrium) price for all fifteen periods of these sessions is equal to the highest dividend across traders. Figures 13.2, 13.3, 13.4 and 13.5 show that the prices in these markets are determined by this highest dividend, and they converge to the fundamental level. A bubble was observed in Session 3 (Fig. 13.1); it did not, however, reappear in the subsequent sessions (Sessions 4–7) after the experimental instructions were modified to include a questionnaire to test the subject comprehension of the experimental instructions, and additional instruction as necessary. In these four sessions, the prices exhibit a strong tendency to converge to the fundamental value.[11]

This result suggests that long-horizon investors play a crucial role in efficiently pricing securities. Even if prices temporarily deviate from the fundamental values due to speculative or noise trading, long-horizon investors' arbitrage tends to bring the prices to their fundamental value. Even in long-horizon sessions, a considerable number of speculative trades took place. In 218 (39 %) of 559 transactions[12] in Sessions 3–7, investors bought shares at prices higher than their own terminal dividend, presumably to seek capital gains, not dividends. These speculative trades, however, do not seem to have destabilized the prices. We observe that price deviations from fundamentals are arbitraged away. Even in Session 3 where we

[11] We conjecture that in Session 3, a bubble formed because at least some of the subjects did not fully understand the instructions about the structure of the security. In period 4 the price rose to 320, although the subjects knew that the maximum possible dividend in this market was only 300. In period 15, one subject sold shares at a price of 1; he later told us that he forgot that the securities earned a liquidating dividend at the end of the last period. In Session 5 (Fig. 13.3), the price dropped below the fundamental value (175) in period 15. The experimenter failed to disable the trading function on the predictors' computers. Although the instructions prohibited predictors from trading, one of the predictors traded anyway, and 14 of the 17 trades in period 15 were sales by this predictor. We confirmed that this predictor's trading had no significant effect on the convergence pattern to the fundamental value during the first 14 periods of the session.

[12] We exclude 3 transactions that resulted from order-of-magnitude typographical errors in Sessions 4 and 6, and 21 predictor's transactions in Session 5 (see, the previous footnote) as errors, from a total of 583 transactions.

experienced a bubble, the price deviations tend to get smaller in the later periods (this is in sharp contrast with the results from short-horizon sessions presented later).

We should add that one of the robustness variations, inequality between the first and higher-order expectations, did not affect our experimental results. Biais and Bossaerts' (1998) model predicts that Keynes' beauty contest bubble arises even when investors have long-term horizons, provided that investors' higher-order beliefs differ from their first-order beliefs. Scheinkman and Xiong (2003) show that when agents have an opportunity to profit from other agents' overvaluation in a market with short-sale constraints, asset prices may exceed the fundamental value. In Sessions 5, 6, and 7, however, where the first and higher-order beliefs might have diverged (also note that our laboratory markets have short-sale constraints), we did not observe any significant gap between the transaction and fundamental prices.[13]

Finally, our result is consistent with Lei et al. (2001). Their experimental study attributes the bubbles observed in the laboratory to the subjects' lack of understanding of the structure of the asset, the nature of the task, or the opportunities available to them. In our laboratory, before we started testing the subjects' understanding of the rules, a bubble was observed in Session 3; in the later four sessions where we used questionnaires (Sessions 4 and 5) and verified and corrected any misunderstanding of the rules (Sessions 6 and 7), no bubbles were observed. The lack of understanding of the instructions is a laboratory artifact that must be carefully guarded against in all experiments.

Overall, the data from the long-horizon sessions (Sessions 4, 5, 6 and 7, but not necessarily from Session 3) support Result 1.

Result 2 *In the short-horizon sessions, the security prices deviate from the fundamental values.*

Figures 13.6, 13.7, 13.8, 13.9, 13.10 and 13.11 show that, in contrast to the long-horizon treatment of Sessions 3–7, short-horizon Sessions 1, 2, and 8–11 exhibit a strong tendency to generate bubbles. While the dividend numbers seem to have played a role in the determination of prices in the early periods of some of these sessions (e.g., Sessions 1, 8 and 10), as these sessions progressed (see Figs. 13.6, 13.7, 13.8, 13.9, 13.10 and 13.11) the prices often ceased to have any meaningful relationship with the dividends (they were about 10–900 % higher or lower than the fundamental values). Investors did not appear to form their expectations about the liquidation value of the securities from terminal dividend, and backward induct the current value from such expectations. As predicted, the role of terminal dividends in the determination of security prices is minimal at best. Further, the possibility of a gap between the first and the second-order beliefs in Sessions 1 and 8 does not seem to affect the existence and magnitude of any gaps between prices and

[13]It is possible that we did not observe bubbles in Sections 5, 6, and 7 because we did not succeed in our attempt to induce our laboratory subjects to develop divergent first and higher order beliefs about dividends. We asked the subjects "not to assume that other subjects have the same dividends as their own" and announced a range containing all investors' dividends, which was wider than the actual distribution of dividends. We cannot rule out the possibility that this procedure does not necessarily induce divergent first and higher-order beliefs.

terminal dividends. Sessions 2 and 10, in which there was no attempt to induce a divergence between the first and the second (higher) order expectations, resulted in large bubbles.

In all six short-horizon sessions both the levels and the paths of prices are indeterminate. In Session 1 (Fig. 13.6), both the transaction as well as the average predicted prices settled down to about 83, about 10 % above the fundamental value, and stayed there throughout. In Session 11 (Fig. 13.11), prices and predictions settled down in the neighborhood of 50, about 33 % *below* the fundamental value, and stayed there till the end. Both these bubbles, one positive and the other negative, were relatively stable, supported by mutual reinforcement between the transaction prices and the price predictions. In Session 1, there was no reason for any investor to pay a price of 83 for these securities except on the basis of the expectation that he/she will sell them at a similar price in later periods, or get the average predicted price of about 83 at the end of the session. A similar argument applies to Session 11. These stable price paths can be interpreted as evidence of rational bubbles, considering the zero discount rate used in the laboratory.[14]

In Session 2 (Fig. 13.7), prices started out low and increased slowly until an explosive spurt from the middle of period 3 to the middle of period 4. It settled noisily around 900, almost seven times higher than the fundamental price, until the end of the session. The fundamental value (130) hardly played any role in determining the transaction prices. Our conjecture is that the investor #5 wanted to push the prices higher, and did so until he depleted most of his cash.[15] After this steep rise in prices the predictors started predicting that price, and there was little reason for the prices to return to their fundamental levels. The investors' expectations of high prices in the future seemed to sustain the realization of high prices. This relatively stable price path after Period 5 can also be regarded as a rational bubble.

On the other hand, Sessions 8, 9, and 10 exhibit rising price paths through most of these sessions. Given zero discount rate in the laboratory, these rising paths are not consistent with rational bubbles prediction. It seems that the actual price increases are followed by increases in price predictions, which in turn help raise the transaction prices. Shiller (2000) argues that this "amplification" mechanism is created by investors' psychological factors and emotions. Our data suggest that the amplification mechanism arises when the investors have short investment horizons and difficulty in backward inducting the value of securities.

In Session 8 (Fig. 13.8), the price grew steadily and by the end of the session it reached above 400, which is about 5 times higher than the fundamentals. In

[14]Bubbles with stable prices are also suggested by Ackley (1983). He states, "... a speculative price—(i.e., a non-equilibrium price) is not always or necessarily a moving price. The price may rest for a considerable period of time at—or fluctuate narrowly around—a level far above (or below) any equilibrium determined by market fundamentals." (p. 6)

[15]During period 4, when prices rose shapely, investor #5 bought 15 shares (out of 30 transactions in this period) and his cash balance went down from 11,966 to 1,353.

Session 9 (Fig. 13.9), the price settled down around 100 for the first half of the session, which was consistent with a rational bubble. However, it grew sharply during the latter periods. In Session 10 (Fig. 13.10), we observed a large bubble. In this session, the terminal dividend (150) was common knowledge among the investors. However, the dividend seemed to play an anchoring role only during the early periods. In periods 3–5, the price rose to a level around 200 and continued to rise, even accelerating in periods 9–10, reaching above 1,500 in periods 13–14, which was about ten times as high as the fundamental value before falling back about 100 points in period 15. It is interesting to note the dynamics of transaction and predicted prices. In period 10 the average price prediction was 342 (dotted line) and the actual transactions averages at 390. In period 11 the predictions rose to 465, and the transactions rose to 520. In period 12 the predictions and the transactions rose to 815 and 995 respectively. Finally the price prediction rose to 1,500 in period 13 when the actual average price was 1,526. Increases in the transaction and the predicted prices appear to reinforce each other.

One may argue that large bubbles in Sessions 2, 8, 9, and 10 arise not from the difficulty of the backward induction under short investment horizons, but from strategic coalition among subjects in our experimental setting of short-horizon sessions. We can suppose that investors attempt to drive prices upward together so as to obtain as much profits as possible from the experimenter at the end of the session. We are unable to exclude this possibility for Session 2 where subjects were paid in proportion to their absolute performance (number of points earned). However, we reject it for Sessions 8, 9, 10 where each subject was paid by his/her relative performance, and the total amount paid to all investors as well as the total amount paid to all predictors, was fixed. Further, the subjects were made fully aware of this relative performance-based compensation scheme in these sessions. In these sessions, neither investors nor predictors benefit from higher prices per se, either individually or collectively. There was no incentive for them to collude to raise prices. Since higher prices are likely to have larger prediction errors, the predictors would have been discouraged from engaging in such behavior. The subjects were recruited independently, and had no knowledge of either the game or their role in it, and could not have arranged a scheme of transfer payments between investors and predictors as an incentive to raise price predictions. Moreover, in Session 9, prices were below the fundamental value in the first eight of the fifteen periods; in Session 11, prices were significantly below the fundamental value through out (only 2 out of 82 transactions were conducted at prices above the fundamental value).

In summary, bubbles were observed in all six sessions of the short-horizon treatment where investors should have had difficulty in the backward induction. Instead of being anchored by investors' beliefs about dividends the prices are determined largely by unanchored anticipations of the future prices. Consequently, the price level and pattern become indeterminate. Our laboratory data provide strong support for Result 2.

Next, we investigate how investors formed their expectations of prices, and what, if any, differences exist between long and short-horizon sessions in this respect.

Result 3 *In the long-horizon sessions, price expectations are consistent with backward induction; in the short-horizon sessions, price expectations are consistent with forward induction.*

We examine the price expectation formation process using one backward induction (fundamental) model and two forward induction (adaptive and trend) models.

The fundamental model assumes that investors form expectations about the current price by backward inducting from future dividends through a simple statistical adjustment process (α being the adjustment coefficient):

$$E_t\left(P_{t+1}\right) - P_t = \alpha\left(D_{max} - P_t\right) \tag{13.5}$$

where D_{max}, the largest (across investors) terminal dividend value in the market, is the fundamental value of the security. Except in Sessions 10 and 11, D_{max} is not announced to the investors and predictors; the dividend range is publicly announced but each investor's dividend is private information. It is possible that subjects rationally predict D_{max} by observing transaction prices if the market is efficient in the sense that prices reflect investors' private information. This strong form efficiency was actually observed in simple repeated experimental asset markets (Forsythe et al. 1982; Plott and Sunder 1982). As our laboratory markets are simple enough (the security pays only a terminal dividend) and repeated for some 12–17 periods under unchanging conditions, the subjects may have a rational prediction on D_{max} and utilize it in forming their expectations of prices (13.5).

This fundamental (backward induction) model is compared with two forward induction models. One is the first-order adaptive model:

$$E_t\left(P_{t+1}\right) = E_{t-1}\left(P_t\right) + \beta'\left(P_t - E_{t-1}\left(P_t\right)\right) \tag{13.6}$$

The expectation at t is formed by adding a fraction (β') of the most recent expectation error to the most recent expectation. To compare it with the other models, we rewrite it as

$$E_t\left(P_{t+1}\right) - P_t = \beta\left(E_{t-1}\left(P_t\right) - P_t\right) \tag{13.7}$$

The other forward induction process is a simple trend model:

$$E_t\left(P_{t+1}\right) - P_t = \gamma\left(P_t - P_{t-1}\right) \tag{13.8}$$

This model is often called an extrapolative expectation model. The expectation at t is formed by adding a fraction (γ) of the most recent price change ($P_t - P_{t-1}$) to the most recent observed price (P_t). When $\gamma > 0$, recent price increases cause investors to expect further price increases in the future.

Finally, we consider a more general specification for expectation formation that include the fundamental, adaptive, and trend factors simultaneously:

$$E_t(P_{t+1}) - P_t = \alpha(D_{max} - P_t) + \beta(E_{t-1}(P_t) - P_t) + \gamma(P_t - P_{t-1}) \quad (13.9)$$

This combined model allows for the possibility that investors use some combinations of the three simple processes to form their expectations. Estimates of multiple regressions (13.9) can also help us detect any biases in the estimates of simple regressions, (13.5), (13.7), and (13.8), due to the omitted variables problem.

We estimated regression equations (13.5), (13.7), (13.8), and (13.9) including constant terms on two samples, one had pooled the data from all five long-horizon sessions and the other had pooled the data from all six short-horizon sessions. As mentioned before, price predictions submitted by the predictors were used as proxies for investors' price expectations. Our null hypothesis is that expectations are static, $E_t(P_{t+1}) = P_t$: α, β, and γ are equal to zero in (13.5), (13.7), and (13.8), respectively, and all three parameters are zero in (13.9).[16]

Two panels of Table 13.3 show the regression results for the long and short-horizon samples respectively. In the long-horizon sessions,[17] the data are more consistent with the backward induction (fundamental) model than with either of the two forward induction (adaptive and trend) models. While the coefficient of ($D_{max} - P_t$) is significantly positive in the fundamental model (13.5), neither coefficient of ($E_{t-1}(P_t) - P_t$) nor ($P_t - P_{t-1}$) is significant in the adaptive model (13.7) and trend model (13.8). In the estimation result of the combined model (13.9), only the fundamental factor ($D_{max} - P_t$) is significant. These results suggest that in long-horizon sessions, the fundamental value of the security not only determines transaction prices but also critically affect the price expectations. The existence of long-horizon traders plays a role in informing market participants of the fundamental values and, at the same time, seems to enable them to expect that future price will converge to the fundamentals. Under such circumstances, investors' speculation caused by future price expectation should have a stabilizing effect on market prices. We conjecture that this mechanism actually worked in our long-horizon sessions; prices converged to the fundamental value even though about 40 % of the trades can be attributed to speculative motives because they represented purchases at a price above personal dividends.

On the contrary, Panel B of Table 13.3 shows that in short-horizon sessions, the forward induction models are supported over the backward induction model. First, while the adaptive factor ($E_{t-1}(P_t) - P_t$) is insignificant in the simple regression (13.7), it is significant at 1 % level in the more general specification (13.9). Second, the trend factor ($P_t - P_{t-1}$) is significantly positive at a 5 % level in the simple

[16] To cope with heteroskedasticity, we used White's (1980) heteroskedastic-consistent standard error to test the significance of the coefficients. Furthermore, we also run weighted least squares regressions and found that the inferences given below remained unchanged.

[17] In period 3 of Session 4, one predictor submitted an abnormally high price prediction (600), after observing an order-of-magnitude typographical error in a trading price in period 2. We regard this expectation data as a noise and exclude from the sample. Inclusion of this datum does not change our results.

Table 13.3 Price expectation model estimates

Model	Const.	$D_{max} - P_t$	$E_{t\text{-}1}\ (P_t) - P_t$	$P_t - P_{t\text{-}1}$	$\overline{R}^2$	N
Panel A: long-horizon sessions						
Fundamental (13.5)	5.705 (5.386)	0.261** (0.098)			0.085	102
Adaptive (13.7)	4.863 (6.075)		0.428 (0.396)		0.045	102
Trend (13.8)	3.318 (5.643)			−0.189 (0.333)	0.002	102
Combined (13.9)	6.646 (6.321)	0.200** (0.074)	0.443 (0.235)	0.178 (0.274)	0.090	102
Panel B: short-horizon sessions						
Fundamental (13.5)	3.543 (4.593)	−0.077 (0.042)			0.092	166
Adaptive (13.7)	19.851** (6.869)		−0.113 (0.149)		0.010	166
Trend (13.8)	5.809 (4.411)			0.467* (0.196)	0.349	166
Combined (13.9)	4.808 (3.244)	0.027 (0.027)	0.543** (0.179)	0.829** (0.164)	0.506	166

** and * indicate that the coefficient is significantly different from zero at the 1 % and 5 % level at two-tailed test, respectively. Figures in *parentheses* are White's (1980) heteroskedastic-consistent standard errors

regression (13.8) and at a 1 % level in the combined regression (13.9). On the other hand, the fundamental factor, ($D_{max} - P_t$), is not significant in either of the two specifications (13.5 or 13.9).[18]

These results indicate that in short-horizon sessions price expectations are formed by forward induction from current or past transaction prices. Extant field and experimental studies generally support an adaptive expectation model, but not a trend (extrapolative) model (Frankel and Froot 1987; Williams 1987; Allen and Taylor 1990; Taylor and Allen 1992). In contrast, our data indicate that the trend factor has a considerable effect on future price expectations in short-horizon sessions—a price rise of 1 raising the expectation of next period price by 0.47. In estimates from the combined model (13.9), this effect is larger at 0.83 (see,

[18]We thank the referee for pointing out that the possibility of session-specific effects should be considered. We re-estimated the original equations after including session dummies for long- as well as short-horizon sessions. The new estimates and their significance level are almost unchanged, with two exceptions: (i) The coefficients of (D_{max} – P_t) for long-horizon sessions, which are significant at 1 percent level in the original estimates, have p-values of 0.063 in (13.5) and 0.064 in (13.9) when session dummy is included. Coefficients of ($E_{t\text{-}1}(P_t)$–P_t) and (P_t-$P_{t\text{-}1}$) are insignificant in the dummy regressions as well as the original regressions. (ii) Coefficient of ($E_{t\text{-}1}(P_t)$–P_t) in (13.9) for short-horizon sessions becomes significant at 5 % level in the dummy regression (p-value = 0.017) compared to 1 % level in the original regression. These results show that there is no qualitative change in our main findings after session-specific effects are controlled for.

Panel B in Table 13.3). Given the random noise in transaction prices, formation of expectations on the basis of the most recent trend holds considerable potential to destabilize prices. A change in price for any reason changes investors' expectation of future prices in the same direction, and it becomes rational to bid the prices up or down in the same direction. These findings are consistent with Hommes et al. (2005) who also show that individual forecasters may coordinate on trend following rules, causing asset prices to fluctuate and deviate from their fundamental. The model estimates suggest that such an amplification mechanism occurred in Sessions 8–10.

Result 4 *The allocative efficiency is high in the long-horizon sessions; it is unpredictable in the short-horizon sessions.*

The allocative efficiency (the percent of securities transferred toward the fundamental value equilibrium allocation) for Sessions 1–5 and 7–9 is plotted in Figs. 13.1, 13.2, 13.3 and 13.5, 13.6, 13.7, 13.8, 13.9 with small dots on the right hand scale.[19] During the three long-horizon sessions (Sessions 4, 5 and 7) almost all the securities ended up in the hands of the high fundamental value traders, and the allocative efficiency approached 100 % at the end of each session. In Session 3, the efficiency hovered around zero during the bubble phase (Periods 1–9). The bubble collapsed in the last five periods as the securities were transferred to the high-dividend investors.[20]

The bubble economies of the short-horizon Sessions 8 and 9 had negative efficiencies (the securities were transferred to low-dividend investors), and Session 2 had an efficiency of only 20 %. When the dividends are replaced by endogenously determined predicted prices, such as the liquidation payoff, there is no reason to expect either high or low efficiencies in the market. Every trader would get the same liquidation payoff from holding the security, and the market exerts little pressure to make the allocation of resources more efficient as defined by the ultimate (period 30) dividends.[21]

[19]In sessions 6 (Fig. 13.4), 10 (Fig. 13.10) and 11 (Fig. 13.11), all investors had identical dividends and the allocative efficiency was undefined.

[20]The efficiency dropped just before the end of Session 3. After the session ended, a high-dividend trader who had accumulated a large number of securities told us that in period 15 he forgot that there was a terminal dividend, panicked, and sold off many securities at a price of 1.

[21]However, in Session 1, the efficiency rose to almost 100 percent. Since the investors could be reasonably sure that they would get the average predicted price that hovered around 83 through almost the entire session, there was no pressure for the securities to be transferred to the higher-dividend investors. In fact, 39 of the 40 securities were held by one of the two high-dividend investors at the end of Period 12. One low-dividend investor held one security, while one high-dividend and one low-dividend investor held no securities. It is therefore plausible that the transfer of securities to one high-dividend investor could have been the outcome of idiosyncratic trading strategies of the investors, and the securities could have just as easily ended up in the hands of any of the other three investors.

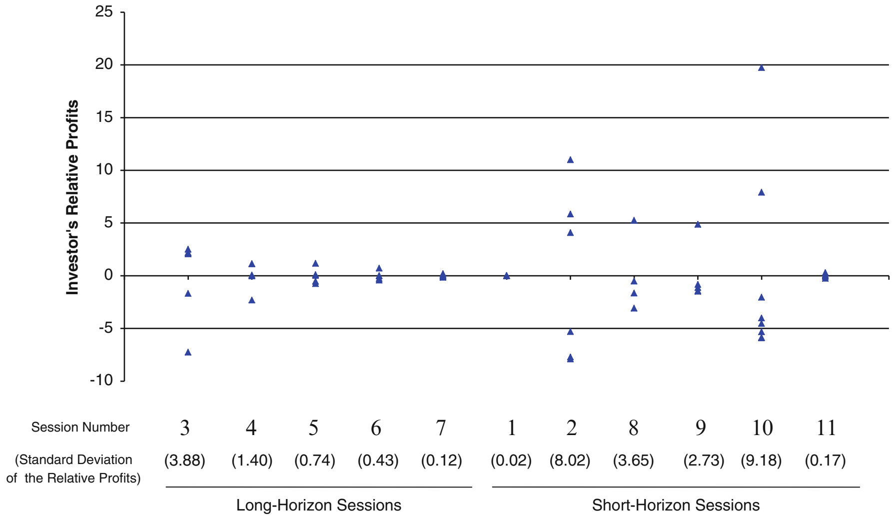


Fig. 13.12 Dispersion of investor profits

The uncertain allocative efficiency of bubble-prone markets is an argument for public policies aimed at discouraging the formation of price bubbles. Bubbles also have significant distributive consequences in the form of increasing the dispersion of wealth among agents.

Result 5 *The cross-sectional dispersion of investor wealth increases with the size of bubbles.*

The cross-sectional distribution of relative profits of individual traders (= individual trader's profit in points / the fundamental value of the initial endowment of 10 shares – the cross-sectional average of this ratio) for each of the 11 sessions are shown in Fig. 13.12. Each triangle marker represents one trader's relative profit. The standard deviations of the relative profits are shown in parentheses under the session numbers.

Of the five long-horizon sessions, only Session 3, which had a large price bubble, shows a significant cross-sectional dispersion of profits (the standard deviation of the relative profits is 3.88). The dispersion in all other sessions is close to zero. Of the six short-horizon sessions, 1 and 11 had small but stable bubbles, and the dispersion of relative profits in these sessions is close to zero. In the other four sessions we observed large price bubbles, all of them show large dispersion of individual profits. The data suggest that the dispersion of wealth increases with the magnitude of the price bubble.

5 Discussion and Concluding Remarks

In the 11 laboratory sessions of this experiment, we observe that when investor decision horizons are short (relative to maturity of the security) and they face a difficult backward induction task, the security prices tend to deviate from fundamentals to form bubbles. In these circumstances, prices lose their dividend anchors and become indeterminate. While some price paths are consistent with rational bubbles, others exhibit positive feedback loops. Market participants tend to form their expectations of future prices through forward induction, using first-order adaptive or trend processes. These results are robust to several variations in experimental conditions; they are also consistent with the agent based models with boundedly rational heterogeneous interacting agents, which typically include fundamentalists with long investment horizons and chartists with short horizons (see LeBaron 2006 and Hommes 2006 for recent surveys).

Our laboratory findings suggest several insights into the stock market environments where bubbles are likely to occur. First, investors' time horizons are critical to asset pricing. The frequency and impact of the failure of backward induction in a market is greater when it is populated or dominated by short-term traders (e.g., day traders); security prices in such markets are more likely to deviate from the fundamental value. In contrast, the existence of long-term investors is crucial for stabilizing market prices near the fundamentals, not only through their arbitrage activity but also through their expectations which are anchored to their estimates of future dividends.

Second, securities with longer maturities—generally, longer duration securities[22]—are more susceptible to bubbles. As the duration of a security increases, investors receive smaller proportion of its value in the form of dividends within their investment horizons; and a greater part of their value depends on the expectations of capital gains, which in turn depends on higher-order expectations. Such higher order expectations might be unstable because of the possibility of the failure of backward induction. Therefore the prices of securities with longer durations are more likely to deviate from the fundamentals. This effect of firm's dividend policy on the volatility and the level of its stock price can be added to the tax, agency, and signaling theory challenges to the Miller and Modigliani's (1961) dividend irrelevance proposition.

Third, bubbles are more likely to occur when the future dividends are more uncertain. In such cases, it is difficult for the investors to conjecture what others think (and what others think about what others think) about the future prospects of dividends. It becomes more likely for the investors to fail to backward induct, and for the prices to be unhinged from their dividend anchors.

These implications—the influence of investment horizon, maturity and duration, and uncertainty on the likelihood and severity of bubbles—can be tested by using data from the field, as well as by conducting further laboratory experiments.

[22]Duration is the first derivative of the present value of a security with respect to the discount rate, or the weighted average time of cash flows associated with the security.

They appear to be consistent with casual observations of stock market behavior. Historically, price bubbles have often been attributed to securities with longer maturity or duration and greater uncertainty about the fundamentals, such as high-growth and new technology stocks (see Blanchard and Watson 1982; Ackley 1983). The so-called "dotcom" bubble is a recent example. It seemed that Internet stocks lost their dividend anchors and suffered from price indeterminacies, as did our laboratory stocks in short-term horizon sessions.

Our experimental result also shows that the allocative efficiency of markets is high in long-term horizon sessions and becomes indeterminate in short-term horizon sessions. Since the efficient allocation of capital, not gambling, is supposed to be the social function of security markets, it is understandable that policy makers view bubbles to be undesirable. Furthermore, the evidence on the tendency of bubbles to increase the cross-sectional dispersion of wealth serves as another reason for policy intervention in markets.

In recent decades, many researchers have presented evidence against the market efficiency. Shiller (1981) concludes that, given the fundamentals, security prices are too volatile. French and Roll (1986) observed that the volatility is greater when the markets are open, and inferred that the market trading itself seems to create volatility. This inference is consistent with forward induction. Recent behavioral finance research argues that emotion and psychological factors have considerable effect on stock prices (Shiller 2000), and that momentum trading strategies sometimes work (Jagadeesh and Titman 2001; Goetzmann and Massa 2002). Our experimental results imply that such phenomena are more likely to be observed when the market is mainly populated by short-term investors who, facing difficulty of the backward induction, resort to forward induction.

Appendix 1: Instruction Sheets for the Subjects for Session 2

General Instructions

This is an experiment in market decision making. The instructions are simple, and if you follow them carefully and make good decisions, you will earn more money, which will be paid to you at the end of the session.

In this session, we conduct a market in which you can trade an object we shall call "shares." You will be assigned the role of either an investor or a predictor. At the beginning of session, your role (investor or predictor) is determined by lots: "I" means investor and "P" means predictor. Your role does not change during the session.

If you are assigned to the role of an investor, you may buy/sell shares in the market. At the end of the session, your total points gained from the market will be converted into U.S. dollars at $ 0.015 per point and paid to you in cash. The more points you earn, the more dollars you will take home with you.

If you are assigned to the role of a predictor, you will predict the average market price at which the investors trade shares each period. What you earn each period depends on the accuracy of your price prediction. The more accurately you predict the share prices, the more dollars you will take home with you.

During this session, neither investors nor predictors are allowed to talk to any other participant. Also, you are to follow the various instructions given by experimenter. Violation of instructions risks forfeiting your earnings.

Investors Instructions

If you are assigned to role of an investor, you may trade shares. At the start of the session, you are given 10 shares. You may sell these shares or keep them until the end of the session. You are also provided with an initial "cash" of 10,000 points. You may use the points to purchase shares or keep them until the end of the session. Trading in shares should follow the rules to be explained later.

The session consists of many market periods of 3 min each. That is, period 1 takes place during the first 3 min of trading, period 2 in the second 3 min, period 3 in the third 3 min, and so on. The number of periods will not exceed 30.

To buy shares you should have cash to pay for them. Buying a share reduces your cash balance by the purchase price. You may sell any shares you have. Selling increases your cash balance by the sale price.

If you keep a share, then you receive a dividend on it at the end of period 30. It is written on the Dividend Card given to you before period 1. Dividend is your private information, and it may be different for different investors. You are not informed of other's dividends. The range of dividends will be announced at the beginning of period 1.

In the likely case that the session ends before period 30, each share you hold at the end of the last period (i.e., at the end of the session) will be converted into points at the average of the predictions of the next period's price made by the predictors.

Your profits in this experiment are the sum of profits from two sources: trading profits and end-of-session payoff. Your trading profits are equal to the change in your cash balances due to trading, which is calculated as (your cash balance at the end of last period – your initial cash provided at the beginning of the experiment). Your end-of-session payoff will depend on the number of shares you hold at the end. If the session lasts for 30 periods, this payoff will be (the number of shares you hold at that time × your dividend per share on the Dividend Card). If the session ends earlier, this payoff will be (the number of shares you hold at that time × average predicted price). Your total profits will be converted from points into U.S. dollars at the rate of $0.015 per point and paid to you in cash (Fig. 13.13).

Investor's Record Sheet

Name______________________ Number _____________ Date ___________

Period	Shares	Cash	Profits
0	10	10,000	
1			
2			
…			
29			
30			
Trading Profits	Last Cash Balance – Initial Cash of 10,000 =		
Final Payoff	Last share holding × Dividend per share or Last share holding × Average Prediction =		
Total Profits	In Points		
Total Profits	In dollars (points × $0.015) =		

Fig. 13.13 Investor's record sheet

Predictors Instructions

Predictors are asked to predict the prices at which the investors trade shares in each 3 min period. At the beginning of each period, you predict the average stock price of the period by writing it down in the column of My Predicted Price on your Price Prediction Sheet. We shall gather this information before starting trading for the period. At the end of each period, we shall write on the board the average predicted price for all to see.

During the period, you can see bids, asks and transaction prices on the big screen in the room.

Your profit for the period depends on the accuracy of your price prediction. Each period, you will earn 100 minus the absolute difference between your prediction and the actual trading price (averaged across all transactions in the market) shown on the big screen. For example, suppose, you predict a price of 960 and the actual average price is 980, you have a prediction error of 20 points, and your prediction earnings will be 100 minus 20 which is 80. On the other hand, if the actual average transaction price turns out to be 900, you have prediction of error of 60 points, and your prediction earnings will be 100–60 = 40.

You should record the actual average price, your prediction error (Absolute Difference), and your earning on your Price Prediction Sheet at the end of each period. At the end of the experiment, you should add your prediction earnings for all periods. Your total earnings are converted from points into U.S. dollars at the rate of $0.015 per point and paid to you in cash (Fig. 13.14).

Price Prediction Sheet

Name ____________ Date _____________ Predictor Number ___________

Period No.	My Predicted Price (PP)	Actual Average Price (AP)	Absolute Difference |PP-AP|	My Earning from Prediction 100 -|PP-AP|	My Cumulative Earnings in Points
1					
2					
…					
29					
30					

Your Dollar Earnings (points × $0.015) = ____________

Fig. 13.14 Price prediction sheet

Fig. 13.15 Trading screen

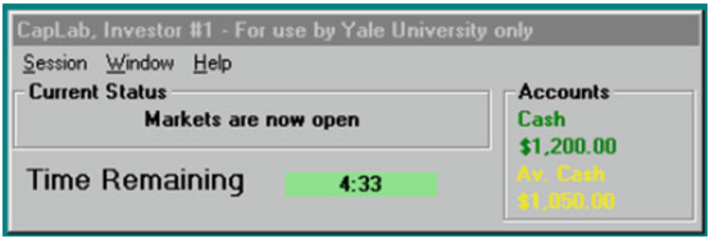


Fig. 13.16 Clock

Instruction Set 2 (Trading Screen Operation)

1. Figure 13.15 shows an active trading screen belonging to Investor 1.
2. Your identification number is shown in the title bar.
3. The number of shares you have are shown in the light blue section under the "Hold." label. This number increases when you buy and decreases when you sell shares.
4. The clock in your CapLab screen (Fig. 13.16) starts ticking down from a preset time allowed for trading, and shows the amount of time remaining till the end of trading.
5. Your cash balance is shown in the CapLab screen in green font in Fig. 13.16.
6. Submit your bids (proposals to buy shares) by entering price and quantity under the "My Bid" column in brown section and pressing the enter key. Do the same

for submitting asks (proposals to sell shares) under the "My Ask" column in deep yellow section. Note that default quantity is 1 share.

7. When you submit a bid (proposal to buy), the computer checks if your bid price is higher than the existing market bid and if you have enough cash to pay for the purchase. If both the answers are affirmative, the market bid is replaced by your bid on the screen in the light brown section.
8. When you submit an ask (a proposal to sell), the computer checks if your ask price is lower than the existing market ask and if you have the shares you propose to sell. If both the answers are affirmative, the market ask is replaced by your ask in the light yellow section of the screen.
9. CapLab warns you whenever your own bid (or ask) is the market bid (or ask) by displaying them in red font on your trading screens.
10. You can change your bid/ask price by one unit with each click on the plus or minus buttons.
11. Please remember that your bid/ask is not submitted until you press the enter key.
12. Caplab shows the identity of the investor who made the market bid/ask next to price and quantity. An S1 label under the Market Bid or Market Ask columns means investor number 1 made it.
13. You can buy shares in two different ways:

    - You can submit a bid under My Bid (brown section) and wait for someone else to accept it
    - If you see a market ask price in light yellow section at which you would like to buy, submit the same price and appropriate quantity under My Bid.

14. You can sell shares in two different ways:

    - You can submit an ask under My Ask (deep yellow section) and wait until someone to accept it.
    - If you see a market bid price (in light brown section) at which you would like to sell, submit the same price and appropriate quantity under My Ask.

15. If bid and ask cross (bid is above ask), transaction is executed at the price equal to the bid or ask that came first.
16. Whenever a transaction takes place, you will see the unit price of the latest transaction on the right hand side in the light blue section under label "L. Price."
17. This market has no book or queue. When a better one overtakes an unaccepted bid/ask, the latter is simply flushed from the system. It does not stay in the memory.
18. At the end of each period, the average trading price for the period will be shown under the heading "Avg. Price" at the right end of the trading screen. You may ignore the "BOP Hold." and "Value" columns.

Appendix 2: Supplementary Materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jedc.2007.01.008

Addendum: Further Analysis (Short-Horizon Session with Two Stocks)[23]

In short-horizon sessions of our experiment, we had a single stock to be traded in each market and observed that the price of the security was no longer anchored to its single dividend, even when it was certain and common knowledge. Price levels and changes became indeterminate; some exhibited positive bubbles of various sizes, while the one exhibited a negative bubble; some of bubbles showed some stability while the others grew rapidly. Indeterminacy seems to be an appropriate label for the great variation in the levels and changes in prices observed across sessions of the experiment.

When backward induction through time from a terminal dividend far into indefinite future generates indeterminacy, one may inquire if cross sectional induction across two or more such securities may still occur. If it does, the absolute level of prices of both the securities would be indeterminate, but the two prices will maintain a predictable relationship. To examine this proposition, we conducted a short-horizon treatment with two stocks, on November 9, 2002 at Yale University (Session A1). Stock 1 had a terminal dividend of 75 and Stock 2 had a terminal dividend of 120 and the rest of the treatment is the same as the one in Sessions 10 and 11. The experimental result is shown in Fig. 13.17. In the figure, we observe huge price bubbles for both the securities. Stock prices started in the 200–400 range in Period 1 and rose sharply in Periods 2 and 3. Some trades occurred around 1,600–2,000 in Period 4, before settling down around 1,000 in Period 5. These prices are far higher than the fundamental values of Stock 1 and Stock 2 (75 and 120, respectively). After Period 6, prices hovered noisily in the 1,000–1,200 range, until the session was terminated in Period 18. These price bubbles are generally similar to those observed in single-security short-horizon Sessions 1, 2, 8–11.

In spite of the indeterminacy of the absolute level of prices of both the securities, their co-movement and mutual relationship is noteworthy. When the price of one security rises (drops), so does the price of the other; product moment correlation between period-by-period average prices of the two securities is 0.693. In addition, the mean of the period-by-period average price of Stock 2 is 86.46 greater than the mean of the period-by-period average price of Stock 1 (Student's t-statistic = 0.871). Note that the terminal dividend of Stock 2 (120) exceeds the terminal dividend of Stock 1 (75) by 45 points.

[23]This addendum has been newly written for this book chapter.

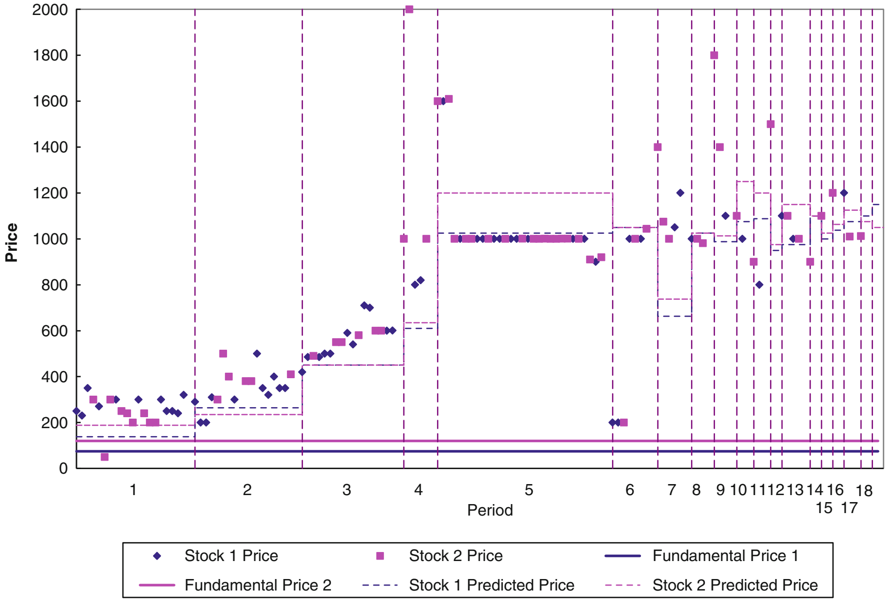


Fig. 13.17 Stock prices for Session A1 (short-horizon session. Two stocks traded)

Why do we observe this cross-sectional anchoring of the prices of the two securities in spite of the price levels becoming so completely unhooked from their dividend anchors placed in distant and indefinite future?

Recall that in our markets, short-term investors cannot backward induct the price from future (terminal) dividend. Without dividend anchors, they may try to find some available anchor for the valuation of shares. One possibility is, as we saw in Table 13.3, that they try the forward induction. Investors use the past and current prices as an anchor even if they have little rational reason to predict that the future will be like the past and current. Another possibility is that investors use the prices of other stocks as an anchor for the valuation of one stock; if they observe some stock prices rise (fall), they raise (or lower) the others as well. In the situation where investors do not have any concrete method of the valuation (such as the backward induction), this cross anchoring may occur just because information of other stocks' prices are available to investors. This possibility explains the co-movement of Stock 1 price and Stock 2 price in Session A1. It also explains the popular use of "comps", such as P/E ratios for other firms in the industry, in professional financial analysis. When nothing else is available, people hang on to "whatever anchor is available at hand" Shiller (2000, p. 137).

References

Ackley G (1983) Commodities and capital: prices and quantities. Am Econ Rev 73:1–16

Allen F, Gorton G (1993) Churning bubbles. Rev Econ Stud 60:813–836

Allen H, Taylor MP (1990) Charts, noise and fundamentals in the London foreign exchange market. Econ J 100:49–59

Allen F, Morris S, Postlewaite A (1993) Finite bubbles with short sale constraints and asymmetric information. J Econ Theory 61:206–229

Allen F, Morris S, Shin SH (2006) Beauty contests, bubbles and iterated expectations in asset markets. Rev Financ Stud 19:719–752

Biais B, Bossaerts P (1998) Asset prices and trading volume in a beauty contest. Rev Econ Stud 65:307–340

Blanchard O, Watson MW (1982) Bubbles, rational expectations and financial markets. In: Wechtel P (ed) Crises in the economic and financial system. Lexington Books, Lanham

De Long JB, Shleifer A, Summers L, Waldmann R (1990a) Noise trader risk in financial markets. J Polit Econ 68:703–738

De Long JB, Shleifer A, Summers L, Waldmann R (1990b) Positive feedback investment strategies and destabilizing rational speculation. J Financ 45:375–395

Dow J, Gorton G (1994) Arbitrage chains. J Financ 49:819–849

Fama EF (1991) Efficient capital markets: II. J Financ 44:1575–1617

Forsythe R, Palfrey TR, Plott CR (1982) Asset valuation in an experimental market. Econometrica 50:537–567

Frankel JA, Froot KA (1987) Using survey data to test standard propositions regarding exchange rate expectations. Am Econ Rev 77:133–153

French K, Roll R (1986) Stock market variances: the arrival of information and the reaction of traders. J Financ Econ 17:5–26

Froot KA, Scharfstein DS, Stein JC (1992) Herd on the street: informational inefficiencies in a market with short-term speculation. J Financ 47:1461–1484

Goetzmann W, Massa M (2002) Daily momentum and contrarian behavior of index fund investors. J Financ Quant Anal 37:375–389

Hirota S, Sunder S (2007) Price bubbles sans dividend anchors: evidence from laboratory stock markets. J Econ Dyn Control 31:1875–1909

Hommes CH (2006) Heterogeneous agent models in economics and finance. In: Tesfatsion L, Judd KL (eds) Handbook of computational economics, vol 2: agent-based computational economics, chapter 23, North-Holland, pp 1109–1186

Hommes CH, Sonnemans J, Tuinstra J, van de Velden H (2005) Coordination of expectations in asset pricing experiments. Rev Financ Stud 18:955–980

Jagadeesh N, Titman S (2001) Momentum. J Financ 56:699–720

LeBaron B (2006) Agent-based computational finance. In: Tesfatsion L, Judd KL (eds) Handbook of computational economics, volume 2: agent-based computational economics, chapter 24, North-Holland, pp 1187–1233

Lei V, Noussair CN, Plott CR (2001) Nonspeculative bubbles in experimental asset markets: lack of common knowledge of rationality vs. actual irrationality. Econometrica 69:831–859

LeRoy SF (2003) Rational exuberance. Mimeo

Miller M, Modigliani F (1961) Dividend policy, growth and the valuation of shares. J Bus 34:411–433

Moore R (1998) The problem of short-termism in British industry, Economic notes no. 81. Libertarian Alliance, London

Plott CR, Sunder S (1982) Efficiency of experimental security markets with insider information: an application of rational-expectations models. J Polit Econ 90:663–698

Scheinkman J, Xiong W (2003) Overconfidence and speculative bubbles. J Polit Econ 111:1183–1219

Shiller RJ (1981) Do stock prices move too much to be justified by subsequent changes in dividends. Am Econ Rev 71:421–436
Shiller RJ (2000) Irrational exuberance. Princeton University Press, Princeton
Smith VL, Suchanek GL, Williams AW (1988) Bubbles, crashes, and endogenous expectations in experimental spot asset markets. Econometrica 56:1119–1151
Stiglitz JE (1990) Symposium on bubbles. J Econ Perspect 4:13–18
Sunder S (1995) Experimental asset markets: a survey. In: Kagel JH, Roth AE (eds) Handbook of experimental economics. Princeton University Press, Princeton
Taylor MP, Allen H (1992) The use of technical analysis in the foreign exchange market. J Int Money Financ 11:304–314
Tirole J (1982) On the possibility of speculation under rational expectations. Econometrica 50:1163–1181
Tirole J (1985) Asset bubbles and overlapping generations. Econometrica 53:1499–1528
Tonello M (2006) Revisiting stock market short-termism. Conference Board, New York
White H (1980) Heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48:817–838
Williams AW (1987) The formation of price forecasts in experimental markets. J Money Credit Bank 19:1–18

Part V
Experimental Markets

Chapter 14
Revenue Non-equivalence Between the English and the Second-Price Auctions: Experimental Evidence

Chew Soo Hong and Naoko Nishimura

Abstract Under second-price sealed bid auctions, when bidders have independent private valuations of a risky object, submitting one's valuation is no longer dominant for non-expected utility bidders. This yields a breakdown in revenue equivalence between English auctions and second-price auctions for non-expected utility bidders. In an experimental auction market selling a single risky object, we find that an English auction yields higher seller revenue than the corresponding second-price auction. Further, the direction of revenue difference is supported by Nash equilibrium bidding behavior of betweenness-conforming non-expected utility bidders, under the additional hypothesis of bidders displaying a weak form of Allais type behavior.

Keywords Auction • Experiment • Allais paradox

JEL Classification Codes A10, C91, D44, D81

1 Introduction

Auctions are widely used to determine the allocation of commodities ranging from art pieces to condominiums. The English oral auction or ascending bid auction is the most familiar form of auction. Vickrey (1961) proposed the second-price

The original article first appeared in the Journal of Economic Behavior & Organization 51: 443–458, 2003. A newly written addendum has been added to this book chapter.

C.S. Hong
Department of Economics, National University of Singapore, Singapore, Singapore
e-mail: chew.soohong@gmail.com

N. Nishimura (✉)
Faculty of Economics, Shinshu University, Matsumoto, Japan
e-mail: nnaoko@shinshu-u.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_14

sealed-bid auction in a seminal paper.[1] In the subsequent literature on auctions, the independent private values (IPV) model has received much attention. Under the IPV model, bidders have independent, private, and deterministic valuations of a risk-free auctioned object. Consequently, bidding one's reservation value constitutes the dominant strategy equilibrium for both the English and the second-price auctions. This gives rise to the *revenue equivalence theorem*, the English and the second-price auctions generate the same revenue. An experimental validation of this equivalence was reported in Coppinger et al. (1980).

This chapter extends the Coppinger-Smith-Titus study by introducing uncertainty into the valuation of the auctioned object. The observed bidding behavior is tested against the implications of expected utility versus non-expected utility preferences. Our test hypotheses follow from the derivation of Nash equilibrium bidding strategies reported in several recent papers that do not require bidders to possess expected utility preferences. For the IPV model under uncertainty, Chew (1989) shows that the symmetric Nash equilibrium bid under the second-price auction is no longer demand-revealing for bidders possessing the more general betweenness-conforming preferences. He provides a condition linked to Allais type behavior for the optimal second-price bids to be bounded by the bidder's reservation price for the auctioned object. In a similar setting, Karni and Safra (1989) show that the Nash equilibrium bidding behavior under the English auction remains demand-revealing if and only if preferences satisfy betweenness. Thus, betweenness does not guarantee the revenue equivalence between the English and the second-price auctions.

Neilson (1994) shows that the Nash equilibrium bids under the second-price auction are lower than bidders' reservation values if their betweenness-conforming preferences satisfy Machina's (1982) Hypothesis II. Chew and Nishimura (2002) obtain, independently, similar Nash equilibrium bidding behavior based on a weaker Allais type behavioral condition which has received greater empirical support (Harless and Camerer 1994). The main theorem in the Chew-Nishimura paper provides the theoretical setting for the experimental test reported in this chapter.

Section 2 describes the model and the testable implications. Section 3 discusses the experimental design and the test results of our hypotheses. Concluding remarks are given in Sect. 4.

2 Bidding Strategies

Consider an auction where n agents bid competitively for a single indivisible object. The valuation of the object is risky and takes the form of a lottery with prizes $\boldsymbol{x} = \mathbf{n}\ (x_1, x_2, \cdots, x_k)$ and probabilities $\boldsymbol{p} = (p_1, p_2, \cdots, p_k)$. Each prize x_j is

[1]In an English auction, the auctioneer gradually raises the floor price. The last remaining bidder wins at the last announced price level. In the second-price auction, the highest bidder wins and pays a price equal to the second highest bid submitted. The reader is referred to McAfee and McMillan (1987) for a survey of various types of auctions.

drawn randomly from an interval $\left[\underline{x}_j, \overline{x}_j\right]$. Each vector $\boldsymbol{x}$ of realizable prizes is independently assigned to each player. For a given realization of $\boldsymbol{x^i}$, the auctioned object may be represented as a distribution $F_{\boldsymbol{x}^i} \equiv \sum_{j=1}^{k} p_j \delta_{x_j^i}$, which belongs to the set D of simple probability distributions with finite supports. When bidder i obtains the auctioned object $F_{\boldsymbol{x}^i}$ at a cost of b, we denote such an outcome by $F_{\boldsymbol{x}^i - b}$.

Let $U^i : D \rightarrow \Re$ be bidder i's continuous, betweenness-conforming utility functional.[2] There may be heterogeneity among utility functionals of bidders, $\left(U^1, \cdots, U^n\right)$. In this setting, bidder i's reservation price θ^i for the auctioned object $F_{\boldsymbol{x}^i}$ is defined by

$$U^i\left(F_{\boldsymbol{x}^i - \theta^i}\right) = U^i\left(\delta_{y^i}\right),$$

where y^i is the bidder i's current wealth. Assume that common knowledge includes all structural details of the auction except for the realized values $\boldsymbol{x}$. This informational setting corresponds to the independent-private-valuations (IPV) model in the literature. Under a similar IPV setting, Karni and Safra show that the English auction has a dominant strategy equilibrium where the bidder's bid b^e always equals his reservation value θ if and only if his preference exhibits the betweenness property. In the second-price auction, since we no longer have dominant strategies, Nash equilibrium provides the natural competitive equilibrium concept. Neilson demonstrates that the bidder's optimal bid b^e under the English auction is higher or lower than his optimal bid b^s in second-price auction depending on whether his preference satisfies Machina's Hypothesis II or its opposite.

For individual preferences satisfying Hypothesis II, the corresponding indifference curves fan out over the entire probability simplex of lotteries with three fixed outcomes. However, the complete fanning-out hypothesis has not found strong empirical support. In particular, Harless and Camerer's experimental study provides partial support of the fanning-out hypothesis, especially when the outcomes of lotteries are restricted to gains relative to the status quo payoff.

Chew and Nishimura (2002) derive the Nash equilibrium bidding behavior of betweenness-conforming players and show that the English auction yields higher revenue for the seller under a weaker condition than Hypothesis II: The indifference curves at the status quo are not steeper than the indifference curves at the higher utility levels. We label the choices induced by this condition as *Allais type* and its opposite as *counter Allais type*. The following theorem taken from Chew and Nishimura summarizes this result and provides the basis of our test hypotheses.

Theorem *Let $U^i : D \rightarrow \Re$, $i \in \{1, \cdots, n\}$, be bidder i's betweenness conforming utility functional, which is smooth and monotone. If each U^i displays Allais type choice behavior, then the equilibrium price in the second-price auction is bounded*

---

[2]The betweenness property requires the utility of a probability mixture between two lotteries to be intermediate in value between the utilities for the respective lotteries constituting the mixture. The class of betweenness- conforming preferences is axiomatized in Chew (1983, 1989), Fishburn (1983), and Dekel (1986).

from above by the equilibrium price in the English auction for any non-deterministic auctioned object. Moreover, if each bidder's utility functional U^i induces the counter Allais type choice, then the equilibrium price in the second-price auction is bounded from below by the equilibrium price in the English auction.

In our experiments reported in the next section, the main objective is to test whether bidding behavior is different for expected utility bidders as opposed to betweenness-conforming ones. The implication of expected utility for bidding behavior is stated as Hypothesis N. The alternative hypothesis, corresponding to Allais type choice behavior, is labeled as Hypothesis A. The test hypothesis opposite to Hypothesis A is labeled as Hypothesis CA.

Hypothesis N (Neutral) The seller's revenue in the English auction equals the seller's revenue in the second-price auction.

Hypothesis A (Allais) The seller's revenue is higher in an English auction than in a second-price auction.

Hypothesis CA (Counter Allais) The seller's revenue is lower in an English auction than in a second-price auction.

3 Experimental Analyses

To check for consistency with the Coppinger-Smith-Titus study, we start by investigating the bidding behavior for a deterministic auctioned object. We then proceed to identify the subjects' risk attitudes in terms of the actual sales prices for the risky objects. This is an intermediate step in our overall test of the implications on bidding behavior of bidders with betweenness-conforming preferences.

3.1 Experimental Design

We report auction experiments conducted at the University of Arizona. Groups of five to six student subjects bid for an object which is a binary lottery that yields a high resale value (H) with probability p and a low resale value (L) with probability 1- p. Each bidder is assigned a pair of H and L values drawn independently and uniformly from different intervals $\left[\underline{H},\ \overline{H}\right]$ and $\left[\underline{L},\ \overline{L}\right]$. The p value and the H and L value intervals are common knowledge. To implement the IPV model, bidders do not know the values assigned to competing bidders. To maintain a competitive environment, communication among the bidders is not permitted.

There are altogether 155 rounds of auctions classified into seven sessions with various combinations of p and the intervals of H and L values. The specific values for p and H and L intervals are summarized in Table 14.1. Each session consists

Table 14.1 Summary of experimental design

Session	Part	p	Number of auction rounds	H value interval	L value interval
A1	A1a	1	2	[1.26, 2.00]	
	A1b	0.50	5	[1.26, 2.00] and [0.75, 1.25]	
	A1c	0.75	5	[0.76, 1.25] and [0.25, 0.75]	
				[1.76, 3.00] and [0.50, 1.75]	
				[1.51, 2.50] and [0.50, 1.50]	
				Above four pairs are rotated for each round	
A2	A2a	1	2	[0.00, 1.00]	
			3	[2.00, 3.00]	
			2	[3.00, 4.00]	
			1	[0.00, 2.00]	
	A2b	0.50	5	[4.00, 5.00] and [0.00, 1.00]	
	A2c	0.75	5	[8.00, 10.00] and [2.00, 3.00]	
				[3.00, 4.00] and [0.00, 2.00]	
				[5.00, 7.00] and [0.00, 3.00]	
				Above four pairs are rotated for each round	
B	Ba	1	12	[0.00, 5.00]	{0}
	Bb	0.5	7	[0.00, 5.00]	{0}
	Bc	0.75	5	[0.00, 5.00]	{0}
C	Ca	1	8	[0.00, 4.00]	{0}
	Cb	0.50	8	[0.00, 8.00]	{0}
	Cc	0.75	8	[0.00, 6.00]	{0}
D	Da	1	8	[0.00, 3.00]	{0}
	Db	0.50	8	[0.00, 6.00]	{0}
	Dc	0.75	8	[0.00, 4.00]	{0}
E	Ea	1	11	[0.00, 5.00]	{0}
	Eb	0.50	10	[0.00, 10.00]	{0}
	Ec	0.90	3	[0.00, 6.00]	{0}
F	Fa	1	10	[0.00, 10.00]	{0}
	Fb	0.50	10	[0.00, 10.00]	{0}
	Fc	0.90	9	[0.00, 6.00]	{0}

of three parts, depending upon the value of p. The first part of each session has $p = 1$. The second part has $p = 0.5$, followed by the third part with p being either 0.75 or 0.90. In addition, each part contains several rounds of auction, as listed in Table 14.1. The same group of student subjects bid through three parts of one session. Each subject receives freshly drawn H and L values at the start of every round of the auction.

In each round, two identical objects are sold, the first by the second-price sealed-bid auction and the second by the English auction. The sale by the English auction is conducted before the outcome of the second-price auction is revealed.

Table 14.2 Bids made by subjects in the seventh round of Cb

Bidder	English auction	Second-price auction	Assigned *H* value
1	0.25	0.75	1.17
2	2.50	0.75	5.95
3	3.50	5.00	7.60
4	3.50	3.25	7.23
5	0.00	0.01	0.72
6	0.00	0.23	0.45

The reason for the sequence is obvious. Running a sealed-bid second-price auction without announcing the outcome does not reveal information about bidder's private valuations. If the sequence is reversed, the price level at which bidders withdraw from bidding in an English auction will reveal information about their private valuations. Table 14.2 illustrates an example of one auction round in session Cb.[3] This sequence rule is not applicable in the certainty parts of sessions B – F, (i.e., Ba, Ca, Da, Ea, and Fa) where only one object is sold via the second-price auction.[4]

The bid increment in the English auction is US$0.05 for sessions A1 and A2, where the ranges of H and L values are relatively narrow, and is US$0.25 for the other sessions. Bids are restricted to be non-negative. All sealed-bids are rounded to the nearest penny.[5] Each subject receives a beginning balance of US$10.00 for participating in the experiment. After each auction round, the outcome (i.e., high or low) of the risky auctioned object is realized according to the specified p, and the winning subject's net profit or net loss is recorded. Subjects are paid at the end of each part of a session unless they become "bankrupt" by having accumulated more than US$10.00 in losses. Such an event did not occur. The participants in sessions C, D, E, and F are experienced, having taken part in sessions A, B, or other similar but different auction experiments.

[3]In Table 14.2, a number in the first column denotes a bidder. The last column gives the realizations of H values assigned to the corresponding bidders. For example, the object will yield for bidder 1 a resale value of $1.17 with 50 % chance and US$0.00 otherwise. Every bidder knows that H values are randomly drawn from the interval [US$0.00, US$8.00] while the L value is $0.00 for all bidders.

[4]In other words, no English auction is conducted in the certainty parts of Sessions B–F. It is only conducted in sessions A1 and A2.

[5]There was no specific instruction for bids not to be in excess of private valuation. Kagel et al. (1987) argue that this sort of instruction has a potential guiding effect, which lowers observed bids.

4 Results

Nomenclature

$B(y,n)$	Cumulative binomial distribution of y observations out of n trials with parameter, $p = 0.5$
b^e	Observed market price in the English-auction
b^s	Observed market price in the second-price sealed-bid auction
$E(\,)$	Average
n	Number of non-zero observations
N	Number of relevant rounds
$\Pr(x < t)$ or $\Pr(x < z)$	Cumulative probability of the corresponding test score
$\Pr(x < W)$	Cumulative probability of the corresponding signed-ranks score
$S(\cdot)$	Standard deviations
t	t-test statistics
var( )	Variance
W+	Score of signed-ranks test for $d > 0$
W−	Score of signed-ranks test for $d < 0$
y	Number of observations of the data of interest
z^* or $z^*(\,)$	z^*-score

4.1 Certainty Case

Each session starts with the certainty part to familiarize the participants with the rules of the game governing the experiment. At the same time, it allows us to observe whether the subjects adopt the dominant strategy of bidding their assigned private values. We then check for revenue equivalence. If bidders adopt such dominant strategies, then the dominant-strategy price is given by the assigned private value of the second highest bidder in the corresponding round. Then we test each of the realized market price data, one from English auction and the other from the second-price sealed-bid auction, against the dominant-strategy price.

The certainty parts of English auctions are conducted in sessions A1 and A2. Figure 14.1a plots the assigned private value of the second highest bidder against the market price realized in the English auction from each of the ten certainty rounds in sessions A1 and A2. Similarly, Fig. 14.1b plots the assigned value of the second highest bidder against the market price in the second-price auction from each of the 59 certainty rounds for all sessions (i.e., A, B, C, D, E, and F).

Table 14.3 displays the results of statistical tests based upon a list of deviations: market price realized in one round – dominant-strategy price in the same round. In terms of non-parametric test results, the sign test and the Wilcoxon signed-ranks test, the data from English auctions indicate that the actual market prices are close to

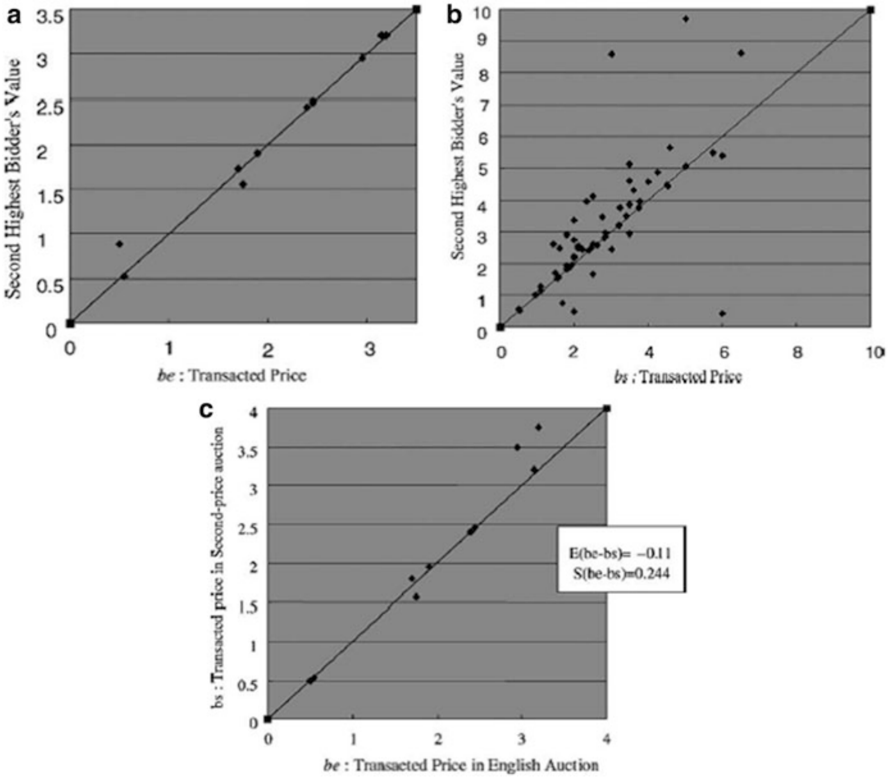


Fig. 14.1 Certainly case. **(a)** English-auction; **(b)** second-price auctions; **(c)** within-round price comparison (sessions A1 and A2)

the dominant-strategy prices. Neither of the two test results is sufficient to reject the null hypothesis that subjects adopt the dominant strategy. In two out of ten auction rounds, the market price strictly exceeds the second highest bidder's value. The signed-ranks test yields a score of 13.50, which is within the lower and upper critical values for a two-sided 10 % significance level. As to the second-price auctions, we observe a tendency for the market prices to be lower than the second-highest bidder's private values. This is supported by both a signed-ranks test and a sign-test at the 1 % level of significance. In 11 out of 55 cases of non-zero differences, the transacted price exceeds the private value of the second highest bidder. The corresponding signed-ranks test score is 304.

In order to obtain further evidence as to whether bidders adopt the dominant bidding strategy in both forms of auction, we need to run a parametric test, which may be a *t*-test. However, our experimental design does not allow a straightforward application of the t-test to our data, because there may be some session-specific effect that carries over to the whole data set. While the English auction data are generated only from part "a" of session A1 and A2, the second-price auction price

Table 14.3 Difference between market price and dominant strategy price in certainty case

	English auction	Second-price auction
$E(d)$	−0.03	−0.34
$S(d)$	0.14	1.35
y: Number of rounds with $d > 0$	2	11
n: Number of rounds with $d \neq 0$	9	55
N	10	59
$B(y, n)$	0.09	0.00
W^+	13.50	304
$P(x < W^+)$	0.18	0.00
z^*-score	−0.54	−1.92
$P(x < z^*)$	0.30	0.027

d actual market price – dominant strategy price

data come from part "a" of all seven different sessions. Each of the seven sessions differs in two aspects of the auction design and implementation.

First, every session has different structural parameters as displayed in Table 14.1. Even among all certainty parts, different intervals for the value of H are used in different sessions. As shown in Table 14.1, the interval can be small, less than US$1 (e.g., [US$1.26, US$2.00] in session A1, part a), or as large as $10 (e.g., [US$0, US$10] in session F, part a).

Second, while seven different groups of subjects participate in seven different sessions, the same group of subjects bid against each other throughout the same session. In each part within that session, the same group of subjects bid through multiple rounds. In the certainty (i.e., $p = 1$) part, for example, the number of rounds is 2, 8, 12, 8, 8, 11 and 10 for parts A1a, A2a, Ba, Ca, Da, Ea and Fa, respectively. For each round within a part, fresh values are randomly drawn and assigned to the subjects. In a given round, a subject assigned a relatively high value is more likely to emerge as the winner. As such, the identity of the winner and the second highest bid, which constitutes the price, are random. Nevertheless, having the same set of subjects in multiple rounds in a given session may generate some group specific effect over the entire data set across various sessions.

Although we cannot distinguish the effects of the aforementioned two kinds of session specific factors, the parameter design factor and the subject group specific factor, we cannot ignore the possible combined effect they may have on our data. This is an issue we need to address when we deal with a data set that is generated from multiple sessions. We handle this by normalizing the data by the square root of the weighted sum of the estimates of session variances across the relevant sessions.

Here, the data of interest are the list of differences between the market price and the corresponding dominant-strategy price from each round, for English auctions and second-price auctions separately. We will normalize the sum of such differences by the square root of the weighted sum of the estimates of session variances across the relevant sessions, namely A1 and A2 for the English auctions, and all seven sessions for the second-price auctions. We refer to the statistic thus obtained as z^*-

score,[6] shown in the bottom two rows of Table 14.3. For the data set plotted in Fig. 14.1a from English auction, the z^*-score is -0.54. The corresponding z^*-score is -1.92 for the data set plotted in Fig. 14.1b from the second-price auction. Neither of these two scores is sufficient to reject the null hypothesis of no deviation in a two-sided test at less than 5 % significance level. For the second-price auction data, however, the null hypothesis can be rejected at the 10 % level of significance.

To investigate further the bidding performance observed in the second-price auctions, we can ask the same question of whether the price data are consistent with the dominant bidding strategy by looking at the actual mean prices versus the theoretical mean prices calculated with a dominant bidding strategy.[7] Table 14.4 shows a list

---

[6] Suppose that we have K groups of two sets of random variables with n_k samples in each group $\left\{\{x_{ki}\}_{i=1}^{n_k}, \{y_{ki}\}_{i=1}^{n_k}\right\}_{k=1}^{K}$, whose mean and variance are $\left\{\mu_k^x, \mu_k^y\right\}_{k=1}^{K}$ and $\left\{\sigma_k^{x2}, \sigma_k^{y2}\right\}_{k=1}^{K}$ for $k \in \{1, 2, \cdots, K\}$, and $n = \sum_{k=1}^{K} n_k$, the total number of samples. There may exist some correlation between two random variables $\{x_{ki}\}, \{y_{ki}\}$, while $\left\{\{x_{ki}\}_{i=1}^{n_k}, \{y_{ki}\}_{i=1}^{n_k}\right\}$ and $\left\{\{x_{hi}\}_{i=1}^{n_h}, \{y_{hi}\}_{i=1}^{n_h}\right\}$, $k \neq h$, are considered independent but not necessarily i.i.d. Let $d_{ki} = x_{ki} - y_{ki}$. Under the hypothesis of expectation of d_{ki} in each group being 0, i.e., $E\left(\sum_{i=1}^{n_k} d_{ki}\right) = 0$, by the central limit theorem, $\sum_{i=1}^{n_k} d_{ki}$ follows the normal distribution with zero mean and variance $\sigma_{d_k^2}$, for $k \in \{1, 2, \cdots, K\}$. Let us define the value z as,

$$z = \frac{\sum_{k=1}^{K} \sum_{i=1}^{n_k} d_{ki}}{\sqrt{\sum_{k=1}^{K} \sigma_{d_k^2} \cdot n_k}}.$$

Then, under the assumption of $E\left(\sum_{i=1}^{n_k} d_{ki}\right) = 0$ and finite variance $\sigma_{d_k^2}$ for all $k \in \{1, 2, \cdots, K\}$, asymptotically z follows the normal distribution $N(0, 1)$. (See for example, Shiryayev (1984).) This z is different from the corresponding statistic when we treat whole n samples as a one big data pool. Here z takes care of the possible different tendency of each group by considering each $\sigma_{d_k^2}$. When these variances are unknown, the appropriate estimate for $\sigma_{d_k^2}$ is $s_{dk}^2 = \frac{\sum_{i=1}^{n_k} \left(d_{ki} - \overline{d}_k\right)^2}{n_k - 1}$, and we obtain the statistic z^* by substituting $s_{d_k^2}$ for $\sigma_{d_k^2}$ in z such as

$$z* = \frac{\sum_{k=1}^{K} \sum_{i=1}^{n_k} (x_{ki} - y_{ki})}{\sqrt{\sum_{k=1}^{K} s_{d_k^2} \cdot n_k}},$$

which follows the standard normal distribution asymptotically. In the main text, we call the value of this statistic z^*-score.

[7] The theoretical mean price pp from session k is given by

$$pp = \frac{n_k - 1}{n_k + 1} \left(\overline{v}_k - \underline{v}_k\right) + \underline{v}_k.$$

Table 14.4 Actual vs. predicted mean prices in the second-price auction in certainty case

Session	N	$E(b^s)$	Predicted $E(b^s)$	$S(b^s\text{-}pp)$[a]	Value range
A1	2	1.875	1.788	0.11	1.26–2.00
A2[b]	2	0.515	0.714	0.00	0.00–1.00
	3	2.783	2.714	0.386	2.00–3.00
	2	3.475	3.714	0.151	3.00–4.00
B	12	2.919	3.333	1.525	0.00–5.00
C	8	2.360	2.857	0.231	0.00–4.00
D	8	1.599	2.000	0.169	0.00–3.00
E	11	2.650	2.857	0.452	0.00–5.00
F	10	4.935	6.666	1.353	0.00–10.00

[a]Here "*pp*" denotes predicted mean price, see footnote 7
[b]One round from session A2 is omitted from the table, since there is no other round which has the same value range to generate any meaningful mean price data

of actual mean market prices from the second-price auctions and corresponding predicted mean prices for all sessions. The z^*-score is −0.51, which is not sufficient to reject the null hypothesis of no difference by a two-sided test at the 10 % significance level. Since only two out of nine non-zero observations are positive, there is not enough evidence to reject the null hypothesis at the 10 % significance level by the sign test. However, the null hypothesis is rejected by the signed-ranks test at the 5 % significance level with the positive score of 3.

Finally, we directly compare prices from the English and the second-price auctions to see whether their theoretically predicted equivalence holds. Such a comparison is feasible only for sessions A1 and A2 since English auction is not run in the certainty part of all other sessions. Figure 14.1c displays the pairs of the corresponding English-auction and second-price auction market prices. The paired t-test yields a score of −1.34, which does not allow us to reject the null hypothesis of no-deviation at less than two-sided 10 % significance level. Taking care of the session specific effect, the resulted z^*-score is −1.42, still not sufficient at two-sided 10 % significance level. The number of rounds with negative deviations is five out of seven non-zero observations, and the signed-ranks test score for the negative deviations is 22. Both results are insufficient to reject the null hypothesis by two-sided test at the 10 % significance level.

Overall, we do not have sufficiently strong evidence to conclude that the subjects do not adopt the dominant bidding strategy in the English auction and the second-price auction. This is consistent with the earlier reported experimental studies (Coppinger et al. 1980; Cox et al. 1982; Kagel et al. 1987; Kagel 1995) that support the dominant strategy bidding behavior for English auctions and similar but weaker support for the second-price auctions. For the latter, it has been observed that the data from second-price auctions generally display a higher degree of dispersion around the dominant strategy prediction. We do not detect overbidding in second-price auctions that was reported in Kagel et al. (1987) and Kagel (1995).

4.2 Uncertainty Case

In this section, we proceed to investigate bidding behavior when uncertainty is involved in the valuation of the auctioned object.

4.2.1 Risk Attitude

First, we ask whether the subjects bid as though they are risk neutral. Our theory predicts that a risk neutral bidder will bid the expected value of the risky object in both the English auction and the second-price auction. In contrast, a risk averse (preferring) bidder will withdraw at a price level lower (higher) than the expected private value at least in the English auction.

Let us first examine the English auction data from the second (denoted by "b", with $p = 0.5$) and third (denoted by "c", with $p = 0.75$ or 0.90) parts of all sessions. We compute the expected private value for each bidder[8] and identify the second highest expected private value as the predicted price for the English auction. We then compare such a predicted price with the corresponding realized market price from each round in the English auction. In a two-sided test against the null hypothesis of risk neutral bidding behavior, the entire data set yields z^*-score of -0.48, too insignificant to reject the null hypothesis. However, this does not imply risk neutral bidding. Recall that the subjects for sessions C, D, E, and F are experienced, having taken part in sessions A1, A2, B, or other similar auction experiments. The z^*-score on the data set from sessions C, D, E and F is -1.47, sufficient to reject the null hypothesis at the two-sided 15 % significance level. Table 14.5 presents such mean deviations for each uncertainty part for all sessions. When we apply the sign test and the signed-ranks test, shown within the parentheses in the lower part of Table 14.5, sessions with the more experienced bidders yield stronger indications of risk averse bidding behavior.

If bidders are risk neutral, then the market price realized in the second-price auction should also be equal to the second highest expected private value. Consequently, the English auction and the second-price auction of the same round should have the same market price. The next subsection offers, among other things, some evidence against such an equivalence, implying that bidders do not behave like risk neutral agents.

[8]The expected private value for bidder i will be $pH_i + (1 - p)L_i$, where H_i and L_i are the respective high and low values randomly drawn from the common, public known ranges for H and L and assigned to bidder i.

Table 14.5 Mean deviations of market prices from second highest expected values[a]

Session/part	Number of rounds	English auctions	
		$E(d_{var})$	$S(d_{var})$
A1b	5	−0.265	0.099
A1c	5	−0.128	0.176
A2b	5	0.293	0.532
A2c	5	0.708	0.221
Bb	7	0.675	0.598
Bc	5	0.180	0.387
Cb	8	−0.283	0.984
Cc	8	0.303	0.477
Db	8	0.336	1.277
Dc	8	−0.298	0.318
Eb	10	0.210	0.997
Ec	3	−0.957	1.242
Fb	10	−1.117	1.327
Fc	9	−0.180	1.586
$Average(E(d_{var}))$		−0.05 (−0.20)	
$S(E(d_{var}))$		1.05 (1.16)	
$z^*(d_{var})$		−0.48 (−1.47)	
$Pr(x < z^*)$		0.31 (0.07)	
y : Number of rounds with $d_{var} < 0$ n : Number of rounds with $d_{var} \neq 0$		47 (35) 95 (63)	
$B(y, n)$		0.50 (0.843)	
W^-: signed-ranks score with dv<0		2189 (1149)	
Pr($x < W^-$)		0.37 (0.83)	

Here d_{var} = market price – expected value of the object
[a]The figures in the parentheses are from sessions Cs and Ds

4.2.2 English Auction Versus Second-Price Auction

In the context of uncertain auction object valuations, we investigate the implications of the Allais type preferences for market prices resulting from the English auctions and the second-price auctions. The test hypotheses – Hypothesis A, Hypothesis CA, and Hypothesis N – are stated in Sect. 2.

Table 14.6 gives, for each uncertainty part of all sessions, the pooled means of the market prices realized in both the English and the second-price auctions. The question is whether the pooled means significantly differ under the two auction forms. The one-tailed t-test yields a t-value of 1.04, supporting Hypothesis A at 15 % level of significance. The higher English pooled means appear in eight out of 13 non-zero observations, supporting Hypothesis A at less than 13 % level of one-tailed significance. The signed-ranks test score for positive difference is 60, supporting of Hypothesis A at the 15 % level of one-tailed significance. Across all

Table 14.6 Pooled means in uncertainty case

		English auction		Second-price auction		
Session/part	Number of rounds	$E(b^e)$	$S(b^e)$	$E(b^s)$	$S(b^s)$	$E(b^e)$-$E(b^s)$
A1b	5	1.17	0.167	1.17	0.105	0
A1c	5	1.66	0.178	1.59	0.193	0.07
A2b	5	3.85	1.269	3.50	2.732	0.35
A2c	5	5.80	4.325	5.82	3.470	−0.01
Bb	7	2.17	1.042	2.42	1.018	−0.25
Bc	5	2.79	1.123	3.13	1.858	−0.34
Cb	8	2.69	1.835	2.61	0.606	0.08
Cc	8	3.53	0.508	3.25	0.253	0.29
Db	8	2.28	1.115	2.05	0.324	0.23
Dc	8	1.84	0.035	2.10	0.206	−0.26
Eb	10	4.15	1.267	3.00	0.597	1.15
Ec	3	3.42	2.021	2.92	0.896	0.50
Fb	10	2.28	0.992	2.94	3.094	−0.67
Fc	9	3.97	1.960	3.26	1.091	0.72

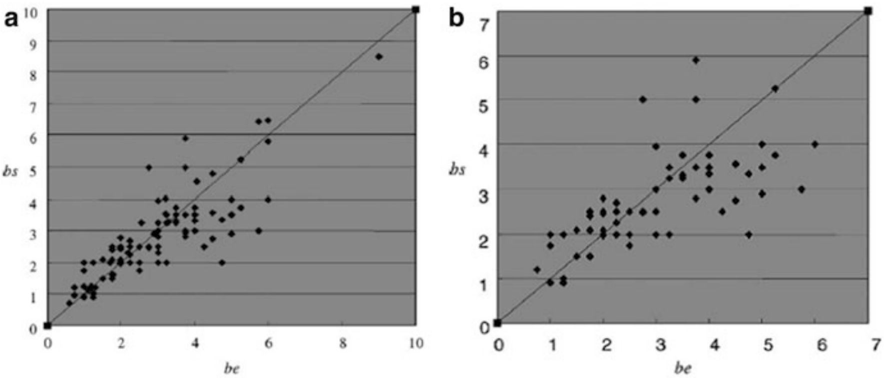


Fig. 14.2 Uncertainty case. English-auction price versus second-price auction price. (**a**) Within-round price comparison (all sessions); (**b**) within-round price comparison (sessions C, D, E, and F)

sessions, the mean prices for the English auction exceed the corresponding second-price auction ones by an average of US$0.14, whereas the average bid increment for the English auctions is US$0.18.

When we narrow our data set to the eight parts of sessions with experienced subjects, namely, C, D, E, and F, the mean price difference between the two auctionforms is more pronounced. The one tailed t-test supports Hypothesis A at the 5 % level of significance with t-value 1.64. The English auction yields a higher mean price in six out of eight cases with non-zero differences, also supporting Hypothesis A at the 4 % level of one-sided significance. The result of signed-ranks test improves to 10 % one-tailed significance in support of Hypothesis A with a

Table 14.7 Within-round price difference $d = b^e - b^s$ in uncertainty case

Session	N	$E(d)$	$S(d)$	t-test	Pr($x<t$)	z*	Pr($x<z^*$)	y	n	$B(y,n)$	W^+	Pr($x<W$)
All	96	0.15	0.835	1.72	0.96	2.05	0.98	46	81	0.90	1927.5	0.90
C, D, E, F	64	0.25	0.975	1.92	0.97	2.34	0.99	34	55	0.97	984.5	0.96

y number of rounds with $d > 0$, n number of rounds with $d \neq 0$, W^+ score of signed-ranks for $d > 0$

score of 27 for the positive differences. The average deviation across sessions C to F is US\$0.26, slightly above the corresponding average bid increment of US\$0.25 for English auctions.

Next, we compare the within-round English auction price and the second-price auction price for the same risky object. Figure 14.2a plots the entire data from all sessions while Fig. 14.2b excludes data from sessions A1, A2, and B. As summarized in Table 14.7, all four tests employed for data from all sessions support Hypothesis A. The paired two-sided t-test of the price differences rejects Hypothesis N at 8 % level of significance in favor of Hypothesis A. The z*-score yields the sharper result in favor of Hypothesis A at a two-sided 4 % level of significance. In 46 out of 81 non-zero observations, the prices in the English auction are higher. The result of sign test represented by the binomial probability score in Table 14.7 also supports Hypothesis A at the one-sided 10 % significance level, as does the signed-ranks test.

Comparing the two rows in Table 14.7 suggests that the evidence to reject Hypothesis N in favor of Hypothesis A is more pronounced when we restrict our data set to sessions C, D, E, and F. The significance level for the t-test improves to a two-sided 6 %, while the z*-score supporting Hypothesis A improves to a two-sided 2 % level of significance. The two-sided sign test and signed-ranks test against Hypothesis N are stronger at 6 % and 8 % significance level, respectively. The corresponding average price deviation of the English auction over the second-price auction is US\$0.25, coinciding with the average bid increment of US\$0.25 in the English auctions.

The observed non-equivalence between the English and the second-price auctions reveals at least two aspects of our subjects' preferences. First, they do display non-risk-neutrality. Second, there is a significant possibility that they possess non-expected utility preferences. By persistently supporting the alternative hypothesis consistent with Allais type choice behavior, our data suggests that subjects' preferences fall into the sub-class of non-expected utility preferences often referred to as Allais type preferences.

5 Conclusion

Our main results are twofold. First, when the auctioned object involves risk, we find experimentally that the English auction yields higher seller revenue than the second-price auction. This is inconsistent with the implications of Nash equilibrium behavior of expected utility bidders. Second, the direction of the observed revenue non-equivalence is consistent with Nash equilibrium behavior of betweenness-conforming bidders under an additional hypothesis that bidders display a weaker form of Allais type choice behavior.

Our results are also corroborated by those reported in a recent paper by Berg et al. (2005). They find that the valuations of risky auctioned objects using the Becker et al. (1964) procedure yield lower values than those inferred from bidding behavior in an English auction. This finding contributed to their view that risk attitudes may not be stable across institutions. Since the Becker-Degroot-Marschak procedure elicits valuations in much the same way as a second-price auction, the analysis in this chapter provides an alternative explanation of Berg, Dickhaut, and McCabe's finding.

Acknowledgments The authors would like to thank John Dinardo, Greg Engl, Cheng Hsiao, Stergios Skaperdas, Yo Sheena for their helpful comments. We acknowledge the research assistance of Vicky Bishop. Financial support from the National Science Foundation (SES 8810833) Research Grant Council (HKUST 6234/97H), and from Seimeikai 1992 is gratefully acknowledged.

Addendum: Follow Up Research on Auction Design Under Risk and Uncertainty[9]

Following Chew and Nishimura (2003)'s application of non-expected utility theory to explain non-equivalence between the English ascending-bid auction and the second-price sealed-bid auction for a risky object, our researches developed in the two venues; one is to reconsider a bidding behavior for a deterministic auctioned object by introducing bidders with reciprocal social preferences. The other is to take another look at risk preferences and ask about *quality* of risks.

The first research question arose from our study (Chew and Nishimura (2004))[10] investigating theoretically the alleged equivalence between the Dutch descending-bid auction and the first-price sealed-bid auction to see whether the bidding behavior of Allais type bidders contribute to generating any difference between the two auctions. Our model predicts that the Allais type bidders would bid higher in the Dutch auction than in the first-price auction. This is contradicted by the literature on laboratory auction experiment reporting the predicted expected price in the Dutch auction is lower than that in the first-price auction.

This led us to leave the issue of risk preferences and ask what else can account for bidders deviating from the prediction by the standard auction theory in terms of observed bidding behavior in laboratories. We revisit the revenue equivalence among four standard auction formats[11] for a deterministic object under the IPV

[9]This addendum has been newly written for this book chapter.

[10]The related work can be found in the recent paper by Nakajima (2011).

[11]Four auction formats are English ascending-bid auction, the second-price sealed-bid auction, Dutch descending-bid auction, and the first-price sealed-bid auction.

setting, when competing bidders have reciprocal social preferences. Nishimura et al. (2011) constructed the *intention-base* model of the kind of Rabin (1993) or Dufwenberg and Kirchsteiger (2004) to investigate the equivalence between English auction and the second-price auction where a bidder with lower value may choose to overbid in order to reduce the winner's surplus by making her pay more than the second highest value. We label such a bid exceeding value as a *spite bid*. The standard auction theory prescribes the best response for a bidder with higher value against such a spite bid just to place a higher bid to win as long as her winning payoff is positive. In contrast, our reciprocity model allows the higher value bidder to retaliate against such a spite bid by placing a bid just below the spite bid and let her opponent win with negative payoff. Such a negative interaction between bidders is more effective in English auction than in the second-price auction, because the ascending calling price in English auction eventually reveals the spite bid, which makes it easier for the higher value bidder to counteract. Thus, the equilibrium price in the second-price auction should be bounded from below by the price in English auction.

When bidders' values are unknown, the above negative interaction between bidders is not likely to occur since they no longer know their relative value positions. Then, the equilibrium prices in two kinds of auction should coincide in the incomplete information setting even under the negative reciprocity hypothesis. These theoretical predictions are tested experimentally, and we confirmed negative reciprocity at work in the complete information setting but not under the incomplete information setting.

Nishimura (2013) extended this approach to reinvestigate the revenue equivalence between the remaining pair of auction formats, namely the first-price sealed-bid auction and Dutch auction. It has been known in the experimental literature (Kagel (1995)) that the revenue from the first-price auction is higher than that from English auction or Dutch auction and that the observed bids in the first-price auction can be approximated by a bid function which is increasing and concave in value. Cox and Oaxaca (1996) argued that it is bidders' constant relative risk aversion that causes the higher revenue in the first-price auction and its equilibrium bid function to be concave. The risk aversion, however, cannot explain the non-equivalence in revenue between Dutch and the first-price auction.

Nishimura (2013) proposes bidders' reciprocal social preferences to explain the revenue non-equivalence as well as the concave bid function in the first-price auction. The crucial difference between two auctions mainly lies in that each step of the descending price ceiling in Dutch auction reveals the lesser spitefulness of one's opponent which allows the higher value bidder to wait for the price to fall further before she makes a retaliatory bid, whereas in the first-price auction there is no such partial revelation of the spite intention. We then proceed to report the results from experimental Dutch and the first-price auctions under both complete and incomplete information settings. The theoretical predictions based on reciprocity are confirmed in the complete information setting, and the same tendency carries over to the incomplete information setting as predicted.

As to the second question to go beyond a decision making under pure individualistic risk or uncertainty, Chew and Sagi (2008) developed a new platform of preferences under risk, called *source preference* which allows an individual to choose one risk over the other with the same objective probabilities depending upon how the risk itself is generated. The model can describe behavioral anomalies such as *familiarity bias*, relating to investors' inclination to concentrate disproportionally on investment opportunities in their own countries than elsewhere.

A strategic interaction is another important source of risk. There has been a stream of research that attempts to explain the inconsistency between choices made under non-strategic risk/uncertainty and those made under strategic risk/uncertainty, starting from Camerer and Karjalainen (1994). Fox and Tversky (1995), Fox and Weber (2002), and Eichberger and Kelsey (2011) consider such inconsistency to arise from the difference between the choices made under risk and under ambiguity, given that human behavior is intrinsically not easy to predict so that individuals perceive their opponent's strategic choice as ambiguous. If this is so, then one should be more willing to face risks arising from non-strategic situation than from strategic situation, because ambiguity aversion is one of the most commonly observed preferential traits. In a study using neuroimaging, Chark and Chew (2013), reported that subjects accept a discount to play coordination strategically rather than randomly, and showed that their results were consistent with the predictions of source-dependent expected utility model.

Our latest study, Chew, Mao, and Nishimura (2014), turns out to be in line with Chark and Chew. We experimentally examine the demand for a sweepstake to see whether we observe a *favorite-longshot b*ias (FLB)[12] widely reported in the racetrack betting literature. A sweepstake awards a large prize with a small probability. In particular, we focus on a variable prize sweepstake in which a single winner receives 90% of the total receipts. Then, the expected value of purchasing a ticket is negative, so that any individual who is risk averse in the usual sense[13] would not purchase a ticket. We find a significant incidence of FLB reflected in sweepstakes purchase over population sizes ranging from 2 to 141, and a greater tendency for FLB among those who exhibit *longshot preference* (LSP)[14] over fixed-odds lotteries. We found, however, mixed support for FLB, that is, subjects showed a greater demand for 28 population sweepstakes than for 141 population sweepstakes including those with LSP and those who are risk averse. Further and intriguingly, we observe significant demands for 2-person sweepstakes even among risk averse subjects. In other words, they are willing to take half-half chance

[12]The favorite-longshot bias is a well observed phenomenon of higher demand for higher odds bet (= longshot) in the racetrack and other competitive gambles including sweepstakes.

[13]We identify risk averse subjects by those who chose a sure outcome of 0 over a lottery with half chance of getting x and half chance of getting $-x$.

[14]An individual is identified as having LSP if she chooses a longshot lottery yielding a large prize with small probability over a lottery with moderate probability of winning a moderate prize. Chew and Tan (2005) characterized LSP and showed that individuals who are risk averse can exhibit LSP under specific functional form of weighted utility and rank-dependent utility which corresponds to cumulative prospect theory for risk with objective probabilities.

risk generated in a two-person sweepstake market while they decline exogenously generated even-chance bets in an individual choice.

Our findings point to the notion that subjects are more willing to face risks generated in a small population sweepstakes which has tighter strategic interdependency than from a large population sweepstakes where strategic interdependency gets diluted. Our observations unveil the existence of an additional element arising from an interactive nature of sweepstake market that induces our subjects to participate in the sweepstake market along with the effect of LSP resisting the effects of risk aversion. Such an additional element may capture a recreational aspect of sweepstake demand which can only be experienced through interaction among participants. After all, the two research questions we started with at the beginning of this addendum appear to have a shared component concerning the role of intentions in a decision making that differentiates quality of the interaction among players and quality of interactive risks.

References

Becker G, Degroot MH, Marschak J (1964) Measuring utility by a single-response sequential method. Behav Sci 9:226–232

Berg J, Dickhaut J, McCabe K (2005) Risk preference instability across institutions: a dilemma. Proc Natl Acad Sci U S A 11:4209–4214

Camerer C, Karjalainen R (1994) Ambiguity-aversion and non-additive beliefs in non-cooperative games: experimental evidence. In: Munier B, Machina MJ (eds) Models and experiments in risk and rationality, vol 29, Theory and Decision Library, Series B, Mathematical and statistical methods. Kluwer, Boston, pp 325–258

Chark R, Chew SH (2013) A neuroimaging study of preference for strategic uncertainty, Working paper, National University of Singapore

Chew SH (1983) A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. Econometrica 51: 1065–1092

Chew SH (1989) Axiomatic utility theories with the betweenness property. Ann Oper Res 19: 273–298

Chew SH, Nishimura N (2002) Equilibrium bidding strategies under the English and the Second-price auctions. In: Rapoport A, Zwick R (eds) Experimental business research. Kluwer, Boston, pp 197–205

Chew SH, Nishimura N (2003) Revenue non-equivalence between the English and the second-price auctions: experimental evidence. J Econ Behav Organ 51:443–458

Chew SH, Nishimura N (2004) Allais preference and paradoxical bidding behavior in the Dutch and the first-price auctions. Report for Grant-in-Aid for Scientific Research 2001–2002

Chew SH, Sagi JS (2008) Small worlds: modeling attitudes toward sources of uncertainty. J Econ Theory 139(1):1–24

Chew SH, Tan G (2005) The market for sweepstakes. Rev Econ Stud 72:1009–1029

Chew SH, Mao J, Nishimura N (2014) In search of the utility from gambling: an experimental study of the demand for sweepstakes. Mimeo

Coppinger V, Smith V, Titus J (1980) Incentives and behavior in English, Dutch, and sealed-bid auctions. Econ Inq 18(1):1–22

Cox J, Oaxaca R (1996) Is bidding behavior consistent with bidding theory for private value auctions? Res Exp Econ 6:131–148

Cox J, Roberson B, Smith V (1982) Theory and behavior of single object auctions. In: Smith V (ed) Research in experimental economics. JAI Press, Greenwich
Dekel E (1986) An axiomatic characterization of choice under uncertainty. J Econ Theory 40: 304–318
Dufwenberg M, Kirchsteiger G (2004) A theory of sequential reciprocity. Games Econ Behav 47:268–298
Eichberger J, Kelsey D (2011) Are the treasures of game theory ambiguous? Econ Theory 48(2–3): 313–339
Fishburn P (1983) Transitive measurable utility. J Econ Theory 31:297–317
Fox CR, Tversky A (1995) Ambiguity aversion and comparative ignorance. Q J Econ 110(3): 585–603
Fox CR, Weber M (2002) Ambiguity aversion, comparative ignorance, and decision context. Organ Behav Hum Decis 88(1):476–498
Harless DW, Camerer CF (1994) The predictive utility of generalized expected utility theories. Econometrica 62:1251–1289
Kagel JH (1995) Auctions: a survey of experimental research. In: Kagel JH, Roth AE (eds) Handbook of experimental economics. Princeton University Press, Princeton, pp 501–586
Kagel JH, Harstad RM, Levin D (1987) Information impact and allocation rules in auctions with affiliated private values: a laboratory study. Econometrica 55:1275–1304
Karni E, Safra Z (1989) Dynamic consistency, revelations in auctions, and the structure of preferences. Rev Econ Stud 56:421–434
Machina M (1982) Expected utility analysis without the independence axiom. Econometrica 50:277–323
McAfee RP, McMillan J (1987) Auctions and bidding. J Econ Lit 25:699–738
Nakajima D (2011) First-price auctions, Dutch auctions, and buy-it-now prices with Allais paradox bidders. Theor Econ 6:473–498
Neilson W (1994) Second price auctions without expected utility. J Econ Theory 62:136–151
Nishimura N (2013) Non-equivalence between Dutch and First-price auctions with reciprocal bidders. Mimeo, first draft (2009)
Nishimura N, Cason T, Saijo T, Ikeda K (2011) Spite and reciprocity in auctions. Games 2(3): 365–411
Rabin M (1993) Incorporating fairness into game theory and economics. Am Econ Rev 83: 1281–1302
Shiryayev AN (1984) Probability. Springer, Berlin
Vickery W (1961) Counterspeculation, auctions, and competitive sealed tenders. J Financ 16:8–37

Chapter 15
An Experimental Test of a Committee Search Model

Yoichi Hizen, Keisuke Kawata, and Masaru Sasaki

Abstract The objective of this chapter is to design a laboratory experiment for an infinite-horizon sequential committee search model in order to test some of the implications obtained by the model in Albrech et al. (J Econ Theory 145:1386–1407, 2010) (AAV). We find that, compared with single-agent search, the search duration is longer for committee search under the unanimity rule, but is shorter for committee search in which at least one vote is required to stop searching. In addition, according to estimates from round-based search decisions, subjects are more likely to vote to stop searching in committee search than in single-agent search. This confirms that agents are less picky in committee search. Overall, the experimental outcomes are consistent with the implications suggested by the AAV model. However, despite the prediction from the AAV model, we could not obtain a significant outcome in relation to the size order of the probabilities of voting to stop searching in committee search for the various plurality voting rules.

Keywords Experiments • Committee search • Plurality voting rules

JEL Classification C91, D83

The original article first appeared in the European Economic Review 61: 59–76, 2013. A newly written addendum has been added to this book chapter.

Y. Hizen
Kochi University of Technology, Kochi, Japan
e-mail: hizen.yoichi@kochi-tech.ac.jp

K. Kawata (✉)
Hiroshima University, Hiroshima, Japan
e-mail: keisuke@hiroshima-u.ac.jp

M. Sasaki
Osaka University and IZA, Suita, Japan
e-mail: sasaki@econ.osaka-u.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_15

1 Introduction

The decision mechanism of agents regarding whether to stop or continue searching has been considered in many fields of economics, including labor economics, monetary economics, macroeconomics, and industrial organization. Very recently, there has been an emerging interest in committee search, in which a decision is made by a group of multiple agents rather than by a single agent. An evolving theoretical literature duly analyzes the properties of decision-making in the case of committee search (Albrecht et al. 2010; Compte and Jehiel 2010). However, to our best knowledge, no corresponding empirical studies have been conducted, mainly because it is generally difficult to collect data on committee search processes. This chapter is thus the first attempt to provide experimental evidence on committee search and to test the theoretical implications obtained by the committee search model in Albrecht et al. (2010) (hereafter AAV). Overall, we find that our experimental outcomes are consistent with those obtained by the AAV model.

In the AAV model, a group engages in search activity to fill a vacant employment position or searches for a new house as a family. The members of the group or family then decide by vote whether to hire a newly encountered worker or to purchase a house. The AAV model assumes that members are homogeneous with respect to preferences and that each member draws a value from an identical and independent distribution across members. The model then compares the member's reservation value in single-agent search and committee search in an environment where the drawn value differs among the members under various plurality voting rules. The main predictions of the AAV model are that members are less picky in committee search than in standard single-agent search in the sense that each member's threshold is lower and that the members' thresholds vary by voting rule. Another implication obtained in the AAV model is that the search duration increases with the number of votes required to stop the committee search process.

The search environment characterized in the model is usually far removed from the environment observed from the micro data. Therefore, the search environment cannot be perfectly duplicated using micro data. However, we can recreate this search environment in the laboratory using controlled treatments. In recent years, many studies have been devoted to experimental analysis of the single-agent search model (Cox and Oaxaca 1989; Harrison and Morgan 1990). This experimental task is empirically tractable and attractive for testing the implications of sequential search models. Experimental studies on the sequential search model have proliferated, covering a range of topics such as the effect of unemployment benefit sanctions on individual search behavior (Boone et al. 2009) and the differences in individual search behavior by attitudes toward loss and risk (Schunk 2009). In addition, Schunk and Winter (2009) explored the reasons why in many of these studies agents stop searching earlier than what is theoretically optimal.

We expand upon this body of work by examining the decision-making processes of multiple agents engaged in a committee search activity. The main feature of our experimental design is that we conduct three types of game to identify exactly the predictions of the AAV model where agents are assumed to be homogeneous with respect to preferences. Game A provides the benchmark as a standard single-agent search task, Game B is a committee search task where three group members have a common value drawn from a distribution, and Game C is a committee search task where three group members each draw different values from the same distribution.[1] The difference between Games A and B is attributable to heterogeneity among members with respect to their risk and loss attitudes, time preferences, and any unobserved characteristics.[2] The difference between Games A and C arises from the heterogeneity among members already mentioned, plus additional heterogeneity in the sense of the different values the other members independently draw from the same distribution. Therefore, the difference between the above two differences is caused only by the second form of heterogeneity among members, in that the values drawn by the other members of the group are different. This is similar to the AAV model. In addition, we design three subgames for each of Games B and C: Subgame 1 adopts a plurality voting rule in which the committee search activity is stopped if at least one member votes to stop searching (the one-vote rule); in Subgame 2, the committee search activity is stopped only if at least two-thirds of members vote to stop searching (the majority rule); and in Subgame 3, the committee search activity is stopped only if all members vote to stop searching (the unanimity rule). The results of these subgames provide evidence concerning the effect of voting rules.

We conducted experimental tests of an infinite-horizon sequential search model with a 5 % probability that the search coercively ends. With this experiment, the focus is on exploring (i) the search duration and (ii) the probability of voting to stop searching in committee search with various plurality voting rules compared with single-agent search. Our finding regarding search duration is that, compared with single-agent search, the search duration is longer for committee search with the unanimity rule but shorter for committee search with the one-vote rule, after controlling for the heterogeneity of preferences among group members regarding risk and loss attitudes, time preferences and any unobserved factors. However, in our experiments, the difference in search duration between single-agent search and committee search with the majority rule is statistically unclear. These outcomes imply that two effects operate to determine this relationship.

The first effect is that it takes more time to reach an agreement in committee search with the majority rule than it does in single-agent search. Thus, on the one hand, the committee search structure with the majority rule lengthens the search duration. However, on the other hand, the second effect is that committee search with the majority rule lowers each subject's reservation value because she or he is less picky, thereby shortening the search duration. In our experiment, under the majority rule, these opposing effects cancel each other out, leading to the conclusion that there is no difference in search duration between single-agent search and committee search with the majority rule. These results imply that search duration

[1]This implies that each member draws a value from an independent and identical distribution.

[2]The heterogeneity of preferences among members in a group is ruled out in the AAV model.

is increasing in the number of votes required to stop the committee search, which supports the first part of Proposition 5 in the AAV model. In addition, a comparison of the search duration between committee search with the unanimity rule and single-agent search shows that the search duration is increasing in group size, holding the unanimity rule fixed. The single-agent search structure is regarded as a special case of the unanimity rule. This supports the implication obtained from the second part of Proposition 4 in the AAV model.

Our second focus is on identifying differences in a subject's willingness to accept a drawn value between single-agent search and committee search. To do this, we estimate the average marginal effects from a probit model to examine the determinants of the probability of voting to stop searching using data from every round-based decision about whether to vote to stop searching. Our findings are that subjects are more likely to vote to stop searching in committee search than in single-agent search, and that this outcome is strongly observed in committee search with the one-vote and majority rules. These estimated results confirm the threshold effect referred to in the AAV model (Proposition 2 therein), in the sense that subjects lower their reservation values in committee search and thus become less picky about the standard of acceptance.

However, our experimental outcome cannot statistically support the AAV model's prediction in terms of the size order of the reservation values among the types of committee search with the various plurality voting rules. The AAV model predicts that the reservation value is either hump-shaped or monotonically increasing in the number of votes required to stop the search (the second part of Proposition 5 in the AAV model). Unfortunately, given our limited sample size, we cannot significantly support this prediction. However, we can say that the probability of voting to stop falls from the one-vote rule to the majority rule and then falls even further from the majority rule to the unanimity rule. That is, the reservation value is quantitatively increasing in the number of votes required to stop the search, which is consistent with the AAV model's prediction. We find that subjects stop searching earliest in committee search with the one-vote rule, followed by the majority rule, and then the unanimity rule. This result arises from the subjects' preferences regarding the two negative externalities (one that the committee search activity stops despite the subject's preference to continue searching, and the other that the committee search activity continues despite a subject's preference to stop searching). For the most part, the one-vote and unanimity voting rules are, respectively, most strongly influenced by the first and second externalities, whereas the majority rule is influenced by both externalities, but only moderately. Our results imply that subjects who participated in this experiment incurred a larger disutility from the negative externality whereby the committee search activity stops despite the subject's preference to continue searching than from the other negative externality whereby the committee search activity continues despite the subject's preference to stop searching.

In summary, our experimental outcomes are consistent with the implications in the AAV model in terms of the relationships between committee and single-agent search for search duration and the probability of voting to stop. However, we do not obtain a significant experimental outcome in terms of the relationship of the probability of voting to stop committee search activities among given plurality voting rules. We also obtain other interesting findings in our comparison of the effects of Games B and C on the probability of voting to stop searching. The probit estimates show that many of the estimated coefficients on Games B are significantly positive relative to the reference group for Game A. This implies that subjects are not homogeneous with respect to their risk and loss attitudes, time preferences, and any unobserved factors, and that the heterogeneity of preferences among group members lowers each member's reservation value. Moreover, because the estimated coefficients on each Game C are significantly positive and larger in magnitude than the corresponding ones on Game B, additional heterogeneity exists in that the values the other members draw are different, which reinforces the incentive to vote to stop searching in an earlier round.

The chapter is organized as follows. Section 2 discusses the implications of the AAV model in detail. Section 3 explains the strategy for identification, and Sect. 4 elaborates upon the experimental design. Section 5 includes descriptive statistics and results of the regression analysis. Section 6 provides some concluding remarks.

2 Model

Albrecht et al. (2010) construct a committee search model in which group members decide whether to stop or continue searching by vote. They assume that group members are risk neutral and homogeneous with respect to preferences, and that each member randomly draws a value from an independent and identical distribution across members. We should note that values are therefore uncorrelated across group members. This model setup is a proxy describing the more realistic environment in which members draw the same value, but do not know how other members evaluate the value or their attitudes toward loss and risk.[3] This setting may differ from a realistic situation, but it is analytically tractable and qualitatively identical. Given the value in hand, each group member votes for or against stopping the search.

Albrecht et al. (2010) found that agents are less picky in committee search than in single-agent search because each member faces two negative externalities: (i) committee search continues under a given voting rule despite an individual's preference for the search to stop; and conversely (ii) committee search stops under a given voting rule despite the individual's preference to continue searching. These negative externalities are attributable to the assumption that each member draws a

[3]Compte and Jehiel (2010) consider the case where members hold the same value but do not know how other members evaluate this value.

value from the independent and identical distribution in the AAV model.[4] Thus, the reservation value is lower in committee search than in single-agent search, thereby leading to a shorter search duration (a higher probability of stopping the search). The AAV model refers to this as the *threshold effect*. However, there is another effect determining search duration; that is, committee search with plurality voting rules either raises or lowers the probability of stopping the search, given any reservation value. This is referred to as the *vote aggregation effect*. Whether the probability of stopping the search for any given reservation value is higher or lower in committee search than in single-agent search depends on the given reservation value in the single-agent search structure, the discount factor and the plurality voting rule applied.[5]

Figure 15.1 illustrates an example of the AAV model and decomposes the probability of stopping the search in the case of single-agent search versus committee search into the two effects, i.e., the threshold and the vote aggregation

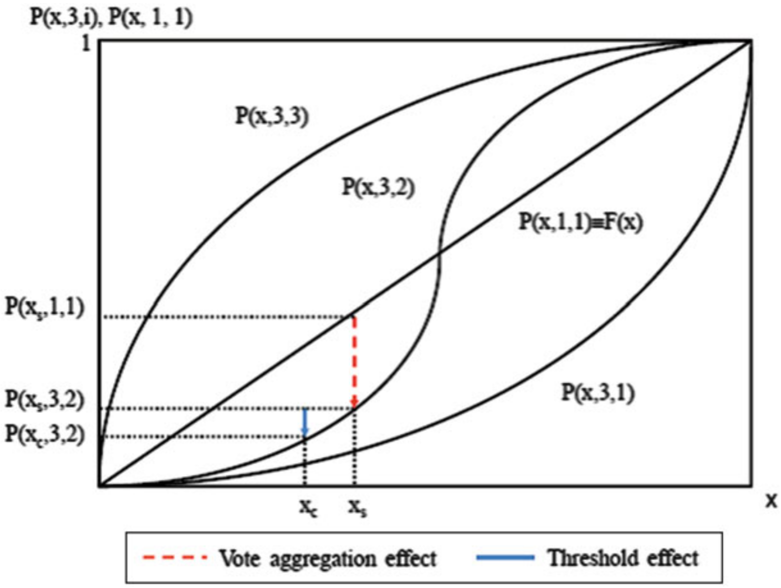


Fig. 15.1 The probability of continuing the search under each plurality voting rule
Note: [1-P(x, 1, 1)] and [1-P(x, 3, i)] represent the probability of stopping the search

[4]It is noted that these negative externalities arise as long as values drawn by group members are not perfectly correlated. The AAV model restricts to a case that draws are made across group members from an independent and identical distribution.

[5]The reservation value and the discount factor are closely related in the standard sequential search model. If an individual discounts the future more, then the individual's reservation value is lower, implying that he or she wants to exit the search earlier.

effects. We assume here that a group consists of three members (we design the group as comprising three members in our experiment). In addition, we assume that an individual member conducts a committee search activity with the other two members and faces a uniform distribution $F(x)$ of drawn value x with a lower bound of $\underline{x}$ and an upper bound of $\overline{x}$. The probability of continuing to search $P(x, 3, i)$ is calculated by the sum of the binomial probabilities that exactly $i - 1$ or fewer among the three members vote to stop, given that the reservation value is x. For example, $P(x, 3, 2)$ indicates the probability that none or only one of the three members votes to stop, in which case this group continues to search under the majority rule. Be aware that $P(x, 1, 1) \equiv F(x)$. $(1 - P(x, 3, i))$ instead represents the probability of stopping the search. In this illustration, the reservation value is lower in committee search with the majority rule (x_c) than in single-agent search (x_s), and the probability of stopping the search is higher in committee search with the majority rule $(1 - P(x_c, 3, 2))$ than in single-agent search $(1 - P(x_s, 1, 1))$. The difference in the probability of stopping the search between committee search with the majority rule and single-agent search is thus decomposed into the threshold effect $(P(x_s, 3, 2) - P(x_c, 3, 2))$ and the vote aggregation effect $(P(x_s, 1, 1) - P(x_s, 3, 2))$.

The vote aggregation effect results in a higher probability of stopping the search in committee search with the one-vote rule than in single-agent search for any given reservation value (see Fig. 15.2). Because the committee search activity stops if any of the three members votes to stop, the probability of stopping the search is higher. This reinforces the shorter search duration that occurs for this type of search. On the other hand, in the comparison of single-agent search versus committee search with

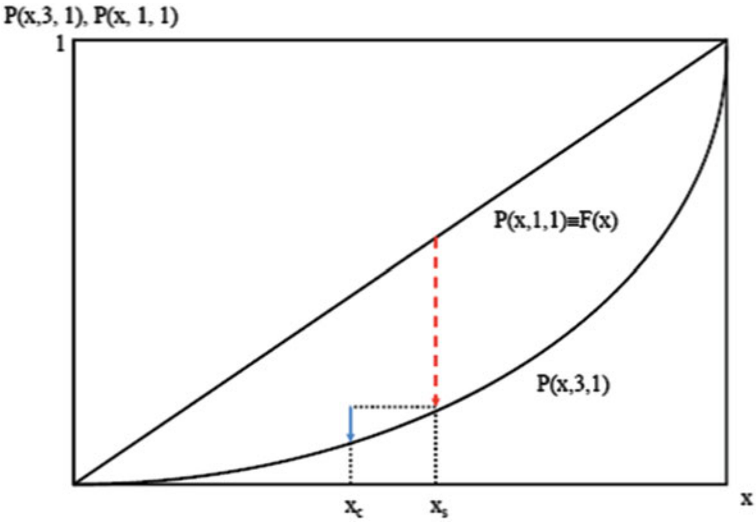


Fig. 15.2 Threshold and vote aggregation effects under the one-vote rule

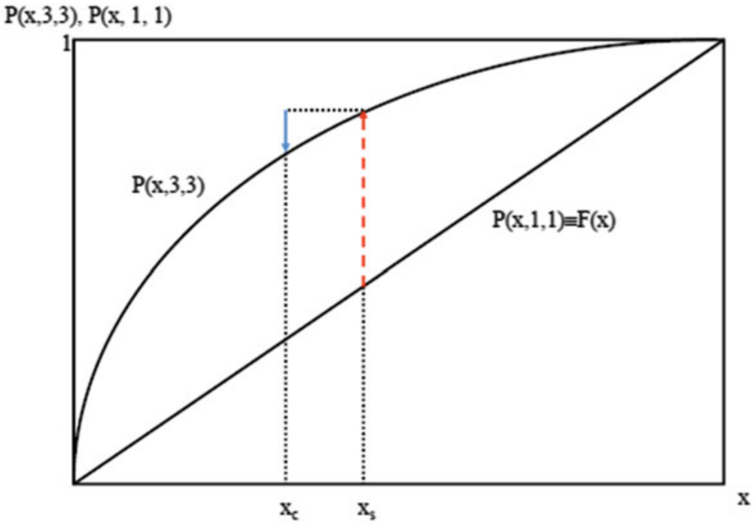


Fig. 15.3 Threshold and vote aggregation effects under the unanimity rule

the unanimity rule shown in Fig. 15.3, the vote aggregation effect induces a lower probability of stopping the search in the case of the committee search activity for any given reservation value because it takes longer for the three members to reach an agreement. Thus, this type of search leads to a longer search duration. Under the unanimity rule in this specific example, the vote aggregation effect is large enough to dominate the threshold effect, thereby resulting in the longer search duration. In a comparison of single-agent search versus committee search with the majority rule, the vote aggregation effect reinforces the threshold effect within the range of lower reservation values (or lower discount factors), implying a shorter search duration. The vote aggregation effect has the opposite effect to the threshold effect within the range of higher reservation values (or higher discount factors). For sufficiently high reservation values, the threshold effect dominates the vote aggregation effect, leading to a shorter search duration. However, for only moderately high reservation values, the magnitude relation between these two effects reverses, resulting in a longer search duration (see Fig. 15.4).

3 Strategy for Identification

This section considers a methodology to identify the implications of the AAV model. In our experiment, we restrict ourselves to the case where the group consists of three members and then compare the experimental outcomes under the various plurality voting rules. We have two main reasons for choosing three-member groups. The first is that this group size is sufficiently large to analyze the search behavior of

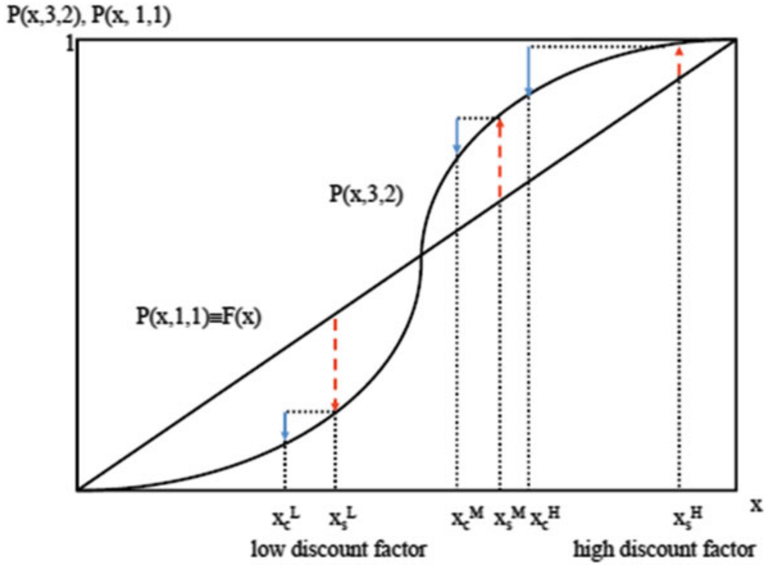


Fig. 15.4 Threshold and vote aggregation effects under the majority rule

individual agents in a committee search activity under the various plurality voting rules described below. The second reason is that this small group size allows us to obtain data from a large number of groups in our laboratory with limited capacity.

3.1 Comparison of the Single-Agent Search Model and the Committee Search Model

This subsection describes a way of testing and comparing the single-agent search model with a committee search model with a variety of plurality voting rules. Recall that in the AAV model, agents are homogeneous with respect to their risk and loss attitudes, time preferences, and any other unobserved characteristics, and that the difference between single-agent and committee search behavior therefore arises only from heterogeneity in the sense that the values drawn by group members are different and unknown to each other.[6] In our experiment, however, there exists yet another source of heterogeneity in the sense that grouped subjects differ in their risk and loss attitudes, time preferences, and any unobserved characteristics. To

[6]The properties of the AAV model remain the same, regardless of whether the different values drawn by the other members are known or unknown. In our experiments, subjects drew the value privately, and therefore no member knew the values that the other members drew.

eliminate the bias arising from the latter form of heterogeneity, we conduct the following three games.

- **Game A:** A single agent independently decides to stop or continue searching.
- **Game B:** A group draws a common value from a distribution (group members know that they hold the same value), and then the group members collectively decide by vote to stop or continue searching.
- **Game C:** Each member of a group draws a value from the independent and identical distribution, and then the group members collectively decide by vote to stop or continue searching.

Using these games, we can observe the search durations of each subject. The difference in the search duration between Games A and B is attributable to the heterogeneity among group members regarding their risk and loss attitudes, time preferences, and any unobserved characteristics. The difference in the search duration between Games A and C is attributable to two types of heterogeneity: heterogeneity in terms of what other members' preferences are, as detailed above, and in the different values that group members draw from the independent and identical distribution. Therefore, the difference-in-differences of the search duration stems only from the heterogeneity in terms of the different values that the group members draw, implying that the heterogeneity in terms of the differences among members regarding their risk and loss attitudes, time preferences, and any unobserved characteristics is eliminated. This method therefore picks up the exact difference in the search duration between single-agent search versus committee search as characterized by the AAV model.

3.2 Comparison of the Different Plurality Voting Rules

The AAV model shows that the probability of stopping the committee search activity varies according to the plurality voting rule applied. To identify the effects of the different plurality voting rules, we conduct the following three subgames.

- **Subgame 1:** The committee search activity is stopped if at least one of the group members votes to stop (the one-vote rule).
- **Subgame 2:** The committee search activity is stopped only if two-thirds or more of the group members vote to stop (the majority rule).
- **Subgame 3:** The committee search activity is stopped only if all the group members vote to stop (the unanimity rule).

In Subgame 1, members are faced with the risk that committee search will stop despite an individual member's preference to continue searching, whereas in Subgame 3, committee search continues even if an individual member wants to stop searching. Subgame 2 is the in-between case of Subgames 1 and 3. Using these subgames, we can identify the effects of the risks presented by the different plurality voting rules.

4 Experimental Design

We conducted our experiment on February 2, 2010, in the experimental laboratory of the Center for Experimental Research in Social Sciences at Hokkaido University, Japan. The experiment consisted of four separate sessions because of the constraints in laboratory capacity. Each session involved the same eight games in a different order (two games of Game A and three subgames each of Game B and Game C). We designed various experiments of an infinite-horizon sequential search model with a 5 % probability that the search coercively ends.[7] When the committee search activity is coercively terminated, the subject unconditionally obtains the value drawn in the previous round. Recall was not allowed in this search model, meaning that the value drawn in the previous round is not available for consideration in the current round, except for the case of coercive termination. Although subjects are not encouraged to search longer in a search environment where recall is not allowed, this design is simple and exactly duplicates the structure of the AAV model.

The games differ in terms of the treatments in each experimental session. The experimental processes with the different treatments for single-agent and committee search are set out below.

4.1 Game A: Single-Agent Search

Game A is the benchmark for a normal single-agent search task. A subject makes a draw from a uniform distribution with a lower bound of zero and an upper bound of 3,000. After making the draw, the subject decides whether to stop or continue searching. If the subject chooses to continue, he or she moves to the next round and makes another draw from the same distribution.

In an infinite-horizon sequential search model, the value of searching for a single agent V_A is given by:

$$V_A = \underbrace{0.95 \frac{R_A}{3{,}000} V_A}_{\text{continuing search}} + \underbrace{0.05 \int_0^{R_A} \frac{u(x)}{3{,}000} dx}_{\text{terminating search}} + \underbrace{\int_{R_A}^{3{,}000} \frac{u(x)}{3{,}000} dx}_{\text{accepting the offer}}, \tag{15.1}$$

where $u(x)$ is the utility function,[8] and R_A is the reservation value. For simplicity, there is no explicit discount over rounds, which encourages subjects to search longer.

[7] We also ran finite-horizon versions of the same experiments at Osaka University, Japan. Specifically, participants were told that the experiments would end after 20 rounds. The results from the finite-horizon experiments are qualitatively similar to those from the infinite-horizon experiments.

[8] The AAV model assumes that agents are risk-neutral.

Our experiments are designed in such a way that a subject's search activity has to be terminated coercively with a probability of 5 % for each round. This probability partially fulfills the role of a search cost that the subject incurs by continuing to search in the next round. The subject then has an incentive to stop searching even in the infinite-horizon sequential search model with no discount over rounds. The sum of the first two terms on the right-hand side in Eq. (15.1) represents the value of rejecting a drawn value. The first of these terms is the value of continuing the search after the subject survives to the next round, and the second of these terms indicates the value of the search activity being terminated after the offer is rejected. When a search activity is coercively terminated, the subject has no choice but to accept the value that she or he rejected in the previous round.[9] The final term represents the value of accepting a drawn value. Because the reservation value of gaining R_A is equivalent to the value of searching in the next round ($R_A = 0.05R_A + 0.95V_A$), we obtain $R_A = V_A$.

4.2 Game B: Committee Search with a Common Value

Game B involves a form of committee search in which all members of a group draw a common value from the uniform distribution with a lower bound of zero and an upper bound of 3,000, and they know that they have the same value. Whether the committee search stops or continues is then decided by vote among group members under various plurality voting rules. Because group members are randomly reshuffled in every game, no member knows who the other two group members are, which rules out the presence of a learning effect regarding other group members' voting behavior.

If group members are homogeneous with respect to preferences, as in the AAV model, the committee search model with a common value is reduced to the single-agent search model shown in Game A. Therefore, the value for a group member of searching by committee V_B is:

$$V_B = \underbrace{0.95 \frac{R_B}{3{,}000} V_B}_{\text{continuing search}} + \underbrace{0.05 \int_0^{R_B} \frac{u(x)}{3{,}000} dx}_{\text{terminating search}} + \underbrace{\int_{R_B}^{3{,}000} \frac{u(x)}{3{,}000} dx}_{\text{accepting}} . \tag{15.2}$$

[9] The experiment is designed to reduce the loss that each subject would have incurred if the search was terminated after she or he had decided to continue to search. This encourages the subject to search longer. In other words, in a design where each subject receives no payment when the search activity is terminated coercively, it is expected that she or he will not engage in the search activity for many rounds.

Because $R_B = 0.05R_B + 0.95V_B$, we obtain $R_B = V_B$. According to Eq. (15.2), if all group members are homogeneous with respect to preferences, they all either accept or reject a drawn value together, regardless of which voting rule is employed.

If there is a difference in the reservation values between Games A and B in our experiment, it is largely attributable to the heterogeneity of preferences in search activity among members. In our experiment, Game B consists of three subgames, each of which differs according to the plurality voting rules explained in Sect. 3.2.

4.3 Game C: Committee Search with Different Values

In Game C, similarly to Game B, group members decide by vote whether to stop or continue searching. However, unlike Game B, each group member separately makes a draw from a uniform distribution with a lower bound of zero and an upper bound of 3,000, which means that the drawn values are identically independent across the group members. Game C, like Game B, consists of three subgames, C-1, C-2, and C-3, as explained in Sect. 3.2.

Suppose that committee search is stopped if at least one group member votes to stop searching (Subgame C-1: one-vote rule). The value for a subject of searching by committee V_{C1} is given by:

$$
\begin{aligned}
V_{C1} = &\underbrace{0.95\left(\frac{R_{C1}}{3{,}000}\right)^3 V_{C1}}_{\text{continuing committee search after the subject votes to continue}} \\
+ &\underbrace{0.05\left(\frac{R_{C1}}{3{,}000}\right)^2 \int_0^{R_{C1}} \frac{u(x)}{3{,}000}dx}_{\text{terminating committee search after the subject votes to continue}} \\
+ &\underbrace{\int_{R_{C1}}^{3{,}000} \frac{u(x)}{3{,}000}dx}_{\text{accepted by self}} + \underbrace{\left[1-\left(\frac{R_{C1}}{3{,}000}\right)^2\right]\int_0^{R_{C1}} \frac{u(x)}{3{,}000}dx}_{\text{accepted by one or both of the others, but not by self}}. \qquad (15.3)
\end{aligned}
$$

The sum of the first two terms on the right-hand side in Eq. (15.3) represents the value of committee search when all group members vote against stopping the search. The first of these two terms denotes that the committee group survives to the next round with a probability of 0.95, but the second term shows that the group has to stop searching, and thus the group obtains the value drawn in the previous round. The sum of the third and fourth terms in Eq. (15.3) indicates the value for the subject of accepting the drawn value. The third term represents that at least one member, including the subject, votes to stop searching, and the fourth term shows that one or both of the other two members vote for stopping the search, although the subject votes against it. Because $R_{C1} = 0.05R_{C1} + 0.95V_{C1}$ according to the reservation value rule, we obtain $R_{C1} = V_{C1}$.

Next, we consider a committee search model in which committee search is stopped only if at least two out of three members vote to stop searching (Subgame C-2: majority rule). The value for a subject of searching by committee V_{C2} is obtained as follows:

$$
\begin{aligned}
V_{C2} = {} & \underbrace{0.95\left[\left(\frac{R_{C2}}{3{,}000}\right)^3 + 3\left(1-\frac{R_{C2}}{3{,}000}\right)\left(\frac{R_{C2}}{3{,}000}\right)^2\right]V_{C2}}_{\text{continuing committee search}} \\
& + \underbrace{0.05\left[\left(\frac{R_{C2}}{3{,}000}\right)^2 + 2\left(1-\frac{R_{C2}}{3{,}000}\right)\frac{R_{C2}}{3{,}000}\right]\int_0^{R_{C2}}\frac{u(x)}{3{,}000}dx}_{\text{terminating committee search after the subject votes to continue}} \\
& + \underbrace{0.05\left(\frac{R_{C2}}{3{,}000}\right)^2\int_{R_{C2}}^{3{,}000}\frac{u(x)}{3{,}000}dx}_{\text{terminating committee search after the subject votes to stop}} + \underbrace{\left(1-\frac{R_{C2}}{3{,}000}\right)^2\int_0^{R_{C2}}\frac{u(x)}{3{,}000}dx}_{\text{accepted by others, but not by self}} \\
& + \underbrace{\left[1-\left(\frac{R_{C2}}{3{,}000}\right)^2\right]\int_{R_{C2}}^{3{,}000}\frac{u(x)}{3{,}000}dx}_{\text{accepted by self and one or both other members}}.
\end{aligned}
\tag{15.4}
$$

The first term on the right-hand side in Eq. (15.4) shows the value of continuing the committee search activity after the committee survives to the next round. $\left[(R_{C2}/3{,}000)^3 + 3(1-(R_{C2}/3{,}000))(R_{C2}/3{,}000)^2\right]$ implies the probability that at least two members vote against stopping the search. The second and third terms represent the value of the committee search activity being coercively terminated. The second term shows the case in which at least two members, including the subject, vote against stopping the search before termination of the committee search activity, whereas the third term deals with the case in which the subject votes for stopping the search, but the other two members vote against it before termination. The fourth and fifth terms indicate the value for the subject of stopping the committee search activity; the fourth term shows the case in which the subject votes against stopping the search, but the other two members vote for it, whereas the fifth term indicates that the subject and one or both of the other two members vote to stop searching. As discussed above, we have $R_{C2} = V_{C2}$.

Finally, we move to the committee search model in which committee search is stopped only if all members of the group vote to stop (Subgame C-3: unanimity rule). The value for a subject of searching by committee V_{C3} is given by:

$$V_{C3} = \underbrace{0.95\left[1-\left(1-\frac{R_{C3}}{3{,}000}\right)^3\right]V_{C3}}_{\text{continuing committee search}} + \underbrace{0.05\int_0^{R_{C3}}\frac{u(x)}{3{,}000}dF(x)}_{\text{terminating committee search after the subject votes to continue}}$$

$$+\underbrace{0.05\left[\left(\frac{R_{C3}}{3{,}000}\right)^2 + 2\left(1-\frac{R_{C3}}{3{,}000}\right)\frac{R_{C3}}{3{,}000}\right]\int_{R_{C3}}^{3{,}000}\frac{u(x)}{3{,}000}dx}_{\text{terminating committee search after the subject votes to stop}}$$

$$+\underbrace{\left(1-\frac{R_{C3}}{3{,}000}\right)^2\int_{R_{C3}}^{3{,}000}\frac{u(x)}{3{,}000}dx}_{\text{accepted by all members}}. \tag{15.5}$$

The first term on the right-hand side in Eq. (15.5) represents the value for the subject of continuing the committee search after her or his committee survives to the next round, whereas the last term denotes the value for the subject of accepting a drawn value when all of the members vote to stop the search unanimously. The second and third terms represent the value of the committee search activity being coercively terminated. The second term indicates the value of the committee search activity being terminated after at least the subject votes against stopping the search. In the third term, committee search is terminated after at least the subject votes for stopping the search, but one or both of the other members vote against it.

4.4 Hypotheses

This subsection sets out our experimental hypotheses developed to test the theoretical implications of the AAV model that assumes that agents are risk-neutral, that is, $u(x) = x$. Because Albrecht et al. (2010) found many novel and interesting implications from their committee search model, we cannot test all possible hypotheses arising from it. Accordingly, we focus the laboratory experiment on testing two sets of hypotheses relating to: (i) the comparison of search duration between single-agent and committee search under various plurality voting rules; and (ii) the comparison of the willingness to stop searching between single-agent and committee search under various plurality voting rules. For the first set of hypotheses (regarding search duration), we test the theoretical implications of the first part of Proposition 5 and the second part of Proposition 4 in Albrecht et al. (2010). For the second set of hypotheses (regarding the willingness to stop searching), we test the implications of Proposition 2 and the second part of Proposition 5.

We begin with the first set of hypotheses.

- **H1 (the first part of Proposition 5):** The average search duration is increasing in the number of votes required to stop committee search. In addition, the average search duration is shorter (longer) in the case of committee search with the one-vote rule (unanimity rule) than in the case of single-agent search.

If this hypothesis is statistically supported, as indicated in the AAV model, then, compared with single-agent search, the probability of stopping the search is higher in committee search with the one-vote rule but lower in committee search with the unanimity rule. The probability of stopping the search in committee search with the majority rule lies between the other plurality voting rules.

- **H2 (the second part of Proposition 4):** The average search duration is increasing in the number of group members, given that the number of votes required to stop searching equals the number of group members.

To test this hypothesis, we compare the average search duration between committee search with the unanimity rule where the search is stopped if all three members of the group vote to stop and single-agent search where the individual search stops if "one out of the one member of the group" decides to stop the search. Our expectation is that the search duration is longer in the case of committee search with the unanimity rule than in the case of single-agent search.

(H1) and (H2) address the combined effects (the threshold and vote aggregation effects) on the average duration of committee search. The next step is to employ the probit model to estimate determinants of individual voting behavior regarding stopping searching, using data from each round-based decision from the eight games. This identifies the threshold effect whereby subjects are less picky in committee search than in single-agent search, as described in Proposition 2 in Albrecht et al. (2010). We then move to the second set of hypotheses as follows.

- **H3 (Proposition 2):** A subject votes for stopping the search in a stochastically earlier round of committee search, regardless of the plurality voting rules, compared with single-agent search.

This test allows us to capture the threshold effect and compares the reservation values between single-agent and committee search using the various plurality voting rules. The final hypothesis we test in our experiment is the implication from the second part of Proposition 5.

- **H4 (the second part of Proposition 5):** The reservation value initially increases in the number of votes required to stop the search. Once this number reaches a sufficiently large value, the reservation value may decrease thereafter. In other words, in committee search, the reservation value is lower under the one-vote rule than under the majority rule, but the reservation value is either higher or lower under the unanimity rule than under the majority rule.

In other words, H4 implies that the probability of voting to stop the committee search activity can be expressed as either everywhere decreasing or U-shaped in the number of votes required to stop. We test the null hypothesis that $R_{C1} \geq R_{C2}$ and $R_{C2} \geq R_{C3}$, or $R_{C1} \geq R_{C2}$ and $R_{C2} \leq R_{C3}$, where as mentioned before, R_i represents the reservation value for game i. This thus allows us to reduce to test the null hypothesis that $R_{C1} \geq R_{C2}$ or the null hypothesis that the probability of voting to stop the community search activity is equal to or lower under the on-vote rule than under the majority rule. We expect that this null hypothesis is significantly rejected.

In addition, we test whether group members are on average homogeneous with respect to their risk and loss attitudes, time preferences, and any unobserved characteristics. If the estimated coefficients on the dummy variables from Game B are not jointly different from zero in the probit estimation of individual voting behavior, we can support the hypothesis that group members are homogeneous with respect to their preferences. This draws our attention to eliminating the bias arising from the heterogeneity of preferences across group members when the four hypotheses are tested. Section 5 provides the estimated results.

4.5 Administration and Payoffs

We conducted four sessions. The order of games in each session was as follows: first session: Games A, B-3, B-2, B-1, C-3, C-2, C-1, and A; second session: Games A, C-1, C-2, C-3, B-3, B-2, B-1, and A; third session: Games A, C-3, C-2, C-1, B-1, B-2, B-3, and A; and fourth session: Games A, B-1, B-2, B-3, C-1, C-2, C-3, and A. We instructed the subjects to play a training game that was the same as Game A once before the experiment began. Although we might not be able to rule out the possibility that this training affected the subjects' behavior in the first Game A, this step was necessary to ensure that the subjects correctly understood the nature of the experiment. We ran Game A twice as the first and last games in each session. This enables us to determine whether an anchoring effect arises in the sense that there is any difference in the search behavior between the first and last Game A. The anchoring effect implies that the subjects' behavior is affected by results that they obtained in previous games. The subjects were 60 undergraduate students from various academic disciplines. We ran the experiments entirely on computers using the software package *Z-Tree* (Fischbacher 2007).[10]

The instruction sheet presented full information about the search task.[11] Following the experiment, the participants answered a questionnaire and the payoff procedures took place. With regard to payoffs, we emphasized that: (i) the subjects' payoff was truncated at JPY0 (EUR0) (i.e., they could not incur losses from the search task) and (ii) they would earn an appearance fee of JPY1,000 (EUR7.9).[12] The performance pay was determined based on the result from one of the eight

[10]The programs were produced by Takanori Kudou, a graduate student of the Engineering Division of Electrical, Electronic and Information Engineering, Osaka University.

[11]The instruction sheet is reproduced in the appendix.

[12]We use the exchange rate of JPY100 to EUR0.79 for February 2, 2010, the date when the experiment was conducted.

games randomly chosen by each subject. The expected total payoff was JPY2,500 (EUR19.75) to 3,000 (EUR23.7). Therefore, because the on-duty time for the experiment was approximately 90 min, the hourly wage was calculated at JPY1,600 (EUR12.64) to 2,000 (EUR15.8). This is approximately twice as much as the average hourly wage for college students in Japan, implying that we set the appropriate way of payoff to encourage subjects to work hard.

5 Results

5.1 Search Duration

We begin with some descriptive statistics before undertaking statistical hypothesis tests of the first set of hypotheses (H1) and (H2). Table 15.1 provides selected descriptive statistics, including the averages and standard deviations of the durations for the infinite-horizon sequential search model. We find that the average search duration for single-agent search in the first Game A differs from that in the last Game A at the 5 % level of significance, implying that there may be an anchoring effect in the sense that a subject's search behavior is influenced by results she or he obtained in the previous experimental games. This may result in an identification bias in the statistical tests. The average search duration is longer for committee search under the majority and unanimity rules when group members have different values (Game C) compared with when they have the same value (Game B), but the reverse is observed when the one-vote rule applies. As the required number of votes to stop searching increases, the average search duration becomes longer, regardless of whether group members draw the same or different values.

Table 15.2 provides the results of t-tests of the difference in search duration between committee search under the different plurality voting rules versus single-agent search. The top three rows of this table compare the average duration for

Table 15.1 Average search durations

	Value	Voting rule	#Searchers	Sample	Mean	S.D.	Max	Min
Game A(1)			1	60	3.233	3.306	18	1
Game B-1	Common	One vote	3	60	2.650	1.571	6	1
Game B-2	Common	Majority	3	60	2.750	1.684	6	1
Game B-3	Common	Unanimity	3	60	3.600	2.981	13	1
Game C-1	Different	One vote	3	60	1.300	0.561	3	1
Game C-2	Different	Majority	3	60	2.850	1.921	7	1
Game C-3	Different	Unanimity	3	60	11.700	8.947	32	1
Game A(8)			1	60	4.633	4.422	21	1
Average				60	4.090	4.992	32	1

Game A(1) and Game A(8) represent the first and last trial of Game A, respectively

Table 15.2 Comparisons with single-agent search (first Game A)

Game B-1	Common	One vote	The search duration is shorter at the 10 % level of significance
Game B-2	Common	Majority	The search duration is insignificantly shorter
Game B-3	Common	Unanimity	The search duration is insignificantly longer
Game C-1	Different	One vote	The search duration is shorter at the 1 % level of significance
Game C-2	Different	Majority	The search duration is insignificantly shorter
Game C-3	Different	Unanimity	The search duration is longer at the 1 % level of significance

Note: We compared the means of the search durations using the one-tailed t-test

single-agent search and committee search in which all group members draw the same value (Game B). We find that the null hypothesis that the average duration of committee search with the one-vote rule is equal to or longer than that of single-agent search is rejected at the 10 % level of significance. Conversely, compared with single-agent search, the average duration is shorter in committee search with the majority rule and longer in committee search with the unanimity rule, but in both these cases, the difference in the average duration between the two models is not significantly different from zero.

According to the AAV model, group members are homogeneous with respect to their risk and loss attitudes, time preferences, and any unobserved characteristics, and therefore share the same threshold. Therefore, when the group members make the same draw and evaluate it in common, neither the threshold nor the vote aggregation effect arises, resulting in no difference in the average duration between the single-agent and committee search models, regardless of which plurality voting rule applies.

Although the differences in the average search duration between single-agent and committee search with the majority and unanimity rules in Game B are statistically insignificant, this does not necessarily mean that group members are homogeneous with respect to their preferences. If group members are heterogeneous in terms of their preferences, they evaluate the same drawn value differently, in a case of which both threshold and vote aggregation effects arise. On the one hand, it takes more time to reach an agreement in committee search with the majority and unanimity rules than in single-agent search, which lengthens the search duration. On the other hand, each subject's reservation value decreases because she or he is less picky, thereby shortening the search duration. It is possible that the two opposing effects cancel each other out, which suggests that group members might have been heterogeneous with respect to their preferences. To correctly test the implications of the AAV model where members are homogeneous, it is necessary to test the difference in the threshold level among group members, which is analyzed in Sect. 5.2.

The bottom three rows in Table 15.2 compare the average duration between single-agent and committee search in which group members draw different values (Game C). Recall that the difference in the average duration between the two models arises from the heterogeneity of preferences among group members, as discussed above, and from the other heterogeneity in terms of the different draws that other members make. The search duration is shorter in committee search with the one-vote rule than in single-agent search at the 1 % level of significance. Similarly, the search duration is longer in committee search with the unanimity rule than in single-agent search at the 1 % level of significance. The average search duration is insignificantly shorter in committee search with the majority rule than in single-agent search. We have not yet identified whether the difference in average duration arises from the heterogeneity among group members with respect to their preferences or from the other heterogeneity in the sense that the draws other members make are different. We can correctly confirm the implications of the AAV model by correspondingly deducting the differences displayed in the top three rows in Table 15.2 from the differences displayed in the bottom three rows. We test the null hypothesis that the *difference-in-differences* of the average search duration, (Game C-1 – Game A) – (Game B-1 – Game A) = Game C-1 – Game B-1, is zero. The same procedure applies to the other two voting rules.

Table 15.3 gives the results of t-tests of the difference in the average duration between single-agent and committee search after controlling for heterogeneity among group members. As shown in the first row of the table, the average search duration is shorter in committee search with the one-vote rule than in single-agent search at the 1 % level of significance. Similarly, the null hypothesis that the average search duration in committee search with the unanimity rule is equal to or shorter than the average duration in single-agent search is rejected at the 1 % level of significance. Another interpretation of this result is that as the number of group members increases from one to three, holding the unanimity rule fixed, the average search duration becomes longer. The single-agent search model is then considered a special case of the committee search model with the unanimity rule. This result is consistent with the second part of Proposition 4 in the AAV model. Looking at the differences between single-agent and committee search with the majority rule in Table 15.3, we cannot significantly reject the null hypothesis that the two average durations are equal. The AAV model predicts that if the reservation value of single-agent search (or the discount factor) is extremely low or extremely high, the average search duration is shorter in committee search with the majority rule, but otherwise it is longer, as illustrated in Fig. 15.4. Therefore, it is not surprising that

Table 15.3 Comparisons of search durations between single-agent search and committee search

First Game A vs Game C-1	The search duration is shorter in Game C-1 at the 1 % level of significance
First Game A vs Game C-2	The search duration is insignificantly longer in Game C-2
First Game A vs Game C-3	The search duration is longer in Game C-3 at the 1 % level of significance

Note: To control for heterogeneity of preferences among group members, we tested the implications of the AAV model by deducting a difference of the search duration between first Game A and Game C from a difference of the search duration between first Game A and Game B. In fact, we compared the means of the search durations between Games B and C using the one-tailed t-test

Table 15.4 Estimates of search durations (OLS)

Search duration			
	[1]	[2]	[3]
Game B-1	−1.1383	− 1.1383*	−1.1383*
	(0.6916)	(0.6564)	(0.6607)
Game B-2	−1.0383	− 1.0383	−1.0383
	(0.6682)	(0.6491)	(0.6521)
Game B-3	−0.1883	− 0.1883	−0.1883
	(0.7324)	(0.7043)	(0.7051)
Game C-1	−2.4883***	− 2.4883***	−2.4883***
	(0.6502)	(0.6118)	(0.6162)
Game C-2	−0.9383	− 0.9383	−0.9383
	(0.6790)	(0.6500)	(0.6543)
Game C-3	7.9117***	7.9117***	7.9117***
	(1.1441)	(1.2278)	(1.2293)
Game A(8)	0.2900	0.2900	0.2900
	(1.0881)	(1.1284)	(1.1336)
Game order	0.1586	0.1586	0.1586
	(0.1242)	(0.1254)	(0.1258)
Risk aversion		−353.1575	
		(503.46)	
Female			0.0282
			(0.3641)
Constant	1.6102	3.169***	3.0311***
	(1.1131)	(0.4777)	(0.6668)
Individual effect	Yes	No	No
N	480	480	480
F-test	4.2342	26.6111	26.5625
R2	0.4781	0.3668	0.3660

Robust standard errors are in parentheses. ***1 % significance, **5 % significance, *10 % significance. The dependent variable is each subject's search durations. Game B-1 (common value + one-vote rule), Game B-2 (common value + majority rule), Game B-3 (common value + unanimity rule), Game C-1 (different values + one-vote rule), Game C-2 (different values + majority rule), Game C-3 (different values + unanimity rule). Game A(8) is the single-agent search game that subjects played in the last trial, which captures the anchoring effect, compared with the first trial of Game A (A(1)). We calculate the absolute risk aversion index using the willing -to-pay price for a lottery with a 25 % chance of winning JPY2,000 (EUR15.8) but a 75 % chance of receiving JPY0 (EUR0)

these outcomes are obtained from our experiment. An examination of the differences in committee search with the various plurality voting rules shows that the average search duration is increasing in the number of votes required to stop the search. This outcome supports the first part of Proposition 5 in the AAV model.

Table 15.4 provides the ordinary least squares estimates of the determinants of search duration in the infinite-horizon sequential search. The dependent variable is each subject's search duration in each game, while the vector of independent variables consists of dummy variables for treatment, game order, each subject's attitude toward risk, a dummy variable for female gender, and/or dummy variables for individual effects. The variable regarding the game order indicates the order in which the subject played the games in each session. The variable for risk attitude is included to partially control for individual heterogeneity. To collect this variable, we administered a questionnaire to all participants after eight games and asked three questions about their attitude toward risk. Of these, we selected one question relating to the price subjects were willing to pay for a lottery with a 25 % chance of winning JPY2,000 (EUR15.8) and a 75 % chance of receiving JPY0 (EUR0).[13,14] We then calculated the index measuring the extent of absolute risk aversion using the method in Cramer et al. (2002).[15] If this index is positive, a subject is considered risk averse; if negative, the subject is considered risk seeking. If the index is exactly zero, the subject is risk neutral.

We control for individual effects in column [1] of Table 15.4, and replace the individual effects with other individual characteristics represented by absolute risk aversion and gender in columns [2] and [3], respectively. The columns in Table 15.4 indicate almost the same results: the coefficient on Game C-1 is negative, whereas that on Game C-3 is positive, both at the 1 % level of significance when the reference group is defined as the first Game A.[16] The coefficient on Game B-1 is negative at a marginal level of significance (10 %) in columns [2] and [3]. From the estimates of search duration, we find that in committee search with the one-vote rule, group members are heterogeneous with respect to their risk and loss attitudes, time preferences, and any other unobserved characteristics. However, we should note that, as mentioned before, both threshold and vote aggregation effects arise if group members evaluate the common drawn value differently. It is then possible that

---

[13]We again use the exchange rate of JPY100 to EUR0.79 on February 2, 2010, corresponding to when the experiment was conducted.

[14]The other two questions were: "With at least what chance of rain do you take an umbrella?" and "What price are you willing to pay for a lottery with a 25 % chance of winning JPY200 (EUR1.6) but a 75 % chance of receiving JPY0 (EUR0)?" In the first question, subjects that responded with a lower value were considered to be more risk averse. We also estimated the determinants of search duration using indices of absolute risk aversion obtained from these questions and obtained similar results.

[15]According to Cramer et al. (2002), the extent of absolute risk aversion is calculated as follows:

$$\frac{0.25 \times 2{,}000 - price}{0.5(0.25 \times 2{,}000^2 - 2 \times 0.25 \times 2{,}000 + price^2)},$$

where *price* implies the price that a subject is willing to pay for the lottery with a 25 % chance of winning JPY2,000 (EUR15.8) but a 75 % chance of receiving JPY0 (EUR0).

[16]Note that the coefficients on games are the same in all columns. This is because, in the estimates of search duration where there are eight observations for search duration for each subject, individual characteristics are perfectly uncorrelated with the treatment variables (games) that are given exogenously.

Table 15.5 Tests for search durations

	Null hypotheses	Difference in the coefficients	Std. err.	t value	p-value	Test
[1]	Coef(C-1)-Coef(B-1)≥0	−1.3500	0.3691	−3.6600	0.0000	Rejected
[2]	Coef(C-2)-Coef(B-2)=0	0.1000	0.4144	0.2400	0.8090	Not rejected
[3]	Coef(C-3)-Coef(B-3)≤0	8.1000	1.1312	7.1600	0.0000	Rejected

Note: Each row represents differences in the search duration between Games B and Games C by each plurality voting rule, using the results from column [1] of Table 15.4, when the reference group is defined as the first Game A. The magnitude of the coefficients on Game B is attributable to the heterogeneity among group members in terms of their preferences. The magnitude of the coefficients on Game C is attributable to the heterogeneity of preferences among group members plus the heterogeneity in terms of what value the other group members draw. Therefore, the differences between the coefficients between Game B and C indicate the marginal effects derived only by the second heterogeneity in terms of what value the other members draw, after controlling for the first heterogeneity of preferences among group members

the threshold and vote aggregation effects are opposite and cancel each other out in Games B-2 and B-3. In this case, the estimated coefficients on these games become insignificant despite the heterogeneity of preferences among group members. We cannot exactly identify whether or not group members are homogeneous in terms of search behavior based on the estimates of search duration in Table 15.4.

Assuming that group members are heterogeneous in terms of their preferences, we test the first set of hypotheses regarding search duration, (H1) and (H2), by deducting the coefficients on Game B from the corresponding values for Game C, using the estimated results from column [1] of Table 15.4. We obtain the same results as Table 15.3, as shown in Table 15.5. The null hypothesis that the search duration in committee search with the one-vote rule is equal to or longer than the search duration in single-agent search is rejected at the 1 % level of significance. Similarly, we reject the null hypothesis that the average search duration in committee search with the unanimity rule is equal to or shorter than the average duration in single-agent search at the 1 % level of significance. This implies that the search duration is increasing in group size, holding the unanimity rule fixed. This result again supports the second part of Proposition 4 in the AAV model (H2). We cannot reject the null hypothesis of no significant difference in the search duration between committee search with the majority rule and single-agent search. As discussed, this result is not surprising because the AAV model shows that the search duration is either shorter or longer in committee search with the majority rule, depending mainly on the reservation value of the single-agent search model and the agent's discount factor. We confirm that the search duration is increasing in the number of votes required to stop committee search. This result is again consistent with the first part of Proposition 5 in the AAV model (H1).

Consider now the estimated coefficients on the other independent variables. The variable regarding the game order is statistically insignificant in all columns of Table 15.4. In the estimates of the search duration, we can say that the anchoring effect whereby a subject's behavior is affected by how she or he behaved in the previous games is minor. We estimate the effects on the search duration of individual characteristics represented by absolute risk aversion and gender. In columns [2] and [3] of Table 15.4, we find that the coefficients on these variables are insignificant.

5.2 The Probability of Voting to Stop

This subsection considers each round-based decision by subjects on whether to vote for stopping the search and then tests the second set of hypotheses regarding the probability of voting to stop the search, (H3) and (H4). Table 15.6 displays the average marginal effects of the probit model to estimate determinants of the vote to stop searching, using the round-based data on search decisions. The dependent variable is dichotomous, taking a value of one if a subject accepts a drawn value in the case of single-agent search or votes to stop searching in the case of committee search, and zero otherwise. The independent variables are dummy variables for treatment, the drawn value, the round, game order, attitude toward risk, a dummy variable for female gender, and/or dummy variables for individual effects. The purpose of these estimations is to capture differences in the probability of voting to stop the search, or in the reservation value between committee search and single-agent search, and then to extract the threshold effect quantitatively.

We control for individual effects in columns [1] and [2] of Table 15.6. Column [1] excludes the independent variable for round whereas column [2] includes this variable. We replace the individual effects with other individual characteristics represented by absolute risk aversion and gender in columns [3] and [4], respectively. In all columns, the coefficients on Game B-2, Game B-3, and all Games C are positive at the 1 % level of significance, compared with the reference group of the first Game A, whereas the coefficients on Game B-1 are significantly positive at the 5–10 % level. Because the coefficients on Games B are significantly positive and different from zero, we can say that heterogeneity exists in terms of preferences among group members, which encourages subjects to vote in favor of stopping the search in a stochastically earlier round. Because the coefficients on committee search are larger when the group members draw different values (Games C) than when the group members draw the same value (Games B), regardless of the plurality voting rule, the remaining heterogeneity regarding the different values that other members draw from the distribution reinforces the incentive to vote to stop searching in a stochastically earlier round. These results confirm the threshold effect, which suggests that the reservation value is lower in committee search than in single-agent search in the AAV model.

Table 15.7 provides the results of the tests of hypothesis (H3) that subjects are less picky in committee search, regardless of which plurality voting rule is employed, than in single-agent search. Each coefficient on Games C represents the average marginal probability of voting to stop in Games C, compared with that in the first Game A, using the estimated results in column [2] of Table 15.6. We should note that, as before, we test (H3) by deducting the coefficients on Games B from

Table 15.6 Average marginal effects on the probabilities of voting to stop the search (probit estimations)

Willing to accept=1	[1] Coef.	[2] Coef.	[3] Coef.	[4] Coef.
Drawn value	0.0003***	0.0003***	0.0003***	0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Round		0.0039***	−0.0016	−0.0017
		(0.0010)	(0.0012)	(0.0012)
Game B-1	0.0460*	0.0566**	0.0664**	0.0798**
	(0.0265)	(0.0271)	(0.0320)	(0.0326)
Game B-2	0.0859***	0.0949***	0.0981***	0.1117***
	(0.0239)	(0.0243)	(0.0288)	(0.0293)
Game B-3	0.1098***	0.1170***	0.1174***	0.1335***
	(0.0343)	(0.0348)	(0.0338)	(0.0332)
Game C-1	0.1153***	0.1311***	0.1360***	0.1500***
	(0.0319)	(0.0318)	(0.0434)	(0.0451)
Game C-2	0.1496***	0.1577***	0.1497***	0.1575***
	(0.0254)	(0.0254)	(0.0323)	(0.0330)
Game C-3	0.1360***	0.1182***	0.1299***	0.1412***
	(0.0218)	(0.0231)	(0.0254)	(0.0253)
Game A(8)	0.0285	0.0317	−0.0037	0.0096
	(0.0321)	(0.0325)	(0.0372)	(0.0373)
Game order	−0.0111***	−0.0116***	−0.0087**	−0.0094**
	(0.0035)	(0.0035)	(0.0042)	(0.0042)
Risk aversion			58.507***	
			(14.3315)	
Female				0.0165
				(0.0124)
Individual effect	Yes	Yes	No	No
N	1963	1963	1963	1963
Pseudo R2	0.7839	0.7896	0.6255	0.6201
Wald chi2	350.95	341.84	451.01	460.35
Log pseudolikelihood	−268.378	−261.329	−465.161	−471.936

Robust standard errors are in parentheses. ***1 % significance, **5 % significance, *10 % significance. The dependent variable represents one if a subject chooses to stop searching, regardless of whether the search is actually ended. Game B-1 (common value + one-vote rule), Game B-2 (common value + majority rule), Game B-3 (common value + unanimity rule), Game C-1 (different values + one-vote rule), Game B-2 (different values + majority rule), Game C-3 (different values + unanimity rule). Game A(8) is the single-agent search game that subjects played in the last trial, which captures the anchoring effect, compared with the first trial of Game A (A(1)). We calculate the absolute risk aversion index using the willing -to-pay price for a lottery with a 25 % chance of winning JPY2,000 (EUR15.8) but a 75 % chance of receiving JPY0 (EUR0)

Table 15.7 Tests for threshold effects

	Null hypotheses	Difference in the coefficients	Std. err.	z value	p-value	Test
[1]	Coef(C-1)-Coef(B-1)≤0	0.0745	0.0307	2.4264	0.0078	Rejected
[2]	Coef(C-2)-Coef(B-2)≤0	0.0628	0.0187	3.3591	0.0000	Rejected
[3]	Coef(C-3)-Coef(B-3)≤0	0.0011	0.0226	0.0505	0.3050	Not rejected

Note: The coefficients on Games B and C show the average threshold effects compared with the first trial of Game A from column [2] of Table 15.6. The magnitude of the coefficients on Game B is attributable to the heterogeneity among group members in terms of their preferences. The magnitude of the coefficients on Game C is attributable to the heterogeneity of preferences among group members plus the heterogeneity in terms of what value the other group members draw. Therefore, the differences between the coefficients between Games B and C indicate the average threshold effects derived only by the second heterogeneity in terms of what value the other members draw, after controlling for the first heterogeneity among group members

those corresponding to Games C, which allows us to control for the heterogeneity of preferences across group members. As shown in Table 15.7, we significantly reject the null hypotheses that the average marginal probability of voting to stop in Game C is equal to or lower than that in Game B under the one-vote and majority rules. This implies that subjects vote to stop the search in a stochastically earlier round under the one-vote and majority rules. We thus support Proposition 2 in the AAV model (H3), stating that subjects are less picky in terms of their acceptance standard in committee search than in single-agent search. However, contrary to our expectations, we cannot reject (H3) in the case of the unanimity rule.

Next, we compare the coefficients on Games C in terms of the magnitude to test the second part of Proposition 5 (H4). The probability of voting to stop searching varies according to the plurality voting rules. Albrecht et al. (2010) showed that the reservation value is either hump-shaped or monotonically increasing in the number of votes required to stop searching; that is, the probability of voting to stop the search is either U-shaped or monotonically decreasing in the number of votes required to stop searching. This implies that at least, the probability of voting to stop under the majority rule must be lower than under the one-vote rule. This is the statistical hypothesis testing that we must undertake.

Table 15.8 displays the results from testing (H4) using the estimated results from column [2] of Table 15.6. The first row in Table 15.8 tests the null hypothesis that the average marginal probability of voting to stop the search in Game C-1 is equal to or lower than that in Game C-2 (that is, the null hypothesis of $R_{C1} \geq R_{C2}$). As before, we control for the heterogeneity of preferences among group members by deducting the coefficients for Games B from the corresponding values for Games C. As shown, we cannot significantly reject this null hypothesis, contrary to our expectations. Nevertheless, we note that the statistical hypothesis testing does not necessarily mean that we *accept* the null hypothesis. Just for the record, we additionally test the null hypothesis that the average marginal probability of voting to stop the search in Game C-2 is equal to or lower than that in Game C-3 (that is, the null hypothesis

Table 15.8 Comparisons of the reservation values among plurality voting rules

	Null hypotheses	Difference in the coefficients	Std. Err.	z value	p-value	Test
[1]	[Coef(C-1)-Coef(B-1)]-[Coef(C-2)-Coef(B-2)]≤0	0.0117	0.0356	0.3292	0.3707	Not rejected
[2]	[Coef(C-2)-Coef(B-2)]-[Coef(C-3)-Coef(B-3)]≤0	0.0616	0.0285	2.1634	0.0154	Rejected

Note: The coefficients on Games B and C also show the average probability effects of voting to stop searching, compared with the first trial of Game A from column [2] of Table 15.6. Similarly to Table 15.7, the differences between the coefficients between Games B and C indicate the average probability effects of voting to stop searching derived only by the second heterogeneity in terms of what value the other members draw, after controlling for the first heterogeneity among group members. In other words, [1] tests the null hypothesis that the "reservation value" is equal to or lower in the majority rule than in the one-vote rule. [2] tests the null hypothesis that the "reservation value" is equal to or lower in the unanimity rule than in the majority rule

of $R_{C2} \geq R_{C3}$), as shown in the second row of Table 15.8. Contrary to the first row, we significantly reject the null hypothesis at the 5 % level of significance. Therefore, the reservation value is lower under the majority rule (Game C-2) than under the unanimity rule (Game C-3). According to these two testing exercises, we cannot reject the joint null hypothesis of $R_{C1} \geq R_{C2}$ *and* $R_{C2} \geq R_{C3}$, and it is also obvious that we cannot reject the joint null hypothesis of $R_{C1} \geq R_{C2}$ *and* $R_{C2} \leq R_{C3}$.

Table 15.8 numerically indicates that the probability of voting to stop falls from the one-vote rule to the majority rule and then falls even further from the majority rule to the unanimity rule, but that its difference between the one-vote rule and the majority rule is not statistically significant. This provides weak evidence to quantitatively support the prediction of the AAV model that a subject votes in favor of stopping the search in a stochastically earlier round under the one-vote rule (Game C-1), followed by the majority rule (Game C-2) and the unanimity rule (Game C-3). In other words, the reservation value is lowest in Game C-1, followed by Game C-2, and then Game C-3 (that is, $R_{C1} < R_{C2} < R_{C3}$). One of the reasons for this result may be that the subject's preferences in relation to the two negative externalities (i.e., the first externality relating to the committee search stopping despite the subject's preference to continue, and the second externality where the committee search continues despite the subject's preference to stop) are different. Recall that the one-vote and unanimity rules are strongly influenced by the first and second externalities, respectively, whereas the majority rule is influenced by both externalities, but only moderately. If the subject incurs a larger disutility from the first externality than the second, she or he tends to vote to stop in a stochastically earlier round under the one-vote rule, followed by the majority rule, and then the unanimity rule, because she or he does not want other members to stop the committee search, despite her or his preference to continue searching. Our results then provide weak evidence that subjects incur a larger disutility from the negative externality involving the stopping of the committee search despite the subject's preference to continue than from the second negative externality relating to the committee search continuing despite the subject's preference to stop.

Other interesting variables also affect the probability of voting to stop searching. The coefficient on drawn value is positive at the 1 % level of significance in all columns of Table 15.6. When a subject draws a higher value, she or he is more likely to vote to accept it. According to column [2] of Table 15.6, the coefficient on round remains positive at the 1 % level of significance. This refutes our expectation that the round is not a determinant of voting to stop searching in an infinite-horizon sequential search. The variable regarding the game order is negative at the 1–5 % levels of significance in all columns. This implies that subjects vote more aggressively in favor of continuing to search in later games.

A final question is whether there are any systematic differences in the probability of voting to stop searching among subjects. To explore this, we employ the individual characteristics, including the extent of absolute risk aversion and gender. The variable indicating the extent of absolute risk aversion is as expected. Its coefficient on absolute risk aversion is positive at the 1 % level of significance in column [3] of Table 15.6, implying that more risk-averse subjects are more likely to vote to stop searching in a stochastically earlier round.

6 Concluding Remarks

This chapter described a laboratory experiment to study an infinite-horizon sequential search-by-committee model and tested some of the implications obtained in Albrecht et al. (2010). To date, there have been no empirical studies on committee search, mainly because of the difficulty in collecting suitable data. Using our laboratory experiment, we collected original data from subjects. This chapter's main contributions are to provide experimental evidence about committee search and then to test the properties of search duration and voting behavior for various plurality voting rules. Our experimental design involved decomposing the source of the difference in search behavior between single-agent search and committee search into effects caused by heterogeneity with respect to preferences among group members and other heterogeneity in terms of the different values other members draw from the identically independent distribution.

Our findings are summarized as follows. After controlling for the heterogeneity of preferences among group members, the average search duration is longer in committee search with the unanimity rule than in single-agent search, whereas the average search duration is shorter in committee search in which at least one vote is required to stop the search than in single-agent search. In a comparison of single-agent search versus committee search with the majority rule, the hypothesis of no difference in search duration is not significantly rejected. This result is not surprising given the properties of the AAV model stating that the duration of committee search is either longer or shorter than that of single-agent search, depending on the value of the discount factor. These results imply that search duration is increasing in the number of votes required to stop committee search. This supports the first part of Proposition 5 in the AAV model (H1). In addition, we found that the search duration is increasing in group size, holding the unanimity rule fixed, when we compare the search duration in single-agent search with that in committee search with the unanimity rule. This result is consistent with the second part of Proposition 4 in the AAV model (H2).

To identify the threshold effect whereby negative externalities caused by committee search involving voting operate to lower a member's reservation value, we estimated the determinants of voting to stop searching. We found that subjects are more likely to vote to stop searching in committee search than in single-agent search. These estimated results confirm the threshold effect, in the sense that agents are less picky in committee search than in single-agent search, as the AAV model indicates in Proposition 2.

The AAV model predicts that the reservation value is either hump-shaped or monotonically increasing in the number of votes required to stop searching. That is, the probability of voting to stop searching is either U-shaped or monotonically decreasing in the number of votes required to stop searching. This implies that at least, the probability of voting to stop under the majority rule must be lower than under the one-vote rule. Unfortunately, our experimental outcome cannot significantly support this prediction in terms of the size order of the reservation values. Comparing the size of the coefficients on games, however, we found that the probability of voting to stop falls from the one-vote rule to the majority rule and then falls even further from the majority rule to the unanimity rule, meaning that the reservation value is lowest under the one-vote rule, followed by the majority rule, and then the unanimity rule. This is weakly and quantitatively consistent with the predictions of the AAV model. This result implies that subjects who participated in the experiment incurred larger disutility from the negative externality involving committee search stopping despite a preference to continue, than from the second negative externality relating to committee search continuing despite a preference to stop.

Overall, our experimental outcomes are consistent with the implications in the AAV model in terms of comparisons of the search duration and the probability of voting to stop searching for committee versus single-agent search. However, the outcomes cannot statistically support the AAV model's prediction according to the different plurality voting rules in terms of comparison of the probability of voting to stop searching in the committee search model.

Of greatest interest to us now in this research topic is whether we would obtain the same experimental results with larger groups. As in the AAV model, we confirmed from the test of (H2) in our experiment that the search duration was longer as the group size increased from one to three under the unanimity rule. Our expectation is that the search duration will become even longer as the group size increases to, say, five under the unanimity rule because the vote aggregation effect operates more strongly. We therefore need to check for consistency in our results for larger groups. This provides one of many directions for our future research.

Appendix

Instructions

Note: Following are the instructions for Session 4 at Hokkaido University. Other sessions differed from this session in terms of the order of the games.

Welcome to our experiment! In this experiment, you will be asked to play eight games. In each game, you will be asked to choose either to accept a value that is randomly selected from a uniform distribution with a lower bound of zero and an upper bound of 3,000, or to refuse this value and move on to the next round to wait for a higher value. If you are willing to accept an offered value, you click on the "Y" displayed on the PC screen; if not, you click on the "N". You can continue to search as long as you want, but please remember that your search activity will be terminated coercively with a probability of 5 %, in which case you will automatically receive the value drawn in the round immediately before termination. Your score will be determined according to the values that you accept.

We would like you to play eight different games. The first game is as follows.

- Game A: In each round, the computer randomly selects a value from a uniform distribution with a lower bound of zero and an upper bound of 3,000. You decide whether to accept the value drawn from this distribution. If you accept the value, then you finish your search and the value is your score. If you do not accept the value, you move on to the next round and observe another value newly drawn by the computer.

The next three games are as follows.

- Game B-1: You are grouped with two other participants. Grouping is done randomly by the computer, and no member knows who the other members are. In this treatment, you play a committee search activity with the other two members. In each round, the computer randomly selects a value for *all* three group members, including you, from a uniform distribution with a lower bound of zero and an upper bound of 3,000. All three group members, including you, receive the same value. You independently decide whether to accept the drawn value. If you prefer to accept the value, you vote for stopping the search, but if you do not accept the value, you vote against stopping. This committee search activity is stopped if at least one member of the group votes for stopping. Otherwise, your group moves on to the next round and observes another value newly drawn by the computer.
- Game B-2: The process of Game B-2 is similar to that of Game B-1 except for the plurality voting rule; that is, this committee search activity is stopped only if at least two-thirds of the members of the group vote for stopping.

- Game B-3: The process of Game B-3 is similar to that of Game B-1 except for the plurality voting rule; that is, this committee search activity is stopped only if all three members of the group vote for stopping.

  The next three games are as follows:

- Game C-1: You are grouped with two other participants. Grouping is done randomly by the computer, and no member knows who the other members are. In this treatment, you play a committee search activity with the other two members. In each round, the computer randomly selects for *each* group member a value from a uniform distribution with a lower bound of zero and an upper bound of 3,000. Each group member therefore has a different value, and you do not know what value the other two members draw, and vice versa. You decide whether to accept the drawn value. If you accept your value, you vote for stopping the search, but if you do not accept your value, you vote against stopping. This committee search activity is stopped if at least one member of the group votes for stopping. Otherwise, your group moves on to the next round, and each member receives another value newly drawn by the computer.
- Game C-2: The process of Game C-2 is similar to that of Game C-1 except for the plurality voting rule; that is, this committee search activity is stopped only if at least two-thirds of the members of the group vote for stopping.
- Game C-3: The process of Game C-3 is similar to that of Game C-1 except for the plurality voting rule; that is, this committee search activity is stopped only if all three members of the group vote for stopping.

  The final game is as follows.

- Game A: In each round, the computer randomly selects a value from a uniform distribution with a lower bound of zero and an upper bound of 3,000. You decide whether to accept the value drawn from this distribution. If you accept the value, then you finish your search and the value is your score. If you do not accept the value, you move on to the next round and observe another value newly drawn by the computer.

Before starting the experiment, we would like you to practice Game A once. Please let us know if you have any questions. We will explain the rule of each game again before it starts. After the experiment, please respond to a questionnaire. You will be paid an appearance fee of JPY1,000. The performance pay will be determined based on one of the scores from the eight games you randomly choose, and your payment will be calculated as JPY1 for each scoring point. Payment processes will take place after the experiment is concluded. Please be quiet and do not communicate with other participants during the experiment. Thank you for your participation.

Addendum: Finite-Horizon Sequence Search Case[17]

This addendum reports results from finite-horizon versions of the same experiments. Our experiments were conducted in the experimental laboratory of the Institute of Social and Economic Research at Osaka University on December 1, 2009. The experiments consisted of three sessions for reasons relating to laboratory capacity. Each session involved the same eight games (two games of Game A and three subgames each for Game B and Game C). Each game would end after 20 rounds. Recall was not allowed in this model, too. The finite-horizon sequential search model differs from the AAV model (Albrecht et al. 2010) in terms of the way each search activity is terminated, but it is more realistic and more easily understandable for subjects in terms of the experimental structure of the search activity.

Each session differed in terms of the order of the games; the order of the first session was Games A, B-3, B-2, B-1, C-3, C-2, C-1, and A; the order of the second session was Games A, C-1, C-2, C-3, B-3, B-2, B-1, and A; and the order of the third session was Games A, B-1, B-2, B-3, C-1, C-2, C-3, and A. The order of the games was changed to control the anchoring effect, whereby search behavior in a game is affected by the previous games' results. The experiment involved a total of 63 undergraduate and graduate noneconomics-major students from Osaka University. The participants were seated at individual desks in each session. The payoff procedures were the same as the ones when the experiment was conducted at Hokkaido University.

We begin with a test of the first hypothesis (H1). Addendum Table 15.9 displays descriptive statistics, including average durations and their standard deviations. The average search duration is longer in the case of committee search with the majority and unanimity rules when group members have different values (Game C) compared with when they have the same value (Game B), but the reverse is observed when the

Table 15.9 Average search durations (finite-horizon sequential search)

	Value	Voting rule	# searchers	Sample	Mean	Sd	Max	Min
Game A(1)			1	63	5.079	4.408	19	1
Game B-1	Common	One vote	3	63	3.762	2.487	9	1
Game B-2	Common	Majority	3	63	3.952	3.553	12	1
Game B-3	Common	Unanimity	3	63	6.524	4.568	20	1
Game C-1	Different	One vote	3	63	1.381	0.580	3	1
Game C-2	Different	Majority	3	63	4.476	3.026	12	1
Game C-3	Different	Unanimity	3	63	11.762	7.141	20	2
Game A(8)			1	63	5.984	5.253	20	1
Average				63	5.365	5.114	20	1

[17]This addendum has been newly written for this book chapter.

Table 15.10 Comparisons with single-agent search (finite-horizon sequential search)

Game B-1	Common	One vote	The search duration is shorter at the 1 % level of significance
Game B-2	Common	Majority	The search duration is shorter at the 5 % level of significance
Game B-3	Common	Unanimity	The search duration is longer at the 10 % level of significance
Game C-1	Different	One vote	The search duration is shorter at the 1 % level of significance
Game C-2	Different	Majority	The search duration is insignificantly shorter
Game C-3	Different	Unanimity	The search duration is longer at the 1 % level of significance

one-vote rule applies. As the required number of votes to stop searching increases, the average search duration is longer, regardless of whether the drawn value is the same or different among group members.

Addendum Table 15.10 shows comparative tests of the search duration between committee search with various plurality voting rules versus single-agent search. The upper three rows of this table compare the average duration for single-agent search and committee search in which all group members draw the same value (Game B). According to the AAV model, if subjects are homogeneous with respect to their risk and loss attitudes, preferences, and other unobserved factors, there is no difference in the average duration between the two search models, regardless of which plurality voting rule applies. Addendum Table 15.10 shows that the null hypothesis of no difference in the average duration between the two search models is significantly rejected under any plurality voting rule. Therefore, we can say that the subjects are not homogeneous with respect to their risk and loss attitudes, preferences, and unobserved factors. To correctly test the implications of the AAV model where members are homogeneous, it is necessary to control these differences that arise from heterogeneity among members.

The lower three rows of Addendum Table 15.10 compare the average duration between single-agent search and committee search in which group members draw different values. Remember that the difference in the average duration between the two models arises from the heterogeneity among group members, as mentioned above, and also from uncertainty in terms of the values that other members draw. The search duration is shorter in committee search with the one-vote rule than in single-agent search at the 1 % level of significance. Similarly, the search duration is longer in committee search with the unanimity rule than in single-agent search at the 1 % level of significance. The search duration is shorter in committee search with the majority rule than in single-agent search, but insignificant in Addendum Table 15.10.

Addendum Table 15.11 shows comparisons of the average duration between single-agent search and committee search, controlling for heterogeneity among group members. According to the first row of Addendum Table 15.11, the search

Table 15.11 Comparison with search durations of committee search (finite-horizon sequnetial search)

Game B-1 vs Game C-1	The search duration is shorter in Game C-1 at the 1 % level of significance
Game B-2 vs Game C-2	The search duration is insignificantly longer in Game C-2
Game B-3 vs Game C-3	The search duration is longer in Game C-3 at the 1 % level of significance

The experiment was conducted at Osaka University

duration is shorter in committee search with the one-vote rule than in single-agent search at the 1 % level of significance. Similarly, the null hypothesis that the average duration in committee search with the unanimity rule is equal to or shorter than the average duration in single-agent search is rejected at the 1 % level of significance. Another interpretation of this result is that as the number of group members increases from one to three, holding the unanimity rule fixed, the average duration becomes longer. These results are consistent with the prediction of the AAV model. Looking at the comparisons between single-agent search and committee search with the majority rule, we cannot significantly reject any null hypothesis that the average duration differs, which does not necessarily contradict the implication from the AAV model.

We estimate effects of various plurality rules on the search duration and the probability of a subject voting for stopping his or her search, using data of the finite-horizon versions. The results are qualitatively similar to those from the finite-horizon versions.

References

Albrecht J, Anderson A, Vroman S (2010) Search by committee. J Econ Theory 145:1386–1407

Boone J, Sadrieh A, van Ours JC (2009) Experiments on unemployment benefit sanctions and job search behavior. Eur Econ Rev 53:37–51

Compte O, Jehiel P (2010) Bargaining and majority rules: a committee search perspective. J Politi Econ 118:189–221

Cox JC, Oaxaca RL (1989) Laboratory experiments with a finite-horizon job-search model. J Risk Uncertain 2:301–30

Cramer JS, Hartog J, Jonker N, Van Praag CM (2002) Low risk aversion encourages the choice for entrepreneurship: an empirical test of a truism. J Econ Behav Organ 48:29–36

Fischbacher U (2007) Z-Tree Zurich toolbox for ready-made economic experiments. Exp Econ 10:171–178

Harrison GW, Morgan P (1990) Search intensity in experiments. Econ J 100:478–486

Hizen Y, Kawata K, Sasaki M (2013) An experimental test of a committee search model. Eur Econ Rev 61:59–76

Schunk D (2009) Behavioral heterogeneity in dynamic search situations: theory and experimental evidence. J Econ Dyn Control 33:1719–38

Schunk D, Winter J (2009) The relationship between risk attitudes and heuristics in search tasks: a laboratory experiment. J Econ Behav Organ 71:347–60

Chapter 16
Equilibrium Refinement Versus Level-k Analysis: An Experimental Study of Cheap-Talk Games with Private Information

Toshiji Kawagoe and Hirokazu Takizawa

Abstract We present the experimental results of cheap-talk games with private information. We systematically compare various equilibrium refinement theories and bounded rationality models such as level-k analysis in explaining our experimental data. As in the previous literature, we find that when interests between sender and receiver are aligned, informative communication frequently arises. While babbling equilibrium play is observed more frequently in conflicting interest cases, a substantial number of players tend to choose truth-telling and credulous play. We also find that level-k analysis outperforms equilibrium refinement theories in explaining this phenomenon. Our results also confirm the existence of the "truth bias" and "truth-detection bias" reported in communication theory.

Keywords Cheap talk • Refinement • Level-k analysis • Truth-bias • Experiment

JEL Classification Codes C72, C92, D82

The original article first appeared in Games and Economic Behavior 66: 238–255, 2009. A newly written addendum has been added to this book chapter.

T. Kawagoe (✉)
Department of Complex Systems, Future University – Hakodate, 116-2 Kameda Nakano cho, Hakodate, Hokkaido 041-1112, Japan
e-mail: kawagoe@fun.ac.jp

H. Takizawa
Graduate School of Management and Information Sciences, Tama University, 4-1-1 Hijirigaoka, Tama City, Tokyo 206-0022, Japan
e-mail: takizawa@tama.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_16

1 Introduction

Game theorists have explored strategic problems arising in information transmission by using cheap-talk games. Cheap-talk games are generally classified into two types: the cheap-talk game with complete information and the cheap-talk game with incomplete information (see the survey paper by Crawford 1998 and Chap. 7 in Camerer 2003).

For cheap-talk games with complete information, the primary focus of research has been on whether a player's intentions regarding his/her future actions can be correctly inferred through pre-play communication. This line of study was initiated by the seminal papers of Farrell (1987, 1988). Since then, from the earliest papers by Cooper et al. (1989, 1992) to recent papers such as Ellingsen and Östling (2006) and Crawford (2007), a number of experimental works have been submitted. Research conducted in this strand has basically proved that one-way pre-play communication works well as a coordination device between players. Recently, the research trend seems to have shifted to analysis of communication when interests between sender and receiver conflict. Ellingsen and Östling (2006), Crawford (2007), and Yamamori et al. (2008) follow this line.

The study of cheap-talk games with incomplete information was initiated by Crawford and Sobel (1982). Since then, research in this area has been concerned with identifying conditions under which the sender's private information can be conveyed to the receiver correctly through cheap talk. Crawford and Sobel (1982) showed that an equilibrium arises with informative communication when interests between sender and receiver are sufficiently aligned. Even in this case, cheap-talk games always have babbling equilibria where communication is totally uninformative, contrary to the fact that informative communication is frequently observed in our daily life. Thus, most theoretical works, the so-called equilibrium refinement theories, among which are Farrell (1993), Matthews et al. (1991), and Rabin and Sobel (1996), focus on how to justify informative separating equilibria. These works are examined later. These papers somehow succeed in selecting separating equilibria in aligned interest cases. On the experimental front, early papers such as Dickhaut et al. (1995), Blume et al. (1998, 2001), and Kawagoe and Takizawa (1999) observed good convergence towards separating equilibrium play in the laboratory in games where both players' interests sufficiently coincide (see also Crawford 1998 and Chap. 7 in Camerer 2003).

As in the research of complete information games, research interest also seems to have shifted to conflicting interest cases in incomplete information games. For example, Holm (2010), Gneezy (2005), Sánchez-Pagés and Vorsatz (2007), and Holm and Kawagoe (2010) study games like "matching pennies." Kawagoe and Takizawa (2005), Cai and Wang (2007), and Wang et al. (2010) examine a series of cheap-talk games à la Crawford and Sobel (1982) in the laboratory and find that overcommunication occurs. That is, senders tend to tell the truth even in games with conflicting interests that have only babbling equilibria. These papers show that bounded rationality models such as AQRE (agent quantal response equilibrium) by McKelvey and Palfrey (1998) and level-k analysis works well in predicting subjects' behavior.

Turning to the communication-theoretic literature, it has been reported that in situations with conflicting interests, although senders usually lie, most receivers believe senders' messages. This phenomenon is called "truth bias" (McCornack and Parks 1986). Burgoon et al. (1994) also show that for novices as well as experts, such as military intelligence instructors, the accuracy rate of detection is much higher for truthful messages than for deceptive ones. This phenomenon is called "truth-detection bias," a well-known but still disputable phenomenon (Vrij 2000; Holm 2010; Holm and Kawagoe 2010).

Against this backdrop, the present chapter tries to contribute to the current literature by presenting the experimental results of cheap-talk games with private information that have systematically varying degrees of preference alignment. In so doing, we identify deviations from the predictions given by equilibrium refinement theories, such as overcommunication, truth bias, and truth-detection bias. We then explain such deviations using a level-k model, a non-equilibrium theory of players' initial responses to games that reflects the strategic thinking of players. This model has been applied to games with and without communication, and has succeeded in explaining a number of anomalous behaviors found in the laboratory (Stahl and Wilson 1995; Nagel 1995; Ho et al. 1998; Camerer et al. 2004; Costa-Gomes et al. 2001;, Costa-Gomes and Crawford 2006; Crawford and Iriberri 2007a; and others).

The games we used have different payoff characteristics for representing different degrees of preference alignment between sender and receiver. Two of them have separating equilibria and the other has only babbling equilibria. Consistent with the previous literature, we find that the more aligned the interests are between sender and receiver, the more frequently the sender tells the truth and the receiver believes the sender's messages to be truthful. While babbling equilibrium play is observed more frequently in conflicting interest cases, a substantial number of receivers still believe senders' messages to be truthful. Thus, our experimental results reveal that subjects' behavior, either in aggregate or in individual level, deviates from the predictions given by equilibrium refinements, when the interests between sender and receiver conflict. Truth bias and truth-detection bias are also clearly observed. However, as we show, such deviations from the standard equilibrium model are explained consistently only by level-k analysis.

The work of Blume et al. (2001) is most closely related to the present chapter, which also reports the experimental results of cheap-talk games with incomplete information.[1] While our results are similar to theirs both in aligned and misaligned preference cases, the games we use have different features from theirs. Since senders' messages have clearly defined literal meanings in our games, truth-telling is clearly distinguished from lying. This enables us to explicitly examine senders'

[1]Our findings were independently obtained at almost the same time. The earliest version of our paper appeared in 1999 (Kawagoe and Takizawa 1999).

and receivers' behavior regarding overcommunication and truth bias, which is the marginal contribution of this chapter. We also comprehensively examine the various equilibrium refinement concepts so far proposed for cheap-talk games such as neologism-proofness, announcement-proofness, and recurrent *mop* in the laboratory.[2]

The organization of the chapter is as follows. The next section presents the games we use in the experiment and compares predictions of play for these games using various theories: sequential equilibrium and its refinements, AQRE, and level-k analysis. The relation between our experiment and communication theory is also discussed. Based on those predictions, we set up several hypotheses. Section 3 presents experimental procedures and results. We also show that level-k analysis provides the most consistent prediction for our data. Concluding remarks are given in Sect. 4.

2 Theory and Hypotheses

2.1 Games and Their Sequential Equilibria

We consider cheap-talk games with private information where a sender sends a payoff-irrelevant message about her type (private information) to a receiver, who then takes a payoff-relevant action. For the purpose of analysis, the games are made as simple as possible. Specifically, the type space for the sender is $T = \{A, B\}$ with the prior probability of each type being 1/2. Knowing her true type, the sender sends a message from her message space $M = \{$"I am type A," "I am type B"$\}$, which we henceforth denote simply as "*a*" and "*b*" respectively as long as no confusion arises. The fact that there is a common language,[3] so that messages have clearly defined literal meanings, means that truth-telling is clearly distinguished from lying. The receiver, observing the sender's message, chooses from his action space $D = \{X, Y, Z\}$. The payoffs for both players are then determined according to the combination of the sender's true type and the receiver's action.

Focusing on cases of interest, we take X and Y to be the best actions for the receiver when the sender types are known to be A and B respectively. Z is introduced as the best action for the receiver when his belief closely approximates the prior belief.[4] In most of what follows, we restrict our main attention to pure strategies or

---

[2]For signaling games, Brandts and Holt (1992, 1993), Banks et al. (1994), and Cadsby et al. (1998) compared performance of various refinements concepts such as the intuitive criterion in laboratory experiments.

[3]Blume et al. (1998) examine how the meaning of the language evolves in the laboratory *without* common language between players.

[4]The labels for the receiver's action we used in the experiments were A, B, and C for X, Y, and Z respectively in Session 1, and they were then permuted from Session 2 on to prevent the labels from working as a coordination device. However, it is convenient to use X, Y, and Z as indicated in the text when we need to classify the play of receivers.

Table 16.1 Sender-receiver game payoff

		Action		
		X	Y	Z
Game 1				
Type	***A***	4, 4	1, 1	3, 3
	B	1, 1	4, 4	3, 3
Game 2				
Type	***A***	3, 4	2, 1	4, 3
	B	2, 1	3, 4	4, 3
Game 3				
Type	***A***	4, 4	1, 1	2, 3
	B	3, 1	2, 4	4, 3

pure-strategy equilibria.[5] Thus, we denote the sender's strategy as (m_A, m_B), where $m_A(m_B)$ is the message sent by type A (B) respectively. Similarly, (r_a, r_b) denotes the receiver's strategy, where $r_a(r_b)$ is the action the receiver takes upon receiving message a (b) respectively.

Even this simplest possible setting can encompass diverse incentive situations between sender and receiver. To specify the payoffs of games used in the experiments, we adopt three general incentive situations as follows:

Case 1. Both sender type A and B want to be correctly identified, inducing the receiver to choose action X and Y respectively;
Case 2. Both sender types want the receiver to play Z, that is, they want to confuse the receiver;
Case 3. Sender type A wants to be correctly identified, while sender type B wants to be misidentified as type A.

These cases correspond to Games 1, 2, and 3 respectively, the payoffs of which are shown in Table 16.1. Games 1 and 2 have two kinds of sequential equilibria: separating equilibria and babbling equilibria. However, both players' interests do not perfectly coincide in Game 2. Game 3 has only babbling equilibria.

To see the difference among these games more clearly, it is illuminating to place them within the framework set by Crawford and Sobel (1982), where the alignment of preferences can be expressed by a single parameter. In their model, sender types are drawn from a unit interval, and the sender's and receiver's preferences over actions are concave. More specifically, the payoff for sender type t from action a is $-[a - (t - d)]^2$ and the receiver's payoff is $-(a - t)^2$, where the alignment of preferences is measured by parameter d. In our games, parameter d is type dependent as in Blume et al. (2001). Let the type space be discretized so that $t = 1/4$ for type A and $t = 3/4$ for type B. Then, Game 1 is characterized by $d = 0$ for both types. While in Game 2, we have $d = 1/5$ for type A and $d = -1/5$

[5]See Kawagoe and Takizawa (1999) for detailed information on sequential equilibria of these games including mixed strategies.

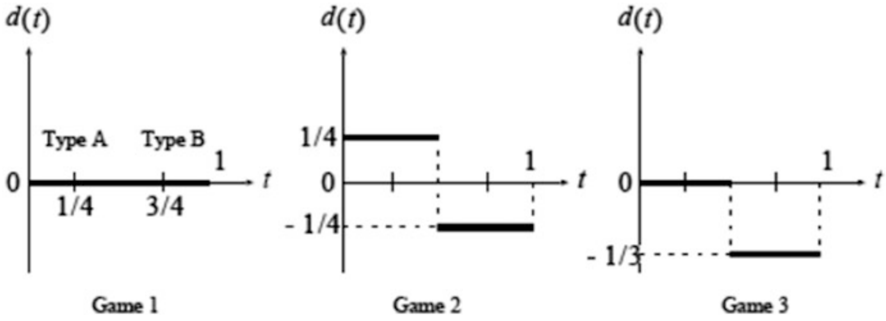


Fig. 16.1 Incentives of both sender types in each game

for type B, in Game 3, we have $d = 0$ for type A and $d = -1/3$ for type B. A characterization of the incentive alignments in our games is shown in Fig. 16.1.

Parameter d being type dependent, it is difficult to introduce into our games a clear-cut partial order regarding the degree of alignment of interests between sender and receiver. As will soon be seen (see Table 16.3), however, all the refinement theories we consider predict babbling equilibria for Game 3, and all the refinement theories except AQRE predict separating equilibrium for Game 1. Game 2 has separating equilibria as well as babbling equilibria, although all refinement theories predict babbling equilibria. From these considerations, Game 2 is somewhere between Games 1 and 3 with respect to alignment of interests between sender and receiver. Game 1 has the highest alignment, Game 2, the second highest, and Game 3, the lowest.

2.2 *Equilibrium Refinements*

It is well known that every cheap-talk game has so-called babbling equilibria in which the sender's messages are totally uninformative. While our intuition suggests that such an implausible outcome should be ruled out in the case of aligned interests, the by-now standard refinement arguments such as Cho and Kreps (1987) intuitive criterion have no bite in cheap-talk games. Accordingly various refinement theories specific to cheap-talk games have been submitted.

Farrell (1993) launched this line of study by proposing "neologism-proofness." In this chapter, he assumes that for any nonempty subset of types, there is a message with literal meaning, "my type is in the subset." Given a putative equilibrium, this kind of message is a "neologism" if it is not used in the equilibrium. A neologism is credible relative to the equilibrium if the sender type in the corresponding subset prefers the neologism to be believed over what they would receive in the putative equilibrium. A neologism-proof equilibrium is an equilibrium relative to which no credible neologism exists.

Amending some shortcomings of neologism-proofness, Matthews et al. (1991) proposed "announcement-proofness." They generalize a possible deviation message to an announcement strategy that consists of a nonempty set of deviant types and a talking strategy that maps from the set of deviant types to a set of (possibly mixed) messages. They then define successively stronger concepts of credible messages.[6] An announcement strategy is weakly credible if: (1) any deviant type prefers sending a message specified by the talking strategy to staying in the original equilibrium; (2) every non-deviant type prefers adopting her equilibrium strategy to sending any message that some deviant type may use in the talking strategy; and (3) no deviant type gains by using messages that are not assigned to her by the talking strategy. A credible announcement strategy further requires, in addition to (1), (2), and (3), that (4) any deviant type having two weakly credible announcement strategies prefers the outcomes obtained using the announcement strategy in question to those using another strategy. In response to the Stiglitz critique,[7] a strongly credible announcement is defined as requiring, in addition to (1), (2), and (4), that (3′) there is an equilibrium in which the set of messages some deviant type may send and the set of messages a non-deviant type may send are mutually exclusive.

Using these concepts of credible message, successively weaker concepts of announcement-proofness are defined. A strongly announcement-proof equilibrium is an equilibrium relative to which no weakly credible announcement exists. An announcement-proof equilibrium is one relative to which no credible announcement exists. An equilibrium relative to which no strongly credible announcement exists is weakly announcement proof.

Rabin and Sobel (1996) develop another theory that takes into account the Stiglitz critique by focusing on the dynamics triggered by an initial deviation. They refer to play that can be observed infinitely often in the dynamics triggered by weakly credible announcement strategy à la Matthews et al. (1991) as "recurrent *mop*." "*Mop*" is an acronym for Matthews, Okuno-Fujiwara, and Postlewaite.

For Game 1, all the refinement theories point to separating equilibria as plausible play. Although Game 2 also has two sequential equilibria, separating and babbling, refinement theory predictions for Game 2 are in sharp contrast with those for Game 1; all the refinement theories predict babbling equilibria for Game 2. This game does not have partial common interest in the sense of Rabin and Sobel (1996). Theoretical predictions for Game 3, whose sequential equilibria are only babbling, are divided. The babbling equilibria pass the test posed by relatively permissible refinement criteria, such as weak announcement-proofness and recurrent *mop*, whereas it fails in more stringent tests. It is interesting to note that players necessarily deviate

[6]Of course, "successively stronger" is not a rigorous expression. However, the meaning should be clear from the explanation below. The same applies to the expression "successively weaker" for the concept of announcement-proofness.

[7]The standard refinement arguments usually stop where a deviant type succeeds in making the receiver believe that she is in some subset of T and respond optimally to this belief. However, the receiver might change his equilibrium belief when he receives the equilibrium message, and change his action accordingly. This might trigger further changes in messages that non-deviant types want to send.

from the babbling equilibrium outcome but they repeatedly come back to it in the deviation dynamics of recurrent *mop*.

McKelvey and Palfrey (1998) propose the concept of AQRE (agent quantal response equilibrium). The limiting set of AQREs, as precision parameter λ goes to infinity, is called limit AQRE, which they proved to be a sequential equilibrium giving a unique prediction of the game. Thus, limit AQRE can also be regarded as a sort of equilibrium refinement. For all the games we consider, AQRE predicts a mixed-strategy babbling equilibrium $\left(\frac{1}{2}a\frac{1}{2}b, \frac{1}{2}a\frac{1}{2}b\right), (Z, Z))$ for almost all values of λ.[8]

2.3 Level-k Analysis

In this subsection, we apply so-called level-k analysis to our games. The level-k model is a non-equilibrium theory of initial responses to games of players, reflecting their strategic thinking. While various types of non-equilibrium behavior are observed in the laboratory, it is to difficult to identify and specify subjects' decision rules among an enormous number of possibilities. The level-k model assumes that each subject's decision rules follow one of a small set of *a priori* plausible types and tries to estimate which type best fits the subject's behavior. Specifically, in level-k analysis, types are defined inductively as follows. Type Lk ($k \geq 1$) anchors its belief in type $L0$ and adjusts its belief through thought experiment. That is, the Lk ($k \geq 1$) player responds optimally to the $L\,(k-1)$ player.[9] This procedure implies that specifying type $L0$ is the key to the analysis, to which we will return soon. The level-k model has been applied to games with or without communication (see Stahl and Wilson 1995; Nagel 1995; Ho et al. 1998; Camerer et al. 2004; Costa-Gomes et al. 2001; Costa-Gomes and Crawford 2006; Crawford and Iriberri 2007a; among others). For application of the level-k model to cheap-talk games with complete information, see Ellingsen and Östling (2006) and Crawford (2007).

Since the games we consider involve incomplete information, some consideration is necessary to extend the previous framework to level-k analysis.[10] First, for the specification of the $L0$ type, we basically follow Ellingsen and Östling (2006) in

[8] A sketch of the proof is as follows. As the sender's payoff for each message is the same in cheap-talk games in general, by definition of the game, AQRE correspondence must assign the same probability to each message for every value of λ, regardless of the receiver's actions. On the other hand, as the receiver cannot distinguish the sender's type from such a message, action Z is the best response due to our choice of payoff function in the class of games we considered. Of course, as the receiver must assign equal probability to every action at $\lambda = 0$, by definition of AQRE correspondence, it is the only exception.

[9] In Ellingson and Östling (2006) terms, this is a SCH (simple cognitive hierarchy) model.

[10] To the best of our knowledge, Camerer et al. (2004) and Crawford and Iriberri (2007b) are the only papers analyzing games with incomplete information.

assuming that the $L0$ sender is the truth-teller and that the $L0$ receiver is the randomizer. Recall that the message space employed in our experiments is designed so that each choice corresponds to truth-telling or lying. In this context, truth-telling seems to be focal, a natural candidate for acting as an anchor in the players' minds. We will shortly say more on the possible specification of the $L0$ type. Second, since play in cheap-talk games can easily involve an off-the-path information set, frequently making several strategies indifferent, we have to decide how to cope with this situation. We assume that type Lk of a player role adheres to the strategy taken by lower type $L\ (k-1)$ of the same player role as long as that strategy continues to be one of the best responses to the play of the $L\ (k-1)$ type of the opponent player role.

In Games 1 and 2, the type $L1$ sender is indifferent in choosing between a and b, keeping the $L0$ type's strategy (a, b). The type $L1$ receiver finds it optimal to play (X, Y) against the type $L0$ sender's strategy (a, b). At this juncture, no higher type of either sender or receiver has an incentive to change his/her strategy. In Game 3, analysis up to the type $L1$ sender and receiver is the same as in Games 1 and 2. While the $L2$ sender of type A has no incentive to change her strategy, type B can gain by changing her strategy from sending message b to message a. Thus, the type $L2$ sender's strategy is (a, a). The type $L3$ receiver then benefits from changing to Z in response to message a, thus (Z, Y). At this juncture, the type A sender does not gain by changing her strategy. So, the strategies of any higher type remain the same. The analyses of Games 1, 2, and 3 are respectively summarized in Table 16.2.

One can think of other specifications of the $L0$ type than those used in the above analysis. There are two other plausible alternatives. First, we may specify the $L0$ receiver as a believing type, (X, Y). It is easy to see, however, that this specification gives exactly the same predictions as described above except for the $L0$ type. The only difference is that behaviors at $L1$ in our specification come down to $L0$. The other possibility is to let both the $L0$ sender and $L0$ receiver be the randomizer. This specification leads to babbling equilibrium play at any higher level. Since this

Table 16.2 Level-k predictions for Games 1, 2, and 3 when the $L0$ sender is the truth-teller and the $L0$ receiver is the randomizer

	Sender	Receiver
Game 1		
***L*1**	(a,b)	(X, Y)
***L*2**	(a,b)	(X, Y)
***L*3**	(a,b)	(X, Y)
Game 2		
***L*1**	(a,b)	(X, Y)
***L*2**	(a,b)	(X, Y)
***L*3**	(a,b)	(X, Y)
Game 3		
***L*1**	(a,b)	(X, Y)
***L*2**	(a,a)	(X, Y)
***L*3**	(a,a)	(Z, Y)

Table 16.3 Predictions of play for Games 1, 2, and 3 by various theories

	Game 1	Game 2	Game 3
Sequential equil.	*S.E* and *B.E.*	*S.E* and *B.E.*	*B.E.*
Refinements	*S.E.*[a]	*B.E.*	*B.E.*
Level-k[b]	(*ab*, *XY*)	(*ab*, *XY*)	(*aa*, *ZY*)

Note: S.E. means separating equilibria and B.E. means babbling equilibria
[a]AQRE predicts B.E. in this case
[b]Predicted play of higher levels is shown

prediction does not much differ from those obtained by other theories, we do not adopt it.

Thus, we assume that the $L0$ sender is the truth-teller and the $L0$ receiver is the randomizer as our default specification in what follows. Then, we will take play adopted by higher-level types as the prediction by the level-k model, since the strategy profiles adopted by higher types stabilize rather quickly in our games. To summarize, in the level-k model, the actions taken by higher-level thinkers are as follows: (1) In Games 1 and 2, the sender tells her type truthfully and the receiver believes the sender's messages. (3) In Game 3, both sender types say they are type *A*, and the receiver plays *Z* upon receiving *a* and *Y* upon receiving *b*. We now have all the predictions of play in our games given by various theories as summarized in Table 16.3.

2.4 Communication-Theoretic Experiments

In order to locate our experiments in the context of communication theory, it is worthwhile briefly reviewing the communication-theoretic literature on deception. Previous research in communication theory focused on deception has centered on nonverbal behaviors associated with uncontrollable psychological processes (Vrij 2000). These studies show that various nonverbal cues such as voice pitch and eye movement are not necessarily reliable signs for detecting deception, and that even well-trained specialists cannot distinguish between truth-telling and lying with more than a 60 % accuracy rate.

Especially related to the present chapter are papers by McCornack and Parks (1986) and Burgoon et al. (1994). As part of a hypothesis that relational development leads to a decrease in the accuracy of deception detection, McCornack and Parks (1986) propose the hypothesis that an increase in the confidence in truth/lie judgment leads to an increase in the presumption of honesty. They named the persistent presumption that the partners are telling the truth "truth bias." They confirmed the existence of truth bias in their experiment in a face-to-face environment. Burgoon et al. (1994) show that for novices as well as experts, such as military intelligence instructors, the accuracy rate of detection is much higher for truthful messages than for deceptive ones. This is called "truth-detection bias," a well-known but still disputable phenomenon (Vrij 2000; Holm 2010; Holm and Kawagoe 2010).

To address such issues as truth bias and truth-detection bias is also our concern in this experiment.

2.5 Hypotheses

As stated in the previous section, all the refinement theories except limit AQRE predict informative communication with separating strategies in Game 1. For Game 2, all the theories except level-k analysis predict babbling equilibrium play. All the theories predict babbling equilibrium play for Game 3, while level-k analysis predicts separating play for lower types. We summarize those predictions by all the refinement theories as the following hypothesis.

Hypothesis 1 (Equilibrium Prediction) *Most play conforms to a separating equilibrium in Game 1, while the majority of play conforms to a babbling equilibrium in Games 2 and 3.*

However, as previous studies of cheap-talk games with private information (Dickhaut et al. 1995; Blume et al. 1998; Cai and Wang 2007; Sánchez-Pagés and Vorsatz 2007) have shown, the sender's tendency toward truth-telling may be observed even in games with conflicting interests such as Games 2 and 3. If this is the case, this phenomenon is overcommunication in the sense that no equilibrium refinement theory predicts such truth-telling equilibria in Game 2 and that there is no truth-telling equilibrium in Game 3. In other words, we mean that overcommunication is the sender's truth-telling tendency relative to equilibrium or refinement prediction. So, if Hypothesis 1 is rejected for these games, we need a hypothesis that is more relaxed. If overcommunication occurs in games with conflicting interests, it enables us to further compare truth-telling propensity with regard to the incentive alignment of the games. This idea is formulated in Hypothesis 2.

Hypothesis 2 (Overcommunication) *Overcommunication occurs in games with conflicting interests. Further, the more aligned the interests are between sender and receiver, the more frequently the sender tells the truth.*

If overcommunication occurs on the sender's side, then there is good reason to anticipate that the receiver will tend to exploit it by believing the sender's messages to be truthful. Then, if Hypothesis 2 holds, we can also give a comparative-statics prediction on the receiver's side. This is Hypothesis 3, which is related to truth bias.

Hypothesis 3 (Truth Bias) *Receivers tend to believe senders' messages to be truthful even in games with conflicting interests. Furthermore, the more aligned the interests are between sender and receiver, the more frequently the receiver believes the sender's message to be truthful.*

In addition, we check whether the accuracy rate of detection differs for truthful and deceptive messages, i.e., whether truth-detection bias is observed in our environment. The following hypothesis is concerned with this.

Hypothesis 4 (Truth-Detection Bias) *The receiver guesses the sender's true type more correctly when the sender tells the truth than when she tells a lie.*

3 Experiment

3.1 Experimental Procedures

Our experiments reported here were conducted at Chuo University, Saitama University, Kyoto Sangyo University, and Toyo University in 1999. Our experiment basically follows the procedure adopted by Cooper et al. (1989, 1992) to create a one-shot, anonymous environment to the extent possible by carefully constructing a matching procedure.[11]

The experimental procedure is as follows. Twenty-six (13 in Session 2) subjects voluntarily participated in each session. When they entered one of two rooms,[12] they were assigned their subject number at random. Each subject was given an envelope in which written instructions, a recording sheet, and questionnaire were enclosed. To eliminate any experimenter effect, instructors other than the authors of the chapter read the instructions aloud and conducted the experiment manually. The instructors knew nothing about the equilibria of the games. The experiment proceeded according to the steps described below.

(1) In each round, the subjects were shown a payoff table of the game they would face in the current round and were privately told whether they were a sender or receiver. With whom they were matched throughout the session could not be identified. (2) Assignment to each subject of games and roles to play and who was matched with whom were randomly determined. (3) In each room, 12 out of 13 subjects made a decision in the experiment, while one subject waited until the next round.[13] (4) The sender was assigned one of two types, *A* or *B*, randomly

[11]In the earliest sessions, we also adopted a lottery reward procedure, first developed by Roth and Malouf (1979) and further extended by Berg et al. (1986), to induce a risk-neutral utility function from subjects. However, as we found no significant difference between lottery and ordinary payment conditions, we used the ordinary payment method in the rest of the session to simplify the instructions. Differences in experimental procedures between sessions are discussed in our previous paper (Kawagoe and Takizawa 1999). Those interested in the details may want to consult that paper.

[12]Before the session began, they were randomly divided into two groups of equal size for Sessions 1, 3, and 4. Separate rooms were assigned for both groups.

[13]This is because of the nature of the matching procedure we adopted. We devised random matching so that each subject plays both player roles and both sender types equally often, matched with a different subject in each round.

with a probability of 1/2. The sender type was only shown to the sender, and the receiver was unable to identify the sender's type before the payoffs for both subjects were determined. (5) The sender was told to choose between two messages, "I am type A" or "I am type B." (6) The receiver was shown the sender's message and told to choose one of three actions, X, Y, or Z.[14] (7) Payoffs for both players were determined by the sender's true type and the receiver's action according to the payoff tables. After all subjects had made a decision, the sender's true type, sent message, action taken by the receiver, and payoffs to both were revealed on a blackboard individually. (8) A session consisted of 13 rounds.[15] (9) Prior to the actual experiment, three rounds of practice experiment were conducted, where equilibria and payoffs of the games were different from those used in the actual experiment. Payoffs earned in this practice did not count toward the final reward calculation.

The above procedure was also explained in the written instructions. Instructions and practice took about half an hour and the session was about 2 h. The average reward for subjects was about 3,000 yen. This was approximately $28 at that time. The average wage for students at that time was 700 yen ($6.5) an hour.

3.2 Experimental Results

3.2.1 Aggregate Data

In this subsection, we examine the aggregate data of subjects' decisions. As our experiments were conducted in a one-shot, anonymous environment, we may regard all the data as independent samples. Table 16.4 shows the aggregate data of our experiment.

First, we consider which equilibrium play was more likely to be played in our games. Recall that by "babbling equilibrium play," we mean that the receiver chooses action Z on the path of play. Furthermore, by "separating play," we mean $((a, b), (X, Y))$ because there is a common language between players, and the messages used in our experiment, "I am type A" and "I am type B," have literal meaning. Figure 16.2 shows the aggregated frequency of choices by senders and receivers in each game.

As is clearly seen in Fig. 16.2, there is a clear tendency toward separating equilibrium play in Games 1 and 2. Thus, the type A (B) sender is inclined to reveal her true type by sending message a (b), and the receiver tends to believe such a message in responding by choosing X (Y) for message a (b). Here, it is somewhat

[14]See Footnote 4.

[15]In Session 1 and 2, we used a within-subject design in which each subject played Game 1, 2, and 3 randomly four times over 13 rounds. (In one round, they just waited, not participating in the game). In Session 3 and 4, we followed a between-subject design in which each subject played either Game 1 or 3 over 13 sessions. As we have noted, there was significant difference in the subjects' behaviors through these design changes.

Table 16.4 Aggregate data

		X	*Y*	*Z*	Total
Game 1					
A	***a***	99	0	13	112
	b	0	4	1	5
B	***a***	3	0	1	4
	b	1	92	20	113
Total		103	96	35	234
Game 2					
A	***a***	20	1	10	31
	b	1	5	2	8
B	***a***	3	1	0	4
	b	5	19	11	35
Total		29	26	23	78
Game 3					
A	***a***	45	9	43	97
	b	1	9	10	20
B	***a***	26	12	40	78
	b	6	19	14	39
Total		78	49	107	234

surprising to find this tendency even in Game 2, for which all the refinement theories predict babbling play. The frequency of babbling equilibrium play increased as the interests between sender and receiver diverged. However, a substantial number of receivers still believed the senders' messages to be truthful in Game 3, which has only babbling equilibria. Thus, our Hypothesis 1 is rejected regarding babbling equilibrium play in Games 2 and 3. These observations can be summarized as follows:

Result 1 *The majority of play in Games 1 and 2 was separating equilibria, and a notable proportion of play was separating in Game 3 even though it has only babbling equilibria.*

Next, to check Hypothesis 2, we consider senders' behavior concerning their truth-telling propensity. Recall that in Sect. 2.5, we define overcommunication as a sender's tendency to reveal her type more truthfully than is supposed from the equilibrium or refinement prediction. So, if the observed play in Games 2 and 3 is closer to separating equilibrium play than to the babbling equilibrium predicted by equilibrium refinement theories, overcommunication occurs.

The proportion of choices made by senders of type *A* (*B*) in each game is depicted on the left-hand side of Fig. 16.2. As one can clearly see, senders of type *A* have a clear tendency to tell the truth in every game, while the frequency of type *B* choosing message *b* decreases as the interests between sender and receiver become more conflicting. In fact, 95.7 % of type *A* senders chose message *a* in Game 1, 79.5 % in Game 2, and 82.9 % in Game 3, whereas 96.6 % of type *B* senders chose message *b* in Game 1, 89.7 % in Game 2, and 33.3 % in Game 3.

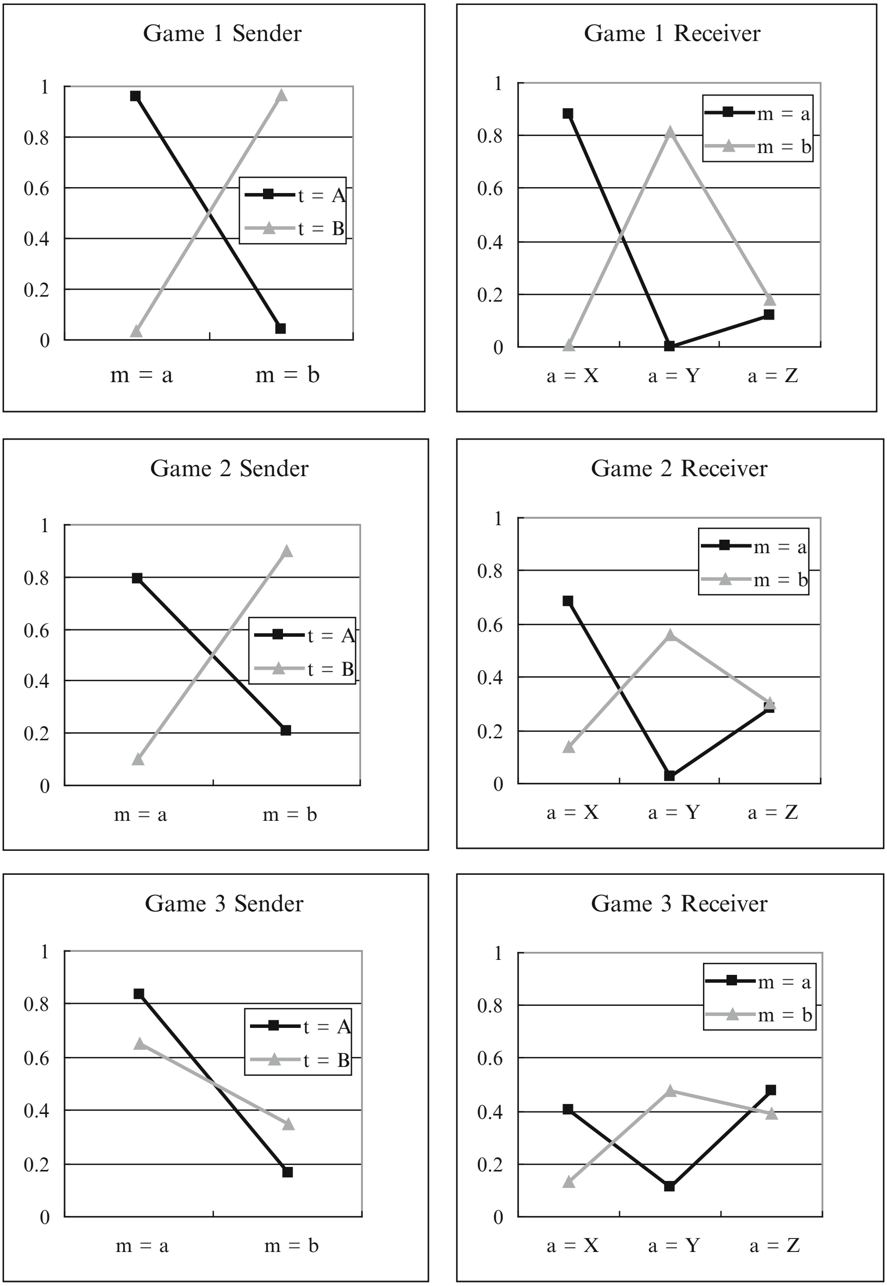


Fig. 16.2 Aggregate behaviors of senders and receivers in each game

Thus, while most type A senders revealed their type truthfully in all the games and the majority of type B senders told the truth in Games 1 and 2, type B senders in Game 3 showed a tendency to hide their identity and to deceive as if they were type A. In fact, 66.7 % of type B senders chose message a in Game 3. We find a significant difference in the frequency of type B choosing message b across games (χ^2 test, p $=$ 0.000), while the difference in the frequency of type A choosing message a across games is not significant (χ^2 test, p $=$ 0.679). But the fact that 33.3 % of type B senders revealed their type truthfully in Game 3 is not negligible. Thus, overcommunication is observed in Game 3 as well as in Game 2. These results are summarized as follows.

Result 2 *Overcommunication is observed in Games 2 and 3. Furthermore, the more aligned the interests are between sender and receiver, the more frequently the sender tells the truth.*

Next, to check Hypothesis 3, we consider receivers' behavior concerning their credulity or "truth-bias" in the terminology of communication theory. The right-hand panels in Fig. 16.2 show the proportion of actions, X, Y, and Z, chosen by receivers who received message a and b respectively. We can easily see that the receivers believed the senders' messages to be truthful more often in Games 1 and 2.

Specifically, 87.9 % of the receivers who received message a chose action X in Game 1, 65.7 % in Game 2, and 40.6 % in Game 3. On the other hand, 81.4 % of the receivers who received message b chose action Y in Game 1, 55.8 % in Game 2, and 47.5 % in Game 3. There is a significant difference in the frequency of action X for message a across games (χ^2 test, p $=$ 0.000). Differences in the frequency of action Y for message b across games are smaller, but significant at a 10 % level (χ^2 test, p $=$ 0.091).

Thus, almost all the receivers believed the senders' messages in Game 1. In Game 2, the majority of the receivers still believed the senders' messages, but the proportion of believers decreased compared with Game 1. In Game 3, the receivers still had a tendency to believe the senders' messages even though the game had only babbling equilibria. Thus, we have the following result:

Result 3 *Truth bias is observed. Furthermore, the more aligned the interests are between sender and receiver, the more frequently receiver believes the sender's message to be truthful.*

Recall that "truth-detection bias" is the tendency for a receiver to detect the sender's true type more correctly when she tells the truth than when she tells a lie. Our data clearly show the existence of such a bias. Here, we say that "the receiver detects truth" if the receiver chooses action X (Y) for message a (b) when the sender is type A (B) and "the receiver detects deception" if the receiver chooses action X (Y) for message b (a) when the sender is type A (B).

In Game 1, 84.9 % of the receivers succeeded in detecting the truth and deception. In Game 2, 59.1 % of the receivers detected the truth and 16.7 % detected deception.

In Game 3, 47.1 % of the receivers detected the truth and 13.3 % detected deception. There is a significant difference between the truth and lie detection rate in each game (χ^2 test, p = 0.006, p = 0.091, and p = 0.000 in Games 1, 2, and 3 respectively). This confirms the "truth-detection bias" stated in Hypothesis 4. We thus have the following:

Result 4 *The receiver guesses the sender's true type more correctly when the sender tells the truth than when she tells a lie. That is, truth-detection bias is observed.*

Thus, our experimental results reveal that subjects' behaviors in aggregate data deviate from the predictions given by equilibrium refinements, especially in Games 2 and 3. As stated in the Introduction, Blume et al. (2001) also report the experimental results of cheap-talk games with incomplete information. Although the games used in their experiment are different from ours, similar results are observed in misaligned interest cases as well as in aligned interest cases. We think this fact strengthens our results. Furthermore, since truth-telling is clearly distinguished from lying by using messages with literal meanings in our experiment, both truth bias and truth-detection bias are clearly identified in our data. As far as aggregate data are concerned, such deviations from standard equilibrium models seem to conform closely to the prediction made by level-k analysis. In the next subsection, we check whether individual data also conform to level-k analysis.

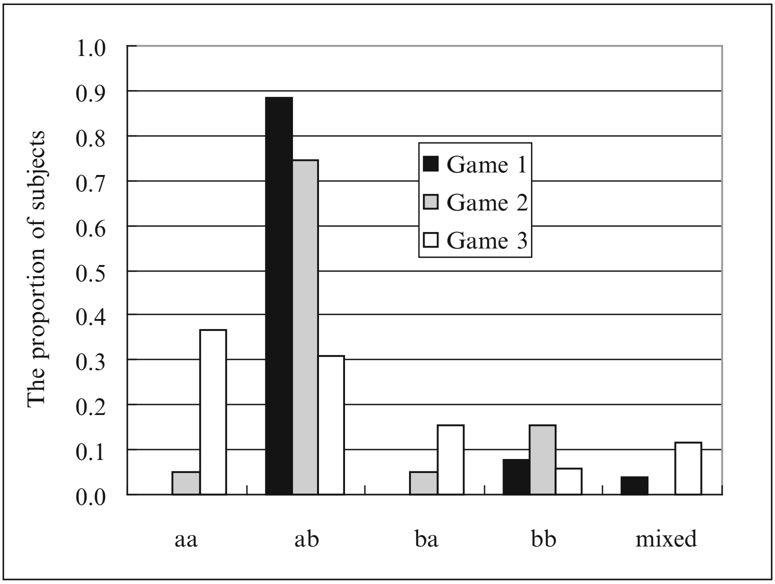
3.2.2 Individual Data

Using data from Sessions 1, 2, and 4,[16] we classified all subjects by behavioral type. For senders, there are five behavioral types: *aa*, *ab*, *ba*, *bb*, and mixed. For receivers, there are ten behavioral types: *XX*, *XY*, *XZ*, *YX*, *YY*, *YZ*, *ZX*, *ZY*, *ZZ*, and mixed. Here, for example, *ab* means a pure strategy where the sender chooses message *a* if type *A* and message *b* if type *B*, XY means a pure strategy where the receiver chooses action *X* for message *a* and action *Y* for message *b*, and the mixed type means that the player chooses more than one pure strategy at a particular information set. In total, 52 subjects in Games 1 and 3 and 39 subjects in Game 2 were classified accordingly. Figure 16.3 shows the relative frequency of those behavioral types observed in our experiment.

Out of the five possible behavioral types, the *ab* type comprised the majority of senders in Games 1 and 2 (88.5 % and 74.4 % respectively) and was still the second majority in Game 3 (30.8 %). In Game 3, the *aa* type was the majority (36.5 %). The mixed type was rarely observed in any game (3.8 % in Game 1, 0.0 % in Game 2, and 11.5 % in Game 3). From these results, the majority of senders in every game belonged to a higher level of rationality, as assumed in level-k analysis (Lk ($k \geq 1$) in Games 1 and 2, and $L2$ or $L3$ in Game 3).

[16]Unfortunately, individual data in Session 3 conducted at Kyoto Sangyo University were lost.

Senders



Receivers

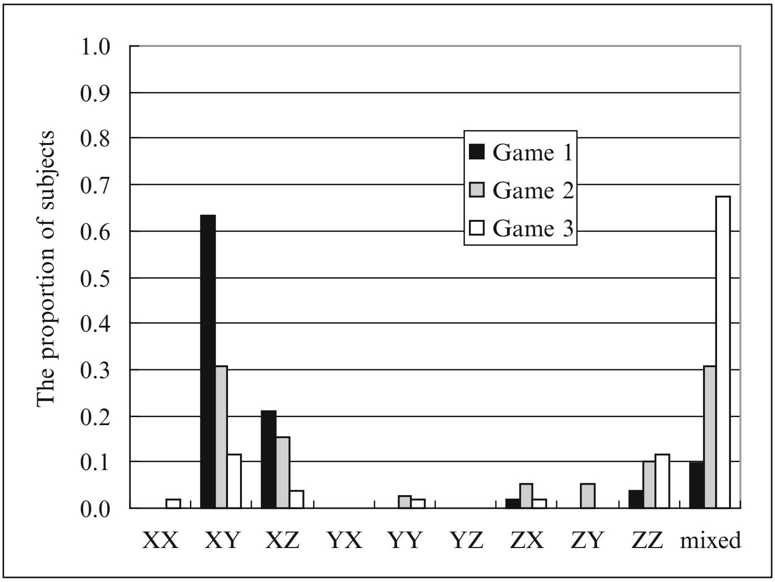


Fig. 16.3 The distribution of each behavioral type

As for the receivers, out of the nine possible pure-strategy behavioral types, *XY* comprised the majority in Games 1 and 2 (63.5 % and 23.1 % respectively), while both *XY* and *ZZ* were the majority in Game 3 (11.5 % for each). While the mixed type was rarely observed in Game 1 (9.6 %), it won the majority in Games 2 and 3 (23.1 % and 67.3 % respectively). Thus, while the receivers in Game 1 belonged to rationality level $k \geq 1$, a substantial amount of $L0$ play was observed in Games 2 and 3.

In Sect. 2.3, we mentioned two other possible specifications of the $L0$ type. The first specification was to let the $L0$ receiver be a credulous type, *XY*. As stated before, this specification basically gives the same prediction as that obtained under the assumption we adopt here. However, it cannot explain the existence of many of the mixed types found on the receiver's side in Games 2 and 3. The other possibility was to let both the $L0$ sender and $L0$ receiver be the randomizer. This specification cannot explain play different from babbling equilibrium play. So, we conclude that our specification of the $L0$ type best fits our data across games.

Thus, these individual data reveal that level-k analysis can explain our experimental data better than any other theories. In particular, that a substantial portion of the senders' behaviors belonged to the *ab* type in Games 2 and 3 and that a non-ignorable portion of the receivers belonged to the *XY* type in Games 2 and 3 are clearly deviations from the predictions given by equilibrium refinements. These facts can be explained consistently only by level-k analysis.

4 Conclusion

The present chapter has reported the experimental results of cheap-talk games with incomplete information. The cheap-talk games we examined have varying degrees of alignment of interests between sender and receiver. The experimental environment is designed to be as simple as possible and a common language is shared between sender and receiver. In these settings, we find that the less aligned the interests are between sender and receiver, the more frequently babbling equilibrium play is observed. However, subjects showed a notable tendency toward separating play even in games with misaligned interests. We then compared the explanatory power of various equilibrium refinement theories with that of level-k analysis. We found that while refinement theories only work in the case of aligned interests, level-k analysis works well in conflicting interest cases as well as in aligned interest cases.

Our experimental results can also be compared with research results accumulated in communication theory. Specifically, our results confirm the existence of "truth bias" and "truth-detection bias." However, it should be stressed that our experimental environments are quite different from those used in communication theory (e.g., McCornack and Parks 1986). First, our experiment was conducted in a one-shot anonymous environment with no room for relational development. Second, no nonverbal cues were available to the receiver, and the sender was restricted to using the simplest possible messages. As truth bias was observed in this environment without any relational development and with the conflicting situation being common knowledge, truth bias was confirmed in a very strong sense.

Thus, our experimental results suggest that truth-telling and truth-guessing are more intrinsic to human communication than is supposed in game theory. However, why this is so is still to be explored. Cai and Wang (2006) and others seem to share the view that bounded rationality gives a good explanation. Sánchez-Pagés and Vorsatz (2007) argue that some individuals may take truth-telling as a social norm. Both elements may be in play. As we have seen, our data are well explained by the level-k model where the $L0$ sender is the truth-teller. This means that some focal or social norm effects are in play to drive boundedly rational players to behave accordingly. Furthermore, the partial presence of boundedly rational or norm-driven players may induce even rational players to mimic boundedly rational play, as Crawford (2003) shows in a complete information game and Kawagoe and Takizawa (2005) show in an incomplete information game.

For the nature of communication, Grice (1989) submits the "cooperative principle" where communication is an attempt to determine truth value through statements exchanged in conversation and where conversation is thus intrinsically a cooperative task.[17] Based on this view, McCornack (1992) sees deception as a violation of one or more maxims of the "cooperative principle." Three major categories of deceptive message, falsification, concealment, and equivocation, are thus identified. Exploring strategic problems by distinguishing among deceptive messages from a game theoretic viewpoint will enrich the study of information transmission in game theory. We intend to address such issues in future research.

Acknowledgments We are especially grateful to Vincent Crawford for his abundant advice and suggestions for revising an earlier version of our paper. We would also like to thank Masahiko Aoki, Bram Cadsby, Colin Camerer, Rémy Juston Coumat, Håkan Holm, Akihiko Matsui, Richard McKelvey, Masahiro Okuno-Fujiwara, Charles Plott, Tatsuyoshi Saijo, Fei Song, and Takehiko Yamato for their helpful comments and encouragement. Part of this research was supported by a grant-in-aid of the Japan Society for the Promotion of Science project for the first 3 years since 1998 in collaboration with Yuji Aruka and Sobei Oda.

Appendix. Instructions

This is an experiment on economic decision making. You can earn some amount of money in cash in this experiment, if you make appropriate choices according to what is explained below.

In this experiment, each group consists of two persons, one of whom we call "*S*-player" and the other "*R*-player." Scores of both players are determined by choices

[17] Rubinstein (2000) and Glazor and Rubinstein (2006) adopt the argument of Grice's pragmatics (1989) in a game theoretical framework and attempt to show that persuasion is profitable and successful.

of both players. We will not inform you who are "*S*-players (*R*-players)" or who are matched with whom at each round. Matching players are determined at random at each round. In each round, one of you has to "wait" and do nothing until the next round.

We repeat such an experimental round several times. When all the rounds finish, the instructors will tell you the end of experiment. Your reward is finally determined based on the score you earned all over the rounds. More detailed experimental procedure follows.

A.1. Experimental Procedure

In this experiment, each round proceeds as follows:

1. Each of you are told whether you are an "*S*-player" or an "*R*-player" at this round.
2. If you are an "*S*-player," you are also told whether you are type *A* or type *B* at this round.
3. "*S*-player" chooses between two alternatives "I am a type *A*" or "I am a type *B*."
4. "*R*-player," informed of the choice of "*S*-player" who is your matched opponent, chooses from among three alternatives "A," "B," and "C."
5. The score is determined according to the type of "*S*-player," which is assigned at the beginning of this round, and the choice by "*R*-player".
6. The final reward is determined based on the score you earned all over the rounds, and then paid in cash.

Let us see the details of each stage more closely.

Step1 Each pair of subjects participates in each decision making, so there are 6 pairs and 1 person has to wait. One subject of a pair is called "*S*-player," while the other subject "*R*-player." Throughout the experiment, you are never told who and who match to form a pair. All that you are told is the number assigned to the pair to which you belong and whether you are an "*S*-player" or "*R*-player." All of these are predetermined according to some random matching rule by the experimenters.

More specifically, at each round a "Payoff table" is distributed to each of those who participate in the experiment. On the table, you will find a payoff table and the number assigned to the pair to which you are belonging at this round. We will later explain how to read the payoff table in more detail. If you are an "*S*-player," "Answer sheet" will also be distributed.

Fill in the blank of your "Recording sheet" with the number of your pair that you have found on the "Answer sheet." Circle the letter "S" in the Player field of your "Recording sheet" if you are an "S-player," "R" if "R-player."

If you are told to wait at this round, write "wait" in the Pair field of your "Recording sheet," and wait silently until the next round.

Step2 Look at the upper half of your "Answer sheet." If you are told to be an "S-player" in Stage 1, you are also told whether you are type *A* or type *B*. Throughout the experiment, the probabilities of being type A and type B are equal. No one except you knows whether you are type A or type B.

If you are an "S-player" and your type is A, circle the letter "A" in the Type field of your "Recording sheet," likewise for the case that your type is B.

Step3 Those who are told to be an "S-player" in the 1st stage choose between "Alternative A" or "Alternative B."

Alternative A: "I am a type A."
Alternative B: "I am a type B."

The choice is completely up to you. While the type of which you are informed in the second stage will not be known to the matched "R-player," the choice you made in the second stage will be known to the opponent.

If you choose "Alternative A," circle the letter A in the Alternative field on your "Recording sheet," likewise for the case that you choose "Alternative B." Also do the same for the "Choice of S-player" field in the lower half of your "Answer sheet" and hand it to the instructors.

Step4 "R-player" chooses among "Alternative A," "Alternative B," and "Alternative C" knowing the choice made by "S-player" at stage 3. You can find the choice of the matched "S-player" on the "Answer sheet."

If you choose "Alternative A," circle the letter A in the Alternative field on your "Recording sheet," likewise for the case that you choose "Alternative B" or "Alternative C." Also do the same for the "Choice of R-player" field on the "Answer sheet" handed to you.

Step5 Both players' scores are determined according to the choice made by "R-player" in stage 4 and *the type revealed to "S-player" in the second stage. Note that the choice by "S-player" in the third stage does not affect scores.*

The score table shows you how both players' scores are determined. The scores that both players get will be shown on the blackboard, so ensure your score at each round. After ensuring your score, write it in the Score field on your "Recording sheet."

Example Suppose you are distributed a payoff table as follows:

	type A		type B	
Alternative A	S	R	S	R
	90	20	60	30
Alternative B	S	R	S	R
	50	10	10	90
Alternative C	S	R	S	R
	80	70	30	50

If "*S*-player" is assigned type *A* in stage 2, look down under the column "type *A*" on this table. If "*S*-player" is assigned type *B*, then look down under the column "type *B*." The left digit in each cell indicates *S*-player's score and the right R-player's.

For example, suppose "*S*-player" is told that his type is type *A* and "*R*-player's" choice is "Alternative A," then "*S*-player" gets 90 and "*R*-player" gets 20 according to this payoff table. If "*S*-player" is told that his type is type *B* and "*R*-player's" choice is "Alternative B," then "*S*-player" gets 10 and "*R*-player" gets 90.

Also suppose that "*S*-player" is told that his type is *A* and "*S*-player" chooses "Alternative B." In this case, if "*R*-player" chooses "Alternative A," then "*S*-player" gets 90 and "*R*-player" gets 20. Next suppose that "*S*-player" is told that his type is *B* and "*S*-player" chooses "Alternative B." In this case, if "*R*-player" chooses "Alternative A," then "*S*-player" gets 60 and "*R*-player" gets 30.

Step6 Stages 1–5 complete a round of the experiment. Your reward in cash in this round is 50 Yen times the score you get in this session. Fill in the Reward field on your "Recording sheet" with the number that is 50 times as large as the score in this round. The total reward in the experiment is the sum of each round's reward plus participation fee, a 1,000 Yen.

A.2. Notices

Please be quiet throughout the experiment. You might be expelled if the instructor thinks it necessary. In that case, you might not be rewarded.
You cannot leave the room throughout the experiment in principle.
Please turn off your pocket bell or cellular phone.
Do not take anything used in the experiment with you.

A.3. Questions

If you have any question concerning the procedure of experiment, raise your hand quietly. An instructor will answer your question in person. In some cases, the content of your question might disallow the instructor to answer it, however.

A.4. Practice

Before conducting the experiment, we have three sessions for practice. These are purely for practice and the results therein will not be counted in your reward. You can always refer to this instruction throughout the experiment.

Please take out "Recording sheet (Practice)" from your envelope and fill in your name and student ID.

We will distribute "Answer sheets (Practice)" and "Score table (Practice)" to those who are to be "*S*-players" in this session. To those who are to be "*R*-players" in this session, only "Score table (Practice)" will be distributed.

"*S*-players" should now circle the letter *S* in the Player field of the "Recording sheet (Practice)" and "*R*-players" the letter *R*.

"*S*-players" now make their choice looking at your own type on the "Answer sheet (Practice)" and the "Payoff table (Practice)." Mark your own type in the Type field of your "Recording sheet (Practice)" and also mark your choice in the Choice field of the "Recording sheet (Practice)." Next mark your choice on the "Answer sheet (Practice)" too. "Answer sheet (Practice)" will be collected later.

Then the lower half of the "Answer sheet (Practice)," on which "*S*-players" have already marked their choices, will be distributed to the matched "*R*-players." "*R*-players" can thus see the choice of "*S*-players," but not their true types. "*R*-players" should now make choice by examining the score table and mark your choice in the Choice field of your "Recording sheet (Practice)." Also mark your choice on the "Answer sheet (Practice)."

Let us now turn to actual experiment. Please fill in your name and student ID on your "Recording sheet."

Addendum: Recent Developments[18]

Since the publication of our paper (Kawagoe and Takizawa 2009), we have seen growing interests in the study of communication between players with conflicting interests. To name only a few, Kartik (2009) theoretically considers a related model of aversion to lying in the context of strategic information transmission between an informed sender and an uninformed receiver. Battigalli et al. (2013) provides an account for the data in Gneezy (2005) experiment on deception based on guilt aversion.

Contribution of our research to the literature is twofold. One is that it is the first paper to report "truth bias" in an economic experiment and to give it a theoretical explanation. The other is that this research is among the first to apply the level-k model to an extensive-form game with incomplete information, thereby contributing to the development of the level-k analysis. Let us look at these points in turn.

Truth bias is a tendency of receiver to believe the truthfulness of the sender's message, a term coined in the communication theory (McCornack and Parks 1986). In contrast to the experiments that had hitherto reported truth bias, our experiment showed that the bias would persist even when the structure of a game was common knowledge. Kawagoe and Holm (2010) also confirmed truth bias in sender-receiver type cheap-talk game with "zero-sum" payoff, where the experiment was conducted using playing cards in face-to-face environment. Even in this competitive

[18]This addendum has been newly written for this book chapter.

environment, they found the tendency of truth bias. These strong results attracted attention of some psychologists. For example, Robert Feldman (2010), a specialist in the study of deception, devoted a whole book to the phenomenon of truth-bias and mentions our research at length in Chap. 2 of his book.

To the best of our knowledge, no theoretical account for the "truth bias" had been provided before us. Our explanation based on level-k analysis was very simple and had a good fit with the data. Taking a plausible L0 type, we constructed upper levels by assuming that the level-k player best responds to the level-(k-1) player, up to level-2.

One may suppose that our model as well as most papers based on level-k analysis crucially depend on the assumption that players of level-k "irrationally" believe that all the other players are of level-(k-1). This question can possibly be important and we considered it in the working paper version of our paper (Kawagoe and Takizawa 2005). There we actually analyzed a level-k model with sophisticated players as in Crawford (2003). The sophisticated player in the game anticipates that he/she faces a population consisting of sophisticated players as well as L0, L1, L2 players, and then assesses subjective belief about the distribution of these types in a way that it is consistent with his/her equilibrium strategy. We showed that truth bias could be a sequential equilibrium even in this game, which means that truth bias can be "rationally" explained.

The second contribution of our paper to the literature concerns the problem of how to apply the level-k model to an extensive-form game like the cheap-talk game. Before our paper, level-k models had been applied to many games, and had succeeded in explaining a number of anomalous behaviors found in the laboratory. However, most of them had been applied to normal-form games. This research is among the first that tried to apply the level-k model to an extensive-form game with incomplete information. There are several questions involved in so doing.

The first question is what strategy should be assumed for L0 players. Randomizing over all pure strategies with uniform distribution had been routinely used in the literature. However, it is not obvious that complete randomization works well in extensive-form games. This problem is very important, because the strategy of L0 player works as an anchor of all the other upper-level players in the level-k analysis.

The second question concerns the problem that arises in reasonably defining player's responses in off-the-play paths in extensive-form games. Constructing level-k strategy requires us somehow to deal with this problem.

Concerning the first problem, Kawagoe and Takizawa (2009) examined two models of the L0 strategy. Later, Kawagoe and Takizawa (2012) extensively studied which L0 assumption works well in explaining controversial centipede game experiments. For the second problem, we made the model probabilistic by introducing the same noise structure in logit form as in the quantal response equilibrium (QRE, McKelvey and Palfrey 1998), which also enabled us to estimate parameters through maximum likelihood method.

Recent studies of communication in games focus on the communication with noisy channel. For example, Blume et al. (2007) study a version of cheap-talk game with noisy channels. In their setting, a message sent by a sender might be changed to the one that the sender doesn't intend to send. With a slightly different motivation, the implications of costly efforts in sending and receiving a message for successful communication are analyzed by Dewatripont and Tirole (2005).

Investigating how truthful communication arises in the environment with possible "misinterpretation" of messages is important. Even in that environment, do we still observe truth bias? This is probably our next research question in this line of study.

References

Banks J, Camerer C, Porter D (1994) An experimental analysis of Nash refinements in signaling games. Game Econ Behav 6:1–31

Battigalli P, Charness G, Dufwenberg M (2013) Deception: the role of guilt. J Econ Behav Organ 93:227–232

Berg JE, Daley LA, Dickhaut JW, O'Brien JR (1986) Controlling preferences for lotteries on units of experimental exchange. Q J Econ 101:281–306

Blume A, DeJong DV, Kim YG, Sprinkle GB (1998) Experimental evidence on the evolution of meaning of messages in sender-receiver games. Am Econ Rev 88:1323–1340

Blume A, DeJong DV, Kim YG, Sprinkle GB (2001) Evolution of communication with partial common interest. Game Econ Behav 37:79–120

Blume A, Board OJ, Kawamura K (2007) Noisy talk. Theor Econ 2:395–440

Brandts J, Holt CA (1992) An experimental test of equilibrium dominance in signaling games. Am Econ Rev 82:1350–1365

Brandts J, Holt CA (1993) Adjustment patterns and equilibrium selection in experimental signaling games. Int J Game Theory 22:279–302

Burgoon JK, Buller DB, Ebesu AS, Rockwell P (1994) Interpersonal deception: V. Accuracy in deception detection. Commun Monogr 61:303–325

Cadsby CB, Frank M, Maksimovic V (1998) Equilibrium dominance in experimental financial markets. Rev Financ Stud 11:189–232

Cai H, Wang JT-Y (2007) Overcommunication in strategic information transmission games. Game Econ Behav 56:7–36

Camerer C (2003) Behavioral game theory: experiments on strategic interaction. Princeton University Press, Princeton

Camerer C, Ho T-H, Chong JK (2004) A cognitive hierarchy theory of one-shot games. Q J Econ 119:861–898

Cho I-K, Kreps D (1987) Signaling games and stable equilibria. Q J Econ 102:179–221

Cooper R, DeJong DV, Forsythe R, Ross TW (1989) Communication in the battle of sexes game. Rand J Econ 20:568–587

Cooper R, DeJong DV, Forsythe R, Ross TW (1992) Communication in coordination games. Q J Econ 107:739–771

Costa-Gomes M, Crawford V (2006) Cognition and behavior in two-person guessing games: an experimental study. Am Econ Rev 96:1737–1768

Costa-Gomes M, Crawford V, Broseta B (2001) Cognition and behavior in normal form games: an experimental study. Econometrica 69:1193–1235

Crawford V (1998) A survey of experiments on communication via cheap talk. J Econ Theory 78:286–298

Crawford V (2003) Lying for strategic advantages: rational and boundedly rational misrepresentations of intentions. Am Econ Rev 93:133–149

Crawford V (2007) Let's talk it over: coordination via preplay communication with level-k thinking. Paper presented at 26th Arne Ryde symposium on communication in games and experiments, , Lund, 24–25 Aug 2007

Crawford V, Iriberri N (2007a) Fatal attraction: salience, naivete, and sophistication in experimental hide-and-seek games. Am Econ Rev 97:1731–1750

Crawford V, Iriberri N (2007b) Level-k auctions: can a non-equilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions? Econometrica 75:1721–1770

Crawford V, Sobel J (1982) Strategic information transmission. Econometrica 50:1431–1451

Dewatripont M, Tirole J (2005) Modes of communication. J Polit Econ 113:1217–1238

Dickhaut JW, McCabe KA, Mukherji A (1995) An experimental study of strategic information transmission. Econ Theory 6:389–403

Ellingsen T, Östling R (2006) Organizational structure as the channeling of boundedly rational pre-play communication. SSE/EFI working paper series in economics and finance, no 634

Farrell J (1987) Cheap talk, coordination, and entry. Rand J Econ 18:34–39

Farrell J (1988) Communication, coordination and Nash equilibrium. Econ Lett 27:209–214

Farrell J (1993) Meaning and credibility in cheap talk games. Game Econ Behav 5:514–531

Feldman R (2010) The liar in your life: the way to truthful relationship. Twelve, New York

Glazor J, Rubinstein A (2006) A study in the pragmatics of persuasion: a game theoretical approach. Theor Econ 1:395–410

Gneezy U (2005) Deception: the role of consequences. Am Econ Rev 95:384–394

Grice P (1989) Studies in the way of words. Harvard University Press, Cambridge, MA

Ho T-H, Camerer C, Weigelt K (1998) Iterated dominance and iterated best response in experimental 'p-beauty contests'. Am Econ Rev 39:649–660

Holm H (2010) Truth and lie detection in bluffing. J Econ Behav Organ 76:318–324

Holm H, Kawagoe T (2010) Face-to-face lying - an experimental study in Sweden and Japan. J Econ Psychol 31:310–321

Kartik N (2009) Strategic communication with lying costs. Rev Econ Stud 76(4):1359–1395

Kawagoe T, Takizawa H (1999) Instability of babbling equilibrium in cheap talk games. Saitama University, Saitama

Kawagoe T, Takizawa H (2005) Why lying pays: truth bias in the communication with conflicting interests. SSRN electronic paper collection, 691641

Kawagoe T, Takizawa H (2009) Equilibrium refinement vs. level-k analysis: an experimental study of cheap-talk games with private information. Game Econ Behav 66:238–255

Kawagoe T, Takizawa H (2012) Level-k analysis of experimental centipede games. J Econ Behav Organ 82:548–566

Matthews S, Okuno-Fujiwara M, Postlewaite A (1991) Refining cheap-talk equilibria. J Econ Theory 55:247–273

McCornack S (1992) Information manipulation theory. Commun Monogr 59:1–16

McCornack S, Parks M (1986) Deception detection and relationship development: the other side of trust. In: McLaughlin M (ed) Communication yearbook, vol 9. Sage, Beverly Hills

McKelvey RD, Palfrey TR (1998) Quantal response equilibria for extensive form games. Exp Econ 1:9–41

Nagel R (1995) Unraveling in guessing games: an experimental study. Am Econ Rev 85:1313–1326

Rabin M, Sobel J (1996) Deviations, dynamics and equilibrium refinement. J Econ Theory 68:1–25

Roth AE, Malouf WK (1979) Game-theoretic models and the role of bargaining. Psychol Rev 86:574–594

Rubinstein A (2000) Economics and language. Cambridge University Press, Cambridge

Sánchez-Pagés S, Vorsatz M (2007) An experimental study of truth-telling in a sender-receiver game. Game Econ Behav 61:86–112

Stahl D, Wilson P (1995) On player's models of other players: theory and experimental evidence. Game Econ Behav 10:218–254

Vrij A (2000) Detecting lies and deceit. Wiley, West Sussex

Wang JT-Y, Spezio M, Camerer CF (2010) Pinocchio's pupil: using eyetracking and pupil dilation to understand truth-telling and deception in games. Am Econ Rev 100:984–1007

Yamamori T, Kato K, Kawagoe T, Matsui A (2008) Voice matters in a dictator game. Exp Econ 11:336–343

Part VI
Behavioral Contract Theory

Chapter 17
Moral Hazard and Other-Regarding Preferences

Hideshi Itoh

Abstract This chapter aims at obtaining new theoretical insights by combining the standard moral hazard models of principal-agent relationships with theories of other-regarding preferences, in particular inequity aversion theory. The principal is in general worse off, as the agent cares more about the wellbeing of the principal. When there are multiple symmetric agents who care about each other's wellbeing, the principal can optimally exploit their other-regarding nature by designing an appropriate interdependent contract such as a "fair" team contract or a relative performance contract. The approach taken in this chapter can shed light on issues on endogenous preferences within organizations.

Keywords Behavioral contract theory • Inequity aversion • Moral hazard

1 Introduction

It is standard in economic analysis to assume that people maximize their wealth and other personal material consumption. This self-interest hypothesis has proved to be correct in many situations and highly useful in the analysis of diverse problems. However, people, including some economists, have recognized that *not all* people are motivated *exclusively* by self-interest. We often care about the wellbeing of others, and find ourselves behaving altruistically, worrying about unfair distribution of wealth, reciprocating the kind (or unkind) behavior of others, and so on. It is these sorts of *other-regarding* behavior that are the concern of this paper.

Recent data from experiments on ultimatum games, gift exchange games, public goods games, trust games, and so on demonstrate that people in fact deviate from

---

The original article first appeared in Japanese Economic Review 55:18–45, 2004. A newly written addendum has been added to this book chapter.

H. Itoh (✉)
Hitotsubashi University, Kunitachi, Tokyo, Japan
e-mail: hideshi.itoh@r.hit-u.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_17

self-interest in systematic ways.[1] "Among experimentalists—and others paying attention to the evidence—the debate over whether there are systematic, non-negligible departures from self-interest is over" (Rabin 2002, p. 666). In particular, the data from a large number of ultimatum game experiments "falsify the assumption that players maximize their own payoffs as clearly as experimental data can. Every methodological explanation you can think of (such as low stakes) has been carefully tested and cannot fully explain the results" (Camerer 2003, p. 43).

From the basis of experimental evidence, researchers in behavioral economics and behavioral game theory have recently developed new theoretical models that deviate from the exclusive self-interest hypothesis by incorporating other-regarding features such as fairness, equity and reciprocity into the framework of the standard utility maximization. Behavioral economics is "an approach to economics which uses psychological regularity to suggest ways to weaken rationality assumptions and extend theory" (Camerer 2003, p. 3). Some of the new models capture other-regarding behavior, and succeed in explaining various experimental results.

The purpose of this chapter is to go one step further. I apply those theories of other-regarding preferences that can explain many experimental results to the standard models of principal-agent relationships with moral hazard, in order to generate new theoretical insights. The traditional literature in contract theory is also based on the self-interest hypothesis. The agent attempts to maximize a function that is increasing in his wealth and other private benefits and is decreasing in the private cost of his action. The principal aims at maximizing the benefit generated by the agent minus her payments to him.[2]

However, a relaxing of the self-interest hypothesis may be particularly important for contract theory. In the first place, the aim of contract theory is to design appropriate incentives, and the way in which people care about others' wellbeing as well as their own is crucial for incentive design. The optimal contract for the self-interested agent may be very different from that for the other-regarding agent.

Second, contract theory offers a major analytical framework for organizational economics (Milgrom and Roberts 1992) and for personnel economics and human resource management (Baron and Kreps 1999; Lazear 1995). Even in these applied fields, the dominant view is that "many institutions and business practices are designed as if people were entirely motivated by narrow, selfish concerns and were quite clever and largely unprincipled in their pursuit of their goals" (Milgrom and

[1] See Camerer (2003) and Fehr and Schmidt (2003) for surveys.

[2] Even in the traditional contract-theoretic models, other-regarding behavior sometimes plays an implicit but important role. In the collusion literature (Holmstrom and Milgrom 1990; Itoh 1993; Tirole 1992), the enforceable side-contracts assumption may reflect other-regarding behavior between agents implicitly. Some of the important results from the multitask analysis (Holmstrom and Milgrom 1991) hinge on the assumption that the agent chooses some positive amount of effort in the absence of incentives. This feature may be explained by task-specific intrinsic motivation, or the agent's caring about the wellbeing of the principal.

Roberts 1992, p. 42). However, organizational researchers in other disciplines have criticized such a view over the years (Perrow 1986; Pfeffer 1997). For example, Perrow (1986) argues that

> there is no innate tendency to either self- or other-regarding behavior in people; either can be evoked depending on the structure. Agency theorists examine the structures favored by capitalism and bureaucracy and find much self-regarding behavior; they then assume that this is human nature. They neglect the enormous amount of neutral and other-regarding behavior that exists (and must, for organizations even to function) and the structures that might increase it. What they take for granted, should be taken as a problem. (Perrow 1986, pp. 234–235)

My optimistic standpoint is that developing "behavioral contract theory" that uses experimental/field evidence and psychological intuition to generalize the standard assumptions, may help to encourage more productive interactions among the students of organizations from various disciplines. As a start, this chapter analyzes how incentives are affected by other-regarding behavior in the stylized principal-agent relationships with moral hazard, and the kinds of preferences that are desirable for the principal.

Third, very few experiments on contract choice by a principal who cannot observe an agent's action have so far been conducted.[3] The establishment of theoretical results on how incentives and other-regarding behavior interact in principal-agent relationships will help promote experimental tests of contract theory in future research.

To study the effects of other-regarding preferences on moral hazard, I use mainly the model of *inequity aversion* developed by Fehr and Schmidt (1999).[4] Their model is an example of the distributional approach to other-regarding preferences in which only the final monetary distribution among players matters. In particular, a player feels guilty when his material payoff is above others' payoffs, while he feels envious when his material payoff is below those of others. In other words, he dislikes either being ahead or being behind. The model is simple but can explain various experimental results and capture fair-minded and reciprocal behavior very well.[5] Furthermore, to cover broader cases of other-regarding preferences, I follow Neilson and Stowe (2003) to extend Fehr and Schmidt (1999) by allowing the player to be *competitive* or *status-seeking*, in the sense that he dislikes being behind, but loves being ahead.

The benchmark moral hazard model is a standard one. A principal hires an agent for a project. The project either succeeds or fails, and the probability distribution

---

[3]Exceptions include Fehr et al. (2001), Güth et al. (1998), and Keser and Willinger (2000), all of which study trade between one principal and one agent. Nalbantian and Schotter (1997) present an interesting experimental examination of a variety of group incentives, although incentive plans are exogenously chosen by the experimenters.

[4]Bolton and Ockenfels (2000) present an alternative formulation of inequity aversion.

[5]Other kinds of distributional preferences as well as an alternative approach to other-regarding preferences are summarized in Sect. 2.

depends on the agent's action. I assume that both principal and agent are risk neutral while the agent is wealth constrained, and hence the contract must satisfy the limited liability constraints. To induce the agent to choose the more productive action, the principal must offer high-powered incentives through a higher rate of pay upon the success of the project. However, higher-powered incentives are more costly to the principal because she has to pass more rents on to the agent.[6]

I first assume that the agent compares his income with that of the principal. This setting corresponds to ultimatum games, gift exchange games, and trust games, the experimental results of which are known to be consistent with the hypothesis that the player in the role of the receiver (agent) cares about the wellbeing of the player in the role of the proposer (principal), although the proposer is not allowed to offer complicated contracts. And Bewley (1999), who sought to learn why wages and salaries seldom fell during recessions through interviews covering more than 300 people, finds some evidence that workers feel that the performance of the company does not justify workers' pay cuts.

Other-regarding preferences interact with moral hazard in nontrivial ways. First, a concern for fairness does not resolve moral hazard: if the incentive compatibility constraint is binding for the self-interested agent, it is also binding for the other-regarding agent. Furthermore, in the "standard" case where the agent's income is below that of the principal, the principal is typically worse off, the more other-regarding the agent is. The logic is simple. Since the agent deplores inequity, the more inequity averse the agent is, the more the principal must pay the agent upon the success of the project in order to satisfy his incentive compatibility constraint.

Although we might reasonably assume that at optimum the agent's income is still lower than that of the principal, we can instead assume the nonstandard situation in which the agent will receive more income than the principal upon conclusion of a successful project. Even in this case, however, the principal will not benefit from inequity aversion, because the inequity-averse agent would be altruistic towards the principal and would want to reduce the probability of success (that is, the probability of his being ahead). However, I show that if the agent has status-seeking preferences so that he enjoys being ahead, the principal is better off the more other-regarding (more competitive) the agent is: the more status-seeking the agent is, the more intrinsically motivated he will be and hence the more monetary incentives the principal can save.

The results obtained are mostly robust to changing specification of other-regarding preferences, so that the agent compares his income net of costs of action, with the principal's wealth, or the principal as well as the agent is other-regarding.

In my view, it is more important to analyze the effects of other-regarding preferences in a multi-agent situation. A large body of sociological literature argues

---

[6]In other words, the *ex ante* participation constraint does not bind. If I did not impose the limited liability constraints, the participation constraint would usually bind, and hence the effects of other-regarding preferences on incentives, which are the focus of the chapter, would be somehow undermined.

that people are most likely to compare themselves to others who are similar in terms of personal characteristics (e.g. age, gender, education), and situations within organization (e.g. job titles, departments, entry cohorts): while an agent may care about the principal's wellbeing, he is likely to care even more about that of others agents. For example, pay attached to job titles and pay differences across job titles are more likely to be perceived as acceptable by employees than pay differences within a single job title (Baron 1988). That is, if multiple candidates for reference actors exist, then "lateral" comparison is more likely to dominate. I thus extend the model to a multi-agent setting, assuming that each agent cares about what he and the other agents are paid while the principal is not his reference actor.

I show that, although I do not assume any technological or stochastic interdependence, the principal can optimally exploit the agents' other-regarding nature by designing an appropriate interdependent contract. Other-regarding preferences lead to the possibility that a team contract (whereby each agent is paid more if the other's project succeeds than if it fails) or a relative performance contract (whereby the agent is paid more if the other project fails than if it succeeds) becomes optimal. I describe the conditions necessary for a team contract and a relative performance contract each to be optimal, and examine how the principal's payoff changes with the extent to which the agents are other-regarding. The "extreme" team contract, under which each agent is paid a positive amount only when both projects succeed, includes the important feature that no agent feels guilty or envious, so that the principal is indifferent concerning how other-regarding the agents are. However, the "extreme" relative performance contract, under which each agent is paid a positive amount only if his project succeeds and the other's fails, becomes optimal for status-seeking agents, or inequity-averse agents who do not feel much guilty about being ahead. The reason such a contract becomes optimal is in contrast to that for the team contract. The relative performance contract leads to great inequality in payments when the performance outcomes differ among agents, which creates strong incentives. This incentive effect is stronger the more other-regarding the agents are, and hence the principal benefits more from the more competitive agents.

The optimal contract for self-interested agents changes drastically when a small degree of other-regarding preferences is introduced. Under the technological assumptions of the model, if the agents are self-interested there is an optimal independent contract in which the payment scheme for each agent depends only on the outcome of his project. However, when the agents become other-regarding, then, however small the changes are, no independent contract remains optimal, and the optimal contract becomes generically unique. This warns us against the use of independent contracts in the analysis of agency relationships with purely self-interested agents.

The analysis of the multi-agent setting seems to suggest that status-seeking agents are desirable, partly because of the absence of productive interaction among agents in the model. However, this depends crucially on the assumption that the agents compare only what they are paid. I show that, if the agents compare their actions as well as their payments, then offering an appropriate extreme team contract to sufficiently inequity-averse agents is most desirable for the principal, because the agents are so intrinsically motivated to avoid being ahead that the principal can induce them to choose appropriate actions without leaving any rent to them.

Some recent literature has started to incorporate inequity aversion into principal-agent models. Englmaier and Wambach (2002) analyze the relationship between a principal and an agent. The agent is inequity averse and the principal serves as a reference actor for the agent. The authors do not impose the limited liability constraints on contracts, and hence in their analysis the participation constraint binds and plays a crucial role. Grund and Sliwka (2002) study rank-order tournaments among inequity-averse agents, and show that inequity-averse agents exert more effort than purely self-interested agents for a given contract, while the first-best effort is not implementable when contracts are endogenous. Neilson and Stowe (2003) analyze the optimal linear contract for other-regarding and risk-averse multiple agents, who are either inequity averse or status seeking. The limited liability constraints are not imposed. More importantly, the principal in their model is only allowed to offer independent contracts.

In a paper written independently and concurrently with this one, Rey Biel (2003) also studies a relationship of a principal with two agents who are inequity averse between them, and shows that the optimal contract is either a team contract or a relation performance contract, although he does not use these terms. In contrast to my model, however, output is deterministic and perfectly informs of the actions chosen by the agents in his model. Furthermore, he assumes exogenously that the participation constraint does not bind. Instead, he allows agents to be asymmetric in terms of output and cost, and studies the implementation of independent production in which only one of the agents produces, as well as of joint production.

The rest of the chapter is organized as follows. In Sect. 2, I briefly summarize the recent theories of other-regarding preferences. In Sect. 3, I study the single-agent case. Section 4 extends this to a multi-agent setting. In Sect. 5 I discuss implications for choice of preferences within organizations. Section 6 concludes.

2 Theories of Other-Regarding Preferences

In this section, I briefly summarize recent developments in modeling other-regarding players by relaxing the pure self-interest assumption.[7] The purpose of each theory "is *not* to explain every different finding by adjusting the utility function just so; the goal is to find parsimonious utility functions, supported by psychological intuition, that are general enough to explain many phenomena in one fell swoop, and also make new predictions" (Camerer 2003, p. 101).

[7] See Camerer (2003, Section 2.8), Fehr and Schmidt (2003), and Sobel (2001) for extensive surveys.

I call people *purely self-interested* if they care about their own material payoff, which I assume is monetary although it can be a vector of consumption goods. On the other hand, people are called *other-regarding*, if they care about others' payoffs, as well as their own.

Suppose there are two players 1 and 2. Let x_i be player i's material payoff. Each player i has preferences over (x_1, x_2), which I assume are represented by a utility function $u_i(x_1, x_2)$. Note that if player i is purely self-interested, his utility function depends only on x_i.

Various other-regarding preferences (alternatively called "interdependent preferences" or "social preferences") can be explained by the following specification: For $i = 1, 2$ and $j \neq i$,

$$u_i(x_1, x_2) = x_i + g_i(x_j - x_i)x_j. \tag{17.1}$$

The main features of this specification are that it is additively separable in the player's own material payoff and his concern for the other, and the latter part is multiplicatively separable in the function of his relative payoff and the other's material payoff.[8]

The two simplest examples of other-regarding preferences of this form are *altruism* and *spite*. Player i is purely altruistic if $g_i(\cdot)$ is constant and positive, while he is purely spiteful if $g_i(\cdot)$ is constant and negative. A slightly more elaborate specification is to assume $g_i(x_j - x_i) = \overline{g}_i$ if $x_i \geq x_j$ and $g_i(x_j - x_i) = \underline{g}_i$ if $x_i < x_j$ with $\overline{g}_i > \underline{g}_i > 0$: Although each player exhibits altruism, he puts more weight on the other's payoff if he is "ahead" ($x_i \geq x_j$) than if he is behind ($x_i < x_j$).[9]

Another category of examples places further restrictions on (17.1) as follows:

$$u_i(x_1, x_2) = x_i + g_i(x_j - x_i). \tag{17.2}$$

Function $g_i(\cdot)$ is defined as

$$g_i(z) = \begin{cases} -\alpha_i v(z) & \text{if z} \geq 0 \\ -\beta_i v(-z) & \text{if z} \leq 0 \end{cases} \tag{17.3}$$

where $v(z)$ is defined for $z \geq 0$ with $v(0) = 0$ and is strictly increasing, and α_i and β_i are constants with $\alpha_i > 0$. This means that, when player i is behind ($x_j - x_i > 0$), he prefers to reduce the inequality in payoffs between the two. If in addition $\beta_i > 0$, the same is true when player i is ahead ($x_j - x_i < 0$). The player with this

[8]See Neilson (2002) and Segal and Sobel (1999) for relevant axiomatization.

[9]An example of such a model is Charness and Rabin (2002).

type of other-regarding preferences is called *inequity averse*. For example, Fehr and Schmidt (1999) assume the following piecewise linear utility function[10]:

$$u_i(x_1, x_2) = \begin{cases} x_i - \alpha_i(x_j - x_i) & \text{if } \mathrm{x_i \leq x_j} \\ x_i - \beta_i(x_i - x_j) & \text{if } \mathrm{x_i \geq x_j}. \end{cases} \tag{17.4}$$

On the other hand, if $\beta_i < 0$, player i prefers to increase the difference in payoffs when he is ahead. This preference can be called *competitive* or *status-seeking*.[11]

In the theories of other-regarding preferences presented above, only the final monetary distribution among players matters. In this sense, these models are often referred to as the distributional approach. The second approach to modeling other-regarding players pays attention to intentions behind behavior. In particular, intention-based reciprocal behavior is regarded as one of the most important other-regarding behaviors. Reciprocal behavior is a response to actions, or intentions behind actions, of others, and can be either positive or negative. Positive reciprocity implies that if actions by others benefit a player, or if he perceives their actions as kind, then he will return the kindness to make the others better off. Negative reciprocity implies that if the others' actions are harmful or are perceived as unkind, then the player will retaliate to make the others worse off.

A seminal paper in the intention-based approach is that of Rabin (1993). He proposes that a player's preferences depend on the strategies of all the players, his beliefs about the others' strategy, and his beliefs about the others' beliefs about his own strategy:

$$u_i(\sigma_i, \sigma_j, \sigma_i') = v_i(\sigma_i, \sigma_j) + g_i(\sigma_i', \sigma_j)v_j(\sigma_i, \sigma_j) \tag{17.5}$$

where σ_i is player i's strategy, σ_j is player i's belief about player j's strategy choice, and σ_i' is player i's belief about what player j believes about player i's strategy choice.[12] Since beliefs enter into preferences, Rabin utilizes psychological game theory (Geanakoplos et al. 1989) and solves for equilibria in strategies and

[10]Bolton and Ockenfels (2000) propose an alternative specification of inequity aversion, which does not take the form in (17.2). They write player i's utility function as $u_i(x_i, s_i)$ with

$$s_i = \begin{cases} \dfrac{x_i}{x_1 + x_2} & \text{if } \mathrm{x_1 + x_2 \neq 0} \\ \dfrac{1}{2} & \text{if } \mathrm{x_1 + x_2 = 0} \end{cases}$$

where $\partial u_i/\partial s_i(x_i, 1/2) = 0$ and $\partial^2 u_i/\partial s_i^2(x_i, s_i) < 0$. $u_i(x_i, s_i)$ is thus maximized at $s_i = 1/2$: Each player will sacrifice to move his share closer to the average if he is either below or above it. Some experimental evidence that compares Bolton and Ockenfels (2000) with Fehr and Schmidt (1999) is discussed in Camerer (2003, Section 2.8.5).

[11]This terminology follows Neilson and Stowe (2003).

[12]See Segal and Sobel (1999) for a related axiomatic approach to generating preferences that can reflect intention-based reciprocity.

beliefs. Rabin's model is only for normal-form games; Dufwenberg and Kirchsteiger (2002) and Falk and Fischbacher (2000) extend his theory to extensive-form games.

These equilibrium models are difficult to use in most applications, because they are very involved and in general have multiple equilibria. Charness and Rabin (2002), Cox and Friedman (2002), and Levine (1998) attempt to develop more tractable models in the spirit of intention-based reciprocity. For example, Levine (1998) assumes that player i differs in the extent (type) to which he cares about the others, and if his type is $\alpha_i > 0$, his preferences are represented by

$$u_i(x_1, x_2) = x_i + \frac{\alpha_i + \lambda\alpha_j}{1 + \lambda} x_j \tag{17.6}$$

where $0 \leq \lambda \leq 1$. This is another example of specification (17.1). The idea is that player i cares more about player j's material payoff if player j cares more about player i (i.e. if α_j is higher).

In this chapter, I adopt the distributional approach, and in particular the theory of inequity aversion *à la* Fehr and Schmidt (1999), for the following reasons. First of all, it is simple and tractable. Second, it is not an *ad hoc* specification tailored for a particular finding. It is a parsimonious specification of other-regarding preferences, supported by psychological intuition, that can explain various types of experimental results with a single function. And it is already a well-studied model; for example, an axiomatic foundation for it has been developed (Neilson 2002). Third, it is argued that "[i]n many situations, reciprocal persons and inequity averse persons behave in similar ways" (Fehr and Fischbacher 2002, p. C3) and these ways seem to be more important than pure altruism and spitefulness. The inequity aversion model can be used as a "shortcut" for studying the effects of important intention-based reciprocal behavior. However, we should be aware that reciprocity and inequity aversion are distinct motives, and often intention matters more, in particular in the domain of punishing behavior, as suggested by recent evidence.[13]

3 The Principal-Agent Model with Other-Regarding Preferences

3.1 Benchmark: The Self-Interest Case

I analyze the effects of other-regarding preferences in the following simple but standard principal-agent framework. The principal hires an agent for engaging

[13] See e.g. Fehr and Schmidt (2003) for a survey. Note that, according to them, "the evidence also suggests that inequity aversion plays an additional, nonnegligible role" (p. 238). Charness and Haruvy (2002) estimate and compare various theories using experimental data on gift exchange games. They argue that distributional concern as well as intention play a significant role in players' decisions.

in a project. Both are assumed to be risk neutral. The agent chooses an action from $A = \{a_0, a_1\}$. When action a_i is chosen, the agent incurs private cost d_i. I assume $d_1 > d_0 = 0$ and denote $d_1 = d$ for simplicity. The project succeeds with probability p_i and generates benefit b_s for the principal, while it fails with probability $1-p_i$ and the benefit is b_f. I assume $1 > p_1 > p_0 > 0$ and $b_s > b_f = 0$, and denote $b_s = b$ for simplicity.

The outcome of the project is verifiable: the principal offers a contract (w_s, w_f) in which the agent is paid w_s if the project succeeds and w_f if it fails. I assume that the contract must satisfy the *limited liability constraints for the agent*:

$$w_j \geq 0, \quad j = s, f. \tag{LL1}$$

I denote the set of feasible contracts by $C = \{(w_s, w_f) \mid (w_s, w_f)\text{satisfies(LL1)}\}$.

The timing of the game is as follows. First, the principal offers a contract. The agent decides whether to accept or reject the contract. If rejected, the game ends and the agent receives the reservation utility $\overline{u}$, which I assume is zero.[14] After accepting the contract, the agent chooses an action. The outcome of the project is then realized and the transfer is made according to the contract. I assume that the contract is binding and cannot be renegotiated.

In the standard setting in which all the parties are purely self-interested, the payoffs to the principal and the agent are given as $u_P = b_j - w_j$ and $u_A = w_j - d_i$, respectively, for action a_i and outcome $j = s, f$.

Throughout the paper, I assume that b is sufficiently large that the principal prefers to implement a_1 to a_0. The incentive compatibility constraint and the participation constraint are, respectively,

$$\Delta_p \Delta_w \geq d \tag{IC1}$$

$$w_f + p_1 \Delta_w \geq d \tag{PC1}$$

where $\Delta_w = w_s - w_f$ and $\Delta_p = p_1 - p_0$. The principal chooses $(w_s, w_f) \in C$, which minimizes the expected payment $w_f + p_1 \Delta_w$ subject to (IC1) and (PC1).

Consider any contract in C that satisfies (PC1) with equality: $w_f + p_1 \Delta_w = d$. This is a first-best solution if the action is enforceable. However, since the action is not verifiable, the contract must satisfy (IC1) as well. To this end, consider a first-best contract $(\overline{w}_s, 0)$ where $p_1 \overline{w}_s = d$. Substituting this contract into the left-hand side of (IC1) yields

$$\Delta_p \overline{w}_s = \Delta_p \left(\frac{d}{p_1}\right) < d$$

[14]The results of the paper will continue to hold if $\overline{u} > 0$, although some care must be taken since the participation constraint will sometimes bind.

and hence no first-best contract satisfies (IC1). (IC1) therefore must bind. The optimal second-best contract (w_s^*, w_f^*) is unique and is given by $w_f^* = 0$ and

$$w_s^* = \frac{d}{\Delta_p}.$$

To demonstrate the uniqueness, suppose $(w_s, w_f) \neq (w_s^*, 0)$ solves the principal's problem. Since (IC1) must bind at optimum, $\Delta_w = w_s - w_f = w_s^*$ and $w_f > 0$ holds. Then the principal's expected payment is $w_f + p_1\Delta_w > p_1 w_s^*$, which contradicts the optimality of (w_s, w_f).

The benchmark results shown above are summarized in the following proposition.

Proposition 17.1 *Suppose both principal and agent are self-interested. The optimal contract is unique and is given by* $(w_s^*, 0)$. *The first-best solution cannot be implemented.*

There exists a moral hazard problem, and in order to mitigate the problem the principal must provide higher-powered incentives and leave rents $p_1 w_s^* - d = p_0 d/\Delta_p > 0$ to the agent.[15]

3.2 Other-Regarding Agent

I now suppose that the agent is other-regarding and cares about the principal's material payoff as well as his own. I assume that the agent has the following utility function: for $i = 0, 1$ and $j = s, f$,

$$u_A = w_j - d_i - \alpha v\left(\max\{b_j - 2w_j, 0\}\right) - \alpha\gamma v\left(\max\{2w_j - b_j, 0\}\right) \tag{17.7}$$

where $\alpha \geq 0$ and γ are constants. This utility function is just a restatement of (17.2), and I assume $v(0) = 0$, $v'(z) > 0$ for all $z > 0$, and $\lim_{z\to\infty} v(z) = \infty$.

The important assumption behind (17.7) is that the agent compares his income w_j with the principal's $b_j - w_j$; in other words, it is assumed that the agent does not take into account the disutility of his action d_i. I later analyze the alternative specification in which the agent compares his "net" material payoff $w_j - d_i$ with the principal's.

Note that, when $2w_j - b_j > 0$ holds, the principal's monetary payoff is smaller than the agent's, and hence the agent is "ahead," while the agent is "behind" when $b_j - 2w_j > 0$ holds. Equation (17.7) can thus be rewritten as follows:

[15]If the principal implements a_0, the optimal contract is $w_s = w_f = 0$ and the principal's expected payoff is $p_0 b$. Since I am assuming that the principal prefers to implement a_1 to a_0, the following condition is assumed implicitly: $\Delta_p b > p_1 d/\Delta_p$.

$$u_A = \begin{cases} w_j - d_i - \alpha\gamma v(2w_j - b_j) & \text{if } w_j \geq b_j - w_j \\ w_j - d_i - \alpha v(b_j - 2w_j) & \text{if } w_j \leq b_j - w_j. \end{cases} \tag{17.8}$$

Parameter α captures the extent to which the agent cares about the principal's material payoff. If $\gamma < 0$, the agent prefers to increase the difference in payoffs when he is ahead; hence he is competitive or status-seeking. On the other hand, if $\gamma > 0$, the agent is inequity averse and his utility is decreasing in the difference in payoffs between the principal and the agent, whether the agent is behind or ahead. Based on experimental evidence, Fehr and Schmidt (1999) assume $\gamma \leq 1$; that is, the agent suffers from inequality more when he is behind than when he is ahead. At this point I do not impose this restriction. Note that Fehr and Schmidt (1999) characterize the utility function by two parameters α and $\beta = \alpha\gamma$. I use γ instead of β, since by changing α I can examine how the extent to which the agent is other-regarding affects incentives.

To simplify the exposition, I assume $w_f = 0$. This assumption can be justified if the principal's limited liability constraints ($w_s \leq b$ and $w_f \leq 0$) are imposed. Alternatively, I will show later in this section that the optimal contract in fact satisfies $w_f = 0$ if an additional assumption is made instead of the principal's limited liability. Since the agent does not suffer from inequality when the project fails, his other-regarding preferences concern how the benefit from the successful project is divided between the principal and the agent. The incentive compatibility constraint is then written as follows:

$$w_s - \alpha\gamma v(2w_s - b) \geq \frac{d}{\Delta_p} \qquad \text{if } w_s \geq \frac{1}{2}b \tag{IC2a}$$

$$w_s - \alpha v(b - 2w_s) \geq \frac{d}{\Delta_p} \qquad \text{if } w_s \leq \frac{1}{2}b \tag{IC2b}$$

where (IC2a) represents the incentive compatibility constraint when the agent is ahead, while (IC2b) is the constraint when he is behind.

I first obtain a necessary and sufficient condition for the existence of a contract satisfying (IC2b).

Lemma 17.1 *There exists a contract that satisfies* (IC2b) *if and only if*

$$\frac{1}{2}\Delta_p b \geq d \tag{17.9}$$

holds.

Proof Since the left-hand side of (IC2b) is increasing in w_s, the existence is guaranteed if (IC2b) is satisfied at $w_s = b/2$, that is, if (17.9) holds. Conversely, if $w_s \leq b/2$ satisfies (IC2b), then

$$\frac{d}{\Delta_p} \leq w_s - \alpha v(b - 2w_s) \leq \frac{1}{2}b$$

holds. □

Define $\hat{w}_s$ implicitly by

$$\hat{w}_s - \alpha v(b - 2\hat{w}_s) = \frac{d}{\Delta_p}. \tag{17.10}$$

Note that $\hat{w}_s$ satisfies $\hat{w}_s \leq b/2$ if (17.9) holds. Since I focus on the implementation of action a_1, I maintain (17.9) for most of the analysis and state it as an assumption. The case in which (17.9) does not hold will be discussed later in this section.

Assumption 17.1 The profit to the principal from the successful project is so large that (17.9) holds.

Proposition 17.2 states that $(\hat{w}_s, 0)$ is the optimal contract under Assumption 17.1.

Proposition 17.2 *Suppose Assumption 17.1 holds. (i)* $(\hat{w}_s, 0)$ *is optimal. (ii) The principal's expected payment under the optimal contract is increasing in* α. *If (17.9) holds with strict inequality, it is strictly increasing in* α.

Proof (i) It is sufficient to show that $(\hat{w}_s, 0)$ satisfies the participation constraint:

$$w_s - \alpha v(b - 2w_s) \geq \frac{d}{p_1}. \tag{PCb}$$

By definition, $\hat{w}_s - \alpha v(b - 2\hat{w}_s) = d/\Delta_p > d/p_1$. (PCb) is hence satisfied. (ii) The principal's expected payment is $p_1\hat{w}_s$. It is thus sufficient to show that $\hat{w}_s$ is increasing in α. Define the left-hand side of (17.10) by $f(\alpha)$. Then

$$\frac{\partial f(\alpha)}{\partial \alpha} = -v(b - 2\hat{w}_s) < 0$$

for $\hat{w}_s < b/2$, which implies that $\hat{w}_s$ is strictly increasing in α. If (17.9) holds with equality, $\hat{w}_s$ does not depend on α. □

Proposition 17.2 implies the following. (i) Similar to the pure self-interest case, at optimum the incentive compatibility constraint binds while the participation constraint does not, and hence the agent earns rents. (ii) The principal is worse off the more the agent cares about the principal's wellbeing. To understand why other-regarding preferences hurt the principal, remember that the left-hand side of (IC2b) is decreasing in α: since the agent suffers from inequity when the project succeeds, the principal must pay more to satisfy the incentive compatibility constraint the more other-regarding is the agent.[16]

[16]The result that the principal prefers lower α will continue to hold if $\overline{u} > 0$ and the participation constraint binds. In this case, the principal has to increase payments for the obvious reason that the participation constraint becomes harder to satisfy the more other-regarding the agent is.

I next show that, even if contracts with $w_f > 0$ are feasible, $(\hat{w}_s, 0)$ is uniquely optimal under an additional assumption.

Proposition 17.3 *Suppose that Assumption 17.1 holds and that contracts with $w_f > 0$ are feasible. Then $(\hat{w}_s, 0)$ is the unique optimal contract if*

$$1 \geq 2\alpha\gamma v'(z) \tag{17.11}$$

is satisfied for all $z > 0$.

Proof Suppose that the conclusion does not hold and (w_s, w_f) with $w_f > 0$ is optimal. This implies that

$$w_f + p_1\Delta_w \leq p_1\hat{w}_s \tag{17.12}$$

holds. Since $\hat{w}_s \leq b/2$ and $w_f > 0$, $w_s < b/2$ must hold, and (w_s, w_f) satisfies the incentive compatibility constraint

$$\Delta_w - \alpha v(b - 2w_s) + \alpha\gamma v(2w_f) \geq \frac{d}{\Delta_p}. \tag{17.13}$$

Combining (17.13) and (17.10) yields

$$\begin{aligned}\Delta_w + \alpha\gamma v(2w_f) &\geq \hat{w}_s + \alpha v(b - 2w_s) - \alpha v(b - 2\hat{w}_s)\\ &\geq \Delta_w + \frac{w_f}{p_1} + \alpha v(b - 2w_s) - \alpha v(b - 2\hat{w}_s).\end{aligned}$$

The second inequality is due to (17.12). Rearranging yields

$$-\frac{w_f}{p_1} + \alpha\gamma v(2w_f) \geq \alpha v(b - 2w_s) - \alpha v(b - 2\hat{w}_s). \tag{17.14}$$

The right-hand side of (17.14) is positive since (17.12) leads to $w_s < \hat{w}_s$. Thus, if the left-hand side is nonpositive, (w_s, w_f) does not satisfy the incentive compatibility constraint, which is a contradiction. Equation (17.11) provides such a condition.[17]

□

A contract with $w_f > 0$ creates inequity even when the project fails. The agent is ahead, since he receives w_f while the principal pays w_f. If the agent is inequity averse ($\gamma > 0$), this change strengthens his incentive to choose a_1, as shown in the incentive compatibility constraint (17.13). However, for this change to benefit the principal, she must decrease w_s to lower Δ_w as well as to raise inequity $b - 2w_s$ when the agent is behind. These changes bring negative effects

[17] Actually, a weaker condition, $1 \geq 2p_1\alpha\gamma v'(z)$, or $1 \geq p_1\alpha\gamma v'(0)$ if $\gamma v''(z) \leq 0$, is sufficient.

on incentives. Condition (17.11) is sufficient for the negative effects to dominate. If condition (17.11) is violated, then $1 < 2\alpha\gamma v'(z)$ for some $z > 0$, which implies that the agent may be willing to transfer some payment back to the principal in order to make him less ahead of the principal. Since this seems to be implausible, I believe condition (17.11) is a reasonable one. In particular, (17.11) holds if the agent is status-seeking ($\gamma < 0$).

3.2.1 When Assumption 17.1 Does Not Hold

The results obtained so far are based on Assumption 17.1. I now examine the case where Assumption 17.1 does not hold: Assume instead (only in this subsection) that

$$\frac{1}{2}\Delta_p b < d. \tag{17.15}$$

By Lemma 17.1, there is no contract $(w_s, 0)$ that satisfies the incentive compatibility constraint in the range $w_s \leq b/2$. The principal thus has to choose $w_s > b/2$ to satisfy (IC2a). The left-hand side of (IC2a) is increasing in w_s if condition (17.11) holds, which I assume throughout this subsection.[18] Then define w_s^+ implicitly by

$$w_s^+ - \alpha\gamma v(2w_s^+ - b) = \frac{d}{\Delta_p}. \tag{17.16}$$

Writing the left-hand side of (17.16) as $g(\alpha)$ and differentiating it with respect to α yields

$$\frac{\partial g(\alpha)}{\partial \alpha} = -\gamma v(2w_s^+ - b)$$

for $w_s^+ > b/2$; and hence w_s^+ is strictly increasing (decreasing) in α if $\gamma > 0$ (respectively $\gamma < 0$). Therefore as in the previous result in Proposition 17.2 (ii), the principal does not benefit from a more "fair-minded" agent if he is inequity averse. However, if the agent is status-seeking, the principal is better off as the agent is more other-regarding. Since $w_s^+ > b/2$, the agent is ahead if the project succeeds, and hence the status-seeking preferences come into play. The status-seeking agent enjoys being ahead, and thus the principal can implement a_1 with a lower cost the more competitive the agent is. This incentive effect arises only when condition (17.15) holds.

[18] If (17.11) does not hold, no contract $(w_s, 0)$ can satisfy (IC2a) and hence none can implement a_1. The principal must increase w_f from zero in order to encourage the agent to choose a_1, by making the agent suffer from the increased inequity facing the unsuccessful project.

3.3 Alternative Specification

I have so far assumed that the agent compares his income w_j to the principal's $b_j - w_j$. However, since the agent knows he incurs the private cost of action d_i, he may be concerned about inequity differently, depending on whether his action is a_0 or a_1.

In this subsection, I assume alternatively that the agent compares his "net" material payoff $w_j - d_i$ to the principal's payoff $b_j - w_j$. Will the principal want the agent to be more inequity averse in this alternative specification?

First suppose $b - w_s > w_s$. Assuming $w_f = 0$ for simplicity, the incentive compatibility constraint and the participation constraint are, respectively,

$$\Delta_p w_s - [p_1 \alpha v(b - 2w_s + d) - p_0 \alpha v(b - 2w_s)] - (1 - p_1)\alpha v(d) \geq d \quad \text{(IC3b)}$$

$$p_1 w_s - p_1 \alpha v(b - 2w_s + d) - (1 - p_1)\alpha v(d) \geq d. \quad \text{(PC3b)}$$

Two modifications should be noted. First, the inequity aversion term under the successful project $\alpha v(b - 2w_s + d)$ contains d. However, if the agent chooses a_0 instead of a_1, the cost of action is zero, and hence the change in the inequity aversion term is

$$p_1 \alpha v(b - 2w_s + d) - p_0 \alpha v(b - 2w_s).$$

This change is larger under the current specification than the corresponding change in (IC2b) that is equal to $\Delta_p \alpha v(b - 2w_s)$, since by choosing a_0 the agent can save the cost of action and reduce the disutility from inequity when he is behind. In other words, when he is behind by choosing a_0, he will not feel as envious as by choosing a_1. The incentive compatibility constraint is thus harder to satisfy than before.

The second change is the new inequity-averse term $\alpha v(d)$ when the project fails. Since the agent incurs the cost of action d, he is again behind, and suffers from inequity. Choosing a_0 instead relieves him of the envy he would suffer if the project fails. This change again makes the incentive compatibility constraint tighter.

The incentive compatibility constraint and the participation constraint therefore become more stringent under this alternative specification than under the original specification, and hence the principal is again worse off the more other-regarding the agent is: the previous results are reinforced.

Next suppose $b - w_s < w_s - d$. The second change remains $\alpha v(d)$ while the first change becomes $p_1 \alpha \gamma v(2w_s - b - d) - p_0 \alpha \gamma v(2w_s - b)$. Now choosing a_1 *reduces* the difference by d when the agent is ahead. If the agent is status-seeking ($\gamma < 0$), both changes again make the constraint harder to satisfy. On the other hand, if the agent is inequity averse ($\gamma > 0$), the first change works so as to *benefit* the principal by making the incentive compatibility constraint easier to satisfy. However, the effect of the second change remains. For example, when

$v(z) = z$, the second change dominates and the incentive compatibility constraint becomes harder to satisfy if and only if $1 - p_1 > p_1\gamma$.[19]

3.4 Other-Regarding Principal

The other-regarding behavior of the player in the role of the principal is more subtle to identify. The principal may behave fairly either because she is fair-minded or because she anticipates that otherwise the agent will respond so as to hurt her. Experimental evidence on proposers' behavior in ultimatum games along with the use of dictator games shows that both explanations are likely to be valid (Forsythe et al. 1994).

To see the effects of other-regarding preferences from the principal's point of view, suppose the principal's utility function is given as follows. For $j = s, f$,

$$u_P = \begin{cases} b_j - w_j - \pi\delta y(b_j - 2w_j) & \text{if } b_j - w_j \geq w_j \\ b_j - w_j - \pi y(2w_j - b_j) & \text{if } b_j - w_j \leq w_j \end{cases} \tag{17.17}$$

where $\pi \geq 0$ and δ are constants, $y(0) = 0$, and $y'(z) > 0$ for $z > 0$.

Suppose Assumption 17.1 holds. The optimal contract when the principal is purely self-interested is $(\hat{w}_s, 0)$. Since $\hat{w}_s \leq b/2$, the principal is ahead when the project succeeds. This contract is still optimal when the principal's preferences are represented by (17.17), if the principal is better off the lower the payment is:

$$1 - 2\pi\delta y'(b - 2w_s) > 0.$$

For example, the condition is satisfied for the status-seeking principal ($\delta < 0$), or for the inequity-averse principal ($\delta > 0$) but with $\pi\delta$ sufficiently close to zero.

On the other hand, if

$$1 - 2\pi\delta y'(b - 2w_s) < 0$$

holds, then the principal will prefer to *increase* the payment for the agent in order to reduce inequity when she is ahead (up to $b/2$). However, increasing the payment beyond $b/2$ puts her behind, and hence she will prefer to *lower* the payment. Therefore $w_s = b/2$ holds at optimum: since the principal is so much inequity averse, the optimal contract attains a precisely equal division between $b_j - w_j$ and w_j for $j = s, f$.

[19] The remaining case is $w_s > b - w_s > w_s - d$. While the effect of the second change is the same, whether or not the first change makes the incentive compatibility constraint harder to satisfy is difficult to tell.

4 Multiple Agents

4.1 The Model

I shall now extend the model of the previous section to a multi-agent setting, and assume that the principal does not belong to the agents' reference group, while each agent cares about the payoff to the other agents. The production technology is the same as in the previous section. The principal hires two agents, denoted by $n = 1, 2$, each of which engages in a project separately. Each agent chooses an action from $A = \{a_0, a_1\}$. Action a_0 costs $d_0 = 0$ while a_1 costs the agent $d_1 = d > 0$. When action a_i is chosen, the project succeeds with probability p_i and generates profit $b_s = b > 0$ while it fails with probability $1 - p_i$ and generate profit $b_f = 0$. I assume that action a_1 is more productive and that $1 > p_1 > p_0 > 0$. There is no correlation.

The timing of the game is as follows. First, the principal offers a contract to the agents. The agents decide simultaneously whether to accept or reject the contract. If rejected by at least one agent, the game ends and each agent receives the reservation utility zero. After both agents accept the contract, they choose actions simultaneously. The outcomes of the projects are then realized and the transfers are made according to the terms of the contract.

Since the outcome of each project is verifiable, let $w^n = (w^n_{jk})_{j,k=s,f}$ be the payment scheme offered to agent n, where w^n_{jk} represents the payment to agent n when his outcome is j and the other agent's outcome is k $(j, k = s, f)$. The payment scheme must satisfy the limited liability constraints

$$w^n_{jk} \geq 0, \qquad j, k = s, f. \tag{LL2}$$

Let $C_n = \{w^n \mid w^n \text{satisfies(LL2)}\}$ be the set of feasible contracts for agent n, and $C = C_1 \times C_2$.

Agent n's utility function is as follows. For $i = 0, 1$, $j, k = s, f$, $n, m = 1, 2$, and $m \neq n$,

$$u^n = w^n_{jk} - d_i - \alpha_n v_n \left(\max\left\{w^m_{kj} - w^n_{jk}, 0\right\}\right) - \alpha_n \gamma_n v_n \left(\max\left\{w^n_{jk} - w^m_{kj}, 0\right\}\right) \tag{17.18}$$

where $\alpha_n \geq 0$, $v_n(0) = 0$, and $v'_n(z) > 0$ for $z > 0$. I assume $|\gamma_n| \leq 1$: being behind by a certain amount changes the agent's utility at least as much as being ahead by an equal amount. Fehr and Schmidt (1999) assume this for inequity-averse agents on the basis of experimental evidence: each inequity-averse agent dislikes inequality at least as much when he is behind as when he is ahead. I extend this assumption to status-seeking agents: each status-seeking agent likes to be ahead no better than he likes to avoid being behind. Although I believe this is a reasonable assumption similar to loss aversion, it does not play an essential role in what follows.

Note that the agents compare the payments from the principal for each pair of project outcomes. The implicit assumption behind this formulation is that each agent either observes or estimates correctly what the other agent is paid, while the other agent's choice of action is not a concern. Since the project outcomes are verifiable and I will soon assume symmetric agents, the former part of the assumption seems reasonable. The latter part may be questionable, however. One could argue that each agent is likely to anticipate the other agent's action correctly and also to compare both actions. Since whether or not actions will be compared is an empirical question, I will first consider this simpler specification, and later turn to the specification in which the agents compare their net payoffs.

To simplify the analysis, I adopt the following assumptions.

Assumption 17.2 (a) The agents are symmetric: $\alpha = \alpha_1 = \alpha_2$, $\gamma = \gamma_1 = \gamma_2$, and $v(\cdot) = v_1(\cdot) = v_2(\cdot)$. (b) The principal chooses a symmetric contract satisfying $w^1 = w^2$. (c) $v(z) = z$ for all $z \geq 0$. (d) $\alpha\gamma \leq 1$.

If Assumption 17.2 (d) fails to hold, an inequity-averse agent, when he is ahead, will want to give up his income in order to reduce the inequality between his and the other agent's payoff, which seems to be implausible.[20] Agent 1's utility function is then rewritten as follows. (Agent 2's utility function can be similarly rewritten.)

$$u^1 = \begin{cases} w^1_{jk} - d_i - \alpha\gamma(w^1_{jk} - w^2_{kj}) & \text{if } \mathrm{w}^1_{jk} \geq \mathrm{w}^2_{kj} \\ w^1_{jk} - d_i - \alpha(w^2_{kj} - w^1_{jk}) & \text{if } \mathrm{w}^1_{jk} \leq \mathrm{w}^2_{kj}. \end{cases} \tag{17.19}$$

As before, I assume that b is large enough for the principal to want to implement (a_1, a_1). She will choose a symmetric contract $(w^1, w^2) = (w, w) \in C$ to minimize the expected payments.

In the Appendix, I show that it is without loss of generality to restrict contracts to those satisfying $w_{fs} = w_{ff} = 0$, that is those which pay the least possible amount to each agent when his project fails. Then the incentive compatibility constraint and the participation constraint are, respectively,

$$p_1 w_{ss} + (1 - p_1)w_{sf} + [p_1 - (1 - p_1)\gamma]\alpha w_{sf} \geq \frac{d}{\Delta_p} \tag{IC4}$$

$$p_1 w_{ss} + (1 - p_1)w_{sf} - (1 - p_1)\alpha(1 + \gamma)w_{sf} \geq \frac{d}{p_1}. \tag{PC4}$$

When the project of an agent fails and the other's project succeeds, the former agent is behind and suffers αw_{sf} from inequity aversion. On the other hand, if his project succeeds and the other's fails, he is ahead and his equity concern is represented by $\alpha\gamma w_{sf}$.

[20] Although not necessary for the results of the paper, $\alpha\gamma \leq 1/2$ may be a more reasonable assumption: otherwise, the agent who is ahead may want to renegotiate *ex post* to transfer his rewards to the other agent.

I call w a *team contract* if $w_{ss} > w_{sf}$, while w is called a *relative performance contract* if $w_{ss} < w_{sf}$. Finally, if $w_{ss} = w_{sf}$, the contract for each agent depends on the outcome of his project only, and hence it is called an *independent contract*. The principal chooses a feasible contract that minimizes the expected payments $2p_1(w_{sf} + p_1(w_{ss} - w_{sf}))$ subject to (IC4) and (PC4).

4.1.1 Benchmark: The Self-Interest Case

Before deriving the optimal contract, I solve the optimal contract for the benchmark case of purely self-interested agents ($\alpha = 0$). It is easy to show that only the incentive compatibility constraint binds, that any feasible contract that solves

$$p_1 w_{ss} + (1 - p_1)w_{sf} = \frac{d}{\Delta_p}$$

is optimal, and that the principal's expected payment is $2p_1 d/\Delta_p$. Note, in particular, that the independent contract $w_{ss} = w_{sf} = d/\Delta_p$ is an optimal contract.

4.2 Analysis

Now consider other-regarding agents ($\alpha > 0$). First, suppose that the incentive compatibility constraint (IC4) binds.

$$p_1 w_{ss} + (1 - p_1)w_{sf} = \frac{d}{\Delta_p} + [(1 - p_1)\gamma - p_1]\alpha w_{sf}. \tag{17.20}$$

Note that the larger the left-hand side of (17.20) is, the higher the principal's expected payment is. Therefore if $(1 - p_1)\gamma > p_1$, the principal will prefer to set $w_{ss} > w_{sf} = 0$, while if $(1 - p_1)\gamma < p_1$, the principal will want to set $w_{ss} = 0 < w_{sf}$.[21] Define $\hat{w}_{ss}$ and $\hat{w}_{sf}$ as follows:

$$\hat{w}_{ss} = \frac{d}{\Delta_p}\frac{1}{p_1} \tag{17.21}$$

$$\hat{w}_{sf} = \frac{d}{\Delta_p}\frac{1}{(1 - p_1) + \alpha[p_1 - (1 - p_1)\gamma]}. \tag{17.22}$$

I analyze two cases separately.

Case 1: $(1 - p_1)\gamma > p_1$

[21] If $(1 - p_1)\gamma = p_1$, any contract satisfying (17.20) is optimal.

The best contract among those under which the incentive compatibility constraint binds is $(w_{ss}, w_{sf}) = (\hat{w}_{ss}, 0)$. By (PC4) and $p_1 > \Delta_p$, this contract satisfies the participation constraint, and hence is the optimal contract. The expected payment is $2p_1 d/\Delta_p$, which is equal to the expected payment under the optimal contract for the purely self-interested agents.

Case 2: $(1 - p_1)\gamma < p_1$

The incentive compatibility constraint binds at $(0, \hat{w}_{sf})$. I now derive the condition for $(0, \hat{w}_{sf})$ to satisfy the participation constraint. Substituting $(0, \hat{w}_{sf})$ into (PC4) yields

$$\frac{d}{\Delta_p} \frac{(1 - p_1)(1 - \alpha(1 + \gamma))}{(1 - p_1) + \alpha[p_1 - (1 - p_1)\gamma]} \geq \frac{d}{p_1}.$$

A straightforward calculation yields the following necessary and sufficient condition for $(0, \hat{w}_{sf})$ to satisfy (PC4):

$$\frac{\ell_s}{\ell_f} \leq \frac{1 - \alpha\gamma}{\alpha} \tag{17.23}$$

where $\ell_s = p_1/p_0$ and $\ell_f = (1 - p_1)/(1 - p_0)$ are likelihood ratios. When conditions $(1 - p_1)\gamma < p_1$ and (17.23) hold, $(0, \hat{w}_{sf})$ is the optimal contract. I call this case Case 2a. The expected payment is calculated as

$$2p_1(1 - p_1)\hat{w}_{sf} = \frac{2d}{\Delta_p} \frac{p_1(1 - p_1)}{(1 - p_1) + \alpha[p_1 - (1 - p_1)\gamma]} \tag{17.24}$$

which is equal to $2p_1 d/\Delta_p$ at $\alpha = 0$, and is decreasing in α and increasing in γ.

When $(0, \hat{w}_{sf})$ fails to satisfy (17.23), the participation constraint (PC4) must bind:

$$p_1 w_{ss} + (1 - p_1)w_{sf} = \frac{d}{p_1} + (1 - p_1)\alpha(1 + \gamma)w_{sf}. \tag{17.25}$$

Since the principal's expected payments is higher the larger the left-hand side is, the principal will want to choose w_{sf} as small as possible. However, $w_{sf} = 0$ does not satisfy the incentive compatibility constraint (IC4), and hence both (IC4) and (PC4) bind. Solving the simultaneous equations (17.25) and (17.20) provides the solution $(\overline{w}_{ss}, \overline{w}_{sf})$ as follows:

$$\overline{w}_{ss} = \frac{d}{\Delta_p} \frac{1 - p_1}{\ell_s p_1} \left(\frac{\ell_s}{\ell_f} - \frac{1 - \alpha\gamma}{\alpha} \right) \tag{17.26}$$

$$\overline{w}_{sf} = \frac{d}{\Delta_p} \frac{1}{\ell_s \alpha}. \tag{17.27}$$

Note that $\overline{w}_{ss} > 0$ if (17.23) does not hold. The expected payment is calculated as

$$2p_1\left[p_1\overline{w}_{ss} + (1-p_1)\overline{w}_{sf}\right] = \frac{2d}{\Delta_p}\frac{p_1(1-p_1)}{\ell_s}\left(\frac{\ell_s}{\ell_f}+\gamma\right). \tag{17.28}$$

which is independent of α and increasing in γ. I call this case Case 2b. This case applies when $\alpha^{-1} - (\ell_s/\ell_f) < \gamma < p_1/(1-p_1)$, a range exists if and only if $\alpha > (1-p_1)p_0/p_1$.

The following proposition summarizes the results obtained.

Proposition 17.4 *The optimal contract $w^* = (w^*_{ss}, w^*_{sf})$ is given as follows.*

Case 1: $w^* = (\hat{w}_{ss}, 0)$ if $\gamma > p_1(1-p_1)$ holds. It is an extreme team contract. The expected payment does not depend on α or γ.

Case 2a: $w^* = (0, \hat{w}_{sf})$ if both $\gamma < p_1(1-p_1)$ and $\gamma \leq \alpha^{-1} - (\ell_s/\ell_f)$ hold. It is an extreme relative performance contract. The expected payment is decreasing in α and increasing in γ.

Case 2b: $w^* = (\overline{w}_{ss}, \overline{w}_{sf})$ if both $\gamma < p_1(1-p_1)$ and $\gamma > \alpha^{-1} - (\ell_s/\ell_f)$ hold. The expected payment is independent of α and increasing in γ.

Other-regarding preferences provide two incentive effects via the incentive compatibility constraint (IC4). An agent is behind if his project is unsuccessful while the other project succeeds. Since the agent faces an incentive to reduce the probability that he will incur disutility αw_{sf} from being behind, this brings about a positive incentive effect. The second effect arises from the situation in which an agent is ahead. If his project succeeds while the other project fails, the inequity-averse ($\gamma > 0$) agent suffers from being ahead, and hence he is discouraged from increasing the probability of success. This second effect is negative for inequity-averse agents. In Case 1 in Proposition 17.4, the second negative effect dominates, because (i) each agent is sufficiently averse to being ahead (γ large); or (ii) each project is relatively uncertain and uncontrollable ($p_1/(1-p_1)$ small) so that the incentive effect of being ahead (due to the other's unsuccessful project) is more important than that of being behind (due to the other's successful project). Since the effects of inequity aversion should be minimized in this case, the extreme team contract is adopted.

Note the important feature of extreme team contracts. They are "fair" in the sense that both agents are always paid exactly the same amount; however, exactly because of this feature, the principal's payoff turns out to be independent of the extent to which the agents care about each other's wellbeing—she neither benefits nor suffers from the agents' other-regarding preferences.

If the project is relatively controllable, or if the extent to which the agent is averse to being ahead is small, the first positive incentive effect dominates the second. This is Case 2a or 2b in Proposition 17.4. The principal will then want to utilize this positive effect in her contract design by adopting a relative performance contract to generate the possibility of inequity. In particular, if the agents are status seeking, the second incentive effect as well as the first effect is positive. (They prefer being ahead.) As long as the participation constraint does not bind (Case 2a), a tournament-like extreme relative performance contract emerges as an optimal contract even though there is no systematic shock, and hence introducing "competition" does not benefit the principal if the agents are self-interested. Note that in Case 2a, as the agents are more other-regarding, the optimal "prize" of the tournament ($\hat{w}_{sf}$) becomes smaller and hence the principal's expected utility increases.

However, the principal must compensate for the disutility from other-regarding preferences as shown in the left-hand side of (PC4). In Case 2b, in which the project outcome is very informative in terms of action choice or the agents are sufficiently other-regarding, the participation constraint binds. The principal then ceases to offer the extreme relative performance contract and chooses a contract in which the agents are paid positive amounts whether the project succeeds or fails.

Now compare the optimal contract in Proposition 17.4 with the optimal contract for the purely self-interested agents ($\alpha = 0$). As discussed above, there are usually many optimal contracts, and the independent contract $w_{ss} = w_{sf} = d/\Delta_p$ is one of them. However, a small amount of other-regarding preferences changes the optimal contract in an important way. The optimal contract is now generically unique, and independent contracts are no longer optimal (except for the insignificant cases), despite technological and stochastic independence. This result alerts the use of independent contracts in the analysis of agency models like ours, even if the agents are assumed to be purely self-interested.

4.3 Correlated Outcomes

One of the well known results from the principal-agent analysis with purely self-interested agents is the optimality of relative performance evaluation when the agents' performances are positively correlated (Holmstrom 1982; Mookherjee 1984). However, the main result in the previous section hints at the possibility that the optimal contract for the other-regarding agents may not be a relative performance contract in the correlated environment. To examine this possibility formally, in this section I extend the model to the case where the results of the projects are correlated.

The project of each agent either succeeds or fails depending not only on his action and idiosyncratic shock (as in the original model) but also on a common shock that affects both projects. The common shock is either good (with probability q) or bad (probability $1 - q$). If the common shock is good, then both projects will succeed irrespective of the agents' actions. However, if the common shock is bad, then the project outcome of each agent will depend on his action and idiosyncratic shock: his project succeeds with probability p_i and fails with probability $1 - p_i$ when his action is a_i, where p_i satisfy the same conditions as before. Each agent's project thus succeeds with probability $q + (1-q)p_i$. Che and Yoo (2001) show that in this setting the optimal contract is an extreme relative performance contract when the agents are purely self-interested.

Now consider other-regarding agents. Assuming $w_{fs} = w_{ff} = 0$, the incentive compatibility constraint and the participation constraint are, respectively,

$$(1-q)[p_1 w_{ss} + (1-p_1)w_{sf} + (p_1 - (1-p_1)\gamma)\alpha w_{sf}] \geq \frac{d}{\Delta_p} \tag{IC5}$$

$$(1-q)[p_1 w_{ss} + (1-p_1)w_{sf} - (1-p_1)\alpha(1+\gamma)w_{sf}] \geq \frac{d - q w_{ss}}{p_1}. \tag{PC5}$$

The principal's expected payments are $2[q w_{ss} + (1-q)p_1(p_1 w_{ss} + (1-p_1)w_{sf})]$. Note that if $q = 0$, this model coincides with the previous one with independent outcomes.

As with (17.21) and (17.22), we define w_{ss}^+ and w_{sf}^+ as follows:

$$w_{ss}^+ = \frac{d}{\Delta_p}\frac{1}{(1-q)p_1}$$

$$w_{sf}^+ = \frac{d}{\Delta_p}\frac{1}{(1-q)[(1-p_1) + \alpha(p_1 - (1-p_1)\gamma)]}.$$

Then $(w_{ss}^+, 0)$ is the extreme team contract and $(0, w_{sf}^+)$ is the extreme relative performance contract such that the incentive compatibility constraint (IC5) binds. And it turns out that if (17.23) holds, $(0, w_{sf}^+)$ satisfies (PC5) as well, and the principal's expected payments are the same as (17.24) which do not depend on q: with the extreme relative performance contract, the principal can filter out the common shock. Note that since $\alpha = 0$ satisfies (17.23), $(0, w_{sf}^+)$ with $\alpha = 0$ is the optimal contract for the purely self-interested agents, as derived by Che and Yoo (2001).

It is also easy to find that the extreme team contract $(w_{ss}^+, 0)$ also satisfies (PC5), and the principal's expected payments are given by

$$2[q + (1-q)p_1^2]w_{ss}^+ = \frac{2d}{\Delta_p}\frac{q + (1-q)p_1^2}{(1-q)p_1}. \tag{17.29}$$

Note that the expected payments are increasing in q. And we know from the previous analysis that if $q = 0$ (no correlation) and $(1-p_1)\gamma > p_1$ (Case 1), then the extreme team contract is optimal and hence the right-hand side of (17.29) is smaller than that of (17.24). Therefore, even in the positively correlated case, if q is sufficiently small, the extreme team contract is still optimal. This result is in contrast to the pure self-interest case.

4.4 Alternative Specification

I now consider the alternative specification that the agents compare their material payoffs net of the costs of actions, rather than their income. For example, employees who work closely may be able to monitor their actions each other; or, even if actions are not mutually observable, they may be able to expect the other agent's actions correctly, and hence their actions are likely to be compared. Agent 1's utility function then changes as follows:

$$u^1 = \begin{cases} w^1_{jk} - d_i - \alpha\gamma(w^1_{jk} - d_i - w^2_{kj} + d_h) & \text{if } w^1_{jk} - d_i \geq w^2_{kj} - d_h \\ w^1_{jk} - d_i - \alpha(w^2_{kj} - d_h - w^1_{jk} + d_i) & \text{if } w^1_{jk} - d_i \leq w^2_{kj} - d_h \end{cases} \tag{17.30}$$

where $j, k = s, f$ and $h, i = 0, 1$, and i (h) is the index for agent 1's (2's) action.

To see how this alternative specification alters the results, suppose that both agents choose a_1. Then their expected utility will not change from the previous model, and hence the participation constraint is the same as (PC4). However, if one of the agents, say agent 1, chooses a_0 while the other chooses a_1, then the comparison is not between w^1_{jk} and w^2_{kj} but between w^1_{jk} and $w^2_{kj} - d$. The incentive compatibility constraint is thus summarized as follows (given symmetric schemes):

$$\begin{aligned} & p_1 w_{ss} + (1 - p_1) w_{sf} + [p_1 - (1 - p_1)\gamma]\alpha w_{sf} \\ & \qquad \geq (1 - \alpha\gamma)\frac{d}{\Delta_p} + (1 - p_0) p_1 \alpha (1 + \gamma)\frac{w_{sf}}{\Delta_p} \qquad \text{if } w_{sf} \leq d \end{aligned} \tag{17.31}$$

$$\begin{aligned} & p_1 w_{ss} + (1 - p_1) w_{sf} + [p_1 - (1 - p_1)\gamma]\alpha w_{sf} \\ & \qquad \geq [1 - \alpha\gamma + (1 - p_0) p_1 \alpha (1 + \gamma)]\frac{d}{\Delta_p} \qquad \text{if } w_{sf} \geq d. \end{aligned} \tag{17.32}$$

First consider an extreme team contract $(w_{ss}, 0)$. Since $w_{sf} = 0 < d$, (17.31) applies and the incentive compatibility constraint is simplified as follows:

$$p_1 w_{ss} \geq (1 - \alpha\gamma)\frac{d}{\Delta_p}. \tag{17.33}$$

Under the extreme team contract, the agent is always ahead by d if he deviates from a_1 and chooses a_0 while the other agent follows a_1. If he is status-seeking ($\gamma < 0$), he will enjoy this deviation, and hence the principal must provide stronger incentives to induce him to choose a_1 under the current specification than under the original specification.

On the other hand, if the agents are inequity averse ($\gamma > 0$), they are more strongly motivated to choose a_1 in order to avoid being ahead, and hence the principal can save on the payments. The principal is better off because each agent feels guilty if he shirks while the other does not, and hence prefers to avoid such guilt. This nonpecuniary incentive replaces the monetary one. And the principal's expected payments are now *decreasing* in α under the extreme team contract; hence when the extreme team contract is adopted, the principal also prefers the agents to be more inequity averse. This result is in contrast to the result under the previous specification, i.e. that the principal implementing extreme team contracts is indifferent in the agents' other-regarding preferences.

Furthermore, if the agents are sufficiently inequity averse, then the participation constraint becomes binding and hence the principal need not leave rents to them: the participation constraint under extreme team contracts is $p_1 w_{ss} \geq d/\Delta_p$, the right-hand side of which is at least as large as the right-hand side of the incentive compatibility constraint (17.33) if and only if $\alpha\gamma \geq 1/\ell_s$.

The new incentive benefit from the agents' caring about their actions also exists under relative performance contracts. However, the benefit is not as large under relative performance contracts as under team contracts. For example, consider the extreme relative performance contract $(0, \hat{w}_{sf})$. A simple calculation shows $\hat{w}_{sf} > d$, and hence the incentive compatibility constraint (17.32) applies. The right-hand side of (17.32) is larger than that of (17.33) if $\gamma > -1$. Under $(0, \hat{w}_{sf})$, when the project of an agent fails and the other project succeeds, the agent is behind. When he is shirking, he compares his payoff zero with the other agent's payoff $\hat{w}_{sf} - d > 0$. It turns out that the extent to which he is behind is not as serious as in a similar situation with unobservable actions (where the difference is $\hat{w}_{fs}$), and hence the new specification brings a negative incentive effect under the extreme relative performance contract. Note that under the team contract the agent choosing a_0 will never be behind and hence this negative effect is absent.

It is therefore likely that the agents comparing their actions as well as their incomes make team contracts more attractive to the principal than relative performance contracts, and hence the extreme team contract is more likely to be optimal. In fact, I show the following results formally.

Proposition 17.5 *(i) If* $(1 - p_1)\gamma > p_1$ *holds, the optimal contract is an extreme team contract. (ii) If* $\alpha\gamma \geq 1/\ell_s$ *holds, the extreme team contract in which the participation constraint binds is optimal.*

Proof (i) Suppose that a contract (w_{ss}, w_{sf}) with $w_{sf} > 0$ is optimal. Define w'_{ss} by $p_1 w'_{ss} = p_1 w_{ss} + (1 - p_1)w_{sf} - \varepsilon$ and consider the extreme team contract $(w'_{ss}, 0)$, where $\varepsilon > 0$. It is easy to see that the new contract satisfies the incentive compatibility constraint (17.31) or (17.32), and the participation constraint (PC4) for ε sufficiently small but positive. And the principal's expected payments are $2p_1^2 w'_{ss} < 2p_1[p_1 w_{ss} + (1 - p_1)w_{sf}]$, which contradicts the optimality of (w_{ss}, w_{sf}). (ii) The discussion preceding the proposition shows that the condition given is necessary and sufficient for the extreme team contract under which the participation constraint binds to satisfy the incentive compatibility constraint (17.33) as well. The principal's expected payments are $2d$. Now consider an arbitrary contract (w_{ss}, w_{sf}). The principal's expected payments are

$$2p_1[p_1 w_{ss} + (1-p_1)w_{sf}] \geq 2p_1 \left(\frac{d}{p_1} + (1-p_1)\alpha(1+\gamma)w_{sf} \right) \geq 2d.$$

The first inequality follows from (PC4). The second inequality holds since $\gamma \geq -1$. (It is strict if $w_{sf} > 0$ and $\gamma > -1$.) The extreme team contract is thus optimal. □

Proposition 17.5, along with Proposition 17.4, implies the following. (i) If an extreme team contract is optimal under the original specification, then the optimal contract is again an extreme team contract under the alternative specification. (ii) There exists a range of parameter values in which, although the optimal contract is not an extreme team contract under the original specification, the extreme team contract becomes optimal under the new specification. The conditions are given by $\alpha\gamma \geq 1/\ell_s$ and $(1-p_1)\gamma < p_1$. It is possible for both of them to hold if $\alpha > p_0(1-p_1)/p_1^2$. In this case, the nonpecuniary incentive is so strong that the incentive compatibility constraint can be ignored and the principal can do as well as if the agents were self-interested and their actions were enforceable.

5 Implications for Endogenous Preferences in Organizations

The introduction of other-regarding preferences into contract theory enables us to analyze how those preferences affect the optimal contract and, more importantly, what kinds of preferences the principal wants the agents to have. In other words, preferences can be part of the "contract" designed by the principal, as Perrow (1986) cited in Sect. 1 above suggests.[22] In this regard, it is interesting to find that even Milgrom and Roberts (1992) take a sympathetic position:

> Furthermore, important features of many organizations can be best understood in terms of deliberate attempts to change the preferences of individual participants to make these factors [such as altruism, exceedingly high regard for others' opinions of one's courage] more salient. As a result, organizationally desired behavior becomes more likely. (Milgrom and Roberts 1992, p. 42)

Although the current paper is just a start and takes the agents' preferences as exogenous, the results provide some interesting implications for desirable preferences.

First, consider the single-agent case. The main result is that the principal in general does not benefit from other-regarding preferences. If the benefit from the successful project is large, the principal prefers a self-interested agent to an inequity averse or status-seeking agent.[23] In particular, the optimal incentive for the self-

[22]Alternatively, the agent could change his preferences strategically. Rotemberg (1994) takes such an approach in studying organizational behavior.

[23]Note that I have not covered purely altruistic agents. Actually modifying the range of the parameter values in the model allows the agent to be unconditionally altruistic (Rey Biel 2003). I call the agent *altruistic* or *efficiency-seeking* if $\alpha < 0$ and $\gamma \leq -1$: the agent's utility is then increasing in the principal's payoff as well as in his income. If $\gamma = -1$, he is purely altruistic in

interested agent does not induce the other-regarding agent to choose a_1: more costly, higher-powered incentives are necessary.

If the benefit from the successful project is small, there may be a case in which the principal benefits from having a more status-seeking agent, since he is motivated to choose a_1 to raise the possibility of earning more than the principal: lower-powered monetary incentives thus suffice.

Next consider the multi-agent case. Suppose that the principal can choose a contract as well as (α, γ) and the agents compare only what they are paid. The principal then wants the agents to have γ as small as possible and α sufficiently large. For example, if only inequity-averse agents ($\gamma \geq 0$) are feasible, then the optimal preference for the principal is $\gamma = 0$ and $\alpha > \ell_f/\ell_s$, i.e. the agent who feels sufficiently envious but does not feel guilty at all. If status-seeking agents are feasible as well, then the most competitive ($\gamma = -1$) and sufficiently other-regarding agent is the best. Note that Case 2b in Proposition 17.4 applies and hence the optimal contract is neither extreme team nor extreme relative performance.

One reason why such a competitive preference pattern is desirable is a lack of productive interaction among agents in the model. However, the analysis reveals another interesting reason. If the principal can change the agents' preferences such that they compare their actions as well as the payments, she will prefer implementing the extreme team contract for the agent with sufficiently inequity-averse preferences (such that $\alpha\gamma \geq 1/\ell_s$, as in Proposition 17.5). Thus, even though the agents work independently, implementation of a fair, team-based pay scheme may benefit the principal if she can change their preferences in a deliberate manner.

6 Concluding Remarks

I have argued that incentives and other-regarding preferences interact in nontrivial ways. When an agent cares about the principal's wellbeing, the principal is in general worse off by having a more inequity-averse agent. When there are two or more symmetric agents who care about each other's wellbeing, the principal can optimally exploit their other-regarding nature by designing an appropriate interdependent contract, such as a "fair" team contract or a relative performance contract, that creates inequality when their performance outcomes are different. The optimal contract depends on the nature of the agents' other-regarding preferences.

I believe behavioral contract theory is a fruitful approach to issues in organization. The approach taken in this chapter can shed light on issues on endogenous pref-

---

the sense that his utility is identical whether he is behind or ahead. If $\gamma < -1$, he emphasizes his income more and the principal's payoff less when he is behind than when he is ahead. It is easy to show that the principal benefits from a more efficiency-seeking agent (with higher $|\alpha|$), for the nonpecuniary incentive to choose a_1 enables the principal to save the monetary incentive. And if the agent is sufficiently efficiency-seeking, no monetary incentive is needed and the participation constraint binds.

erences, as suggested by sociologists and even organizational economists. Future research should deal with organizational contexts or problems more explicitly, to determine how various aspects of organizations affect the members' preferences. Another promising research theme is how members with various preferences should be grouped. To this end, extending the analysis to the cases not considered in the current paper, such as those of productive externalities, asymmetric agents, and more than two agents, is high on the list.

Acknowledgements I am grateful to Munetomo Ando, Shingo Ishiguro, Shinsuke Kambe, Michihiro Kandori, Hitoshi Matsushima, Masahiro Okuno-Fujiwara, Pedro Rey Biel, Takashi Shimizu, Katsuya Takii, an anonymous referee, and the participants at Contract Theory Workshop (CTW), Contract Theory Workshop East (CTWE), and University of Tokyo Microeconomics Workshop for helpful comments. The research for this chapter was supported by Grant-in-Aid for Scientific Research (KAKENHI 13630005) from JSPS and MEXT of the Japanese Government.

Appendix

In this appendix I show that in the model in Sect. 4, it is without loss of generality to restrict attention to contracts with $w^n_{fs} = w^n_{ff} = 0$. To this end, consider a symmetric contract $(w, w) \in C$ in which at least one equality of $w_{fs} \geq 0$ and $w_{ff} \geq 0$ is strict. Note that to simplify the notation I drop the superscripts from the contracts. If $w_{sf} \geq w_{fs}$, the incentive compatibility constraint and the participation constraint are, respectively,

$$(w_{sf} - w_{ff}) + p_1(w_{ss} + w_{ff} - w_{sf} - w_{fs}) + [p_1 - (1 - p_1)\gamma]\alpha(w_{sf} - w_{fs}) \geq \frac{d}{\Delta_p} \tag{17.34}$$

$$\begin{aligned} w_{ff} &+ p_1(w_{sf} - w_{ff}) + p_1(w_{fs} - w_{ff}) + p_1^2(w_{ss} + w_{ff} - w_{sf} - w_{fs}) \\ &- (1 - p_1)p_1\alpha(1 + \gamma)(w_{sf} - w_{fs}) \geq d. \end{aligned} \tag{17.35}$$

Similarly, if $w_{sf} < w_{fs}$, they are

$$(w_{sf} - w_{ff}) + p_1(w_{ss} + w_{ff} - w_{sf} - w_{fs}) + [p_1\gamma - (1 - p_1)]\alpha(w_{fs} - w_{sf}) \geq \frac{d}{\Delta_p} \tag{17.36}$$

$$\begin{aligned} w_{ff} &+ p_1(w_{sf} - w_{ff}) + p_1(w_{fs} - w_{ff}) + p_1^2(w_{ss} + w_{ff} - w_{sf} - w_{fs}) \\ &- (1 - p_1)p_1\alpha(1 + \gamma)(w_{fs} - w_{sf}) \geq d. \end{aligned} \tag{17.37}$$

The expected payment to each agent is

$$W = w_{ff} + p_1(w_{sf} - w_{ff}) + p_1(w_{fs} - w_{ff}) + p_1^2(w_{ss} + w_{ff} - w_{sf} - w_{fs}). \quad (17.38)$$

Denote by $C^* \subset C$ the set of feasible contracts that satisfy the incentive compatibility constraint and the participation constraint.

Lemma 17.2 *For a given contract* $(w, w) \in C^*$ *with* $w_{sf} < w_{fs}$*, there exists a contract* $(w', w') \in C^*$ *such that* $w'_{sf} > w'_{fs}$ *holds and the principal's expected payment under* (w', w') *is the same as that under* (w, w).

Proof Define the new contract by $w'_{ss} = w_{ss}$, $w'_{ff} = w_{ff}$, $w'_{sf} = w_{fs}$, and $w'_{fs} = w_{sf}$. Obviously, the new contract satisfies $w'_{sf} > w'_{fs}$ and the participation constraint. And the incentive compatibility constraint (17.34) is satisfied because (w, w) satisfies (17.36) and $p_1 - (1 - p_1)\gamma \geq p_1\gamma - (1 - p_1)$ holds for $\gamma \leq 1$. Finally, the expected payment under the new contract is equal to W. □

By Lemma 17.2, from now on I focus on contracts with $w_{sf} \geq w_{fs}$. Define a new pay scheme for each agent, $\hat{w} = (\hat{w}_{jk})$, that satisfies $\hat{w}_{fs} = \hat{w}_{ff} = 0$, $\hat{w}_{sf} = w_{sf} - w_{fs}$, and

$$p_1\left(\hat{w}_{sf} + p_1(\hat{w}_{ss} - \hat{w}_{sf})\right) = W. \quad (17.39)$$

Condition (17.39) implies that the principal's expected payment under the new contract $(\hat{w}, \hat{w})$ is equal to that under (w, w). It is easy to show $\hat{w}_{ss} > 0$.

I next show that the new contract $(\hat{w}, \hat{w})$ satisfies the participation constraint

$$p_1\left(\hat{w}_{ss} + p_1(\hat{w}_{ss} - \hat{w}_{sf})\right) - p_1(1 - p_1)\alpha(1 + \gamma)\hat{w}_{sf} \geq d. \quad (17.40)$$

The left-hand side is equal to

$$W - p_1(1 - p_1)\alpha(1 + \gamma)(w_{sf} - w_{fs})$$

which is the left-hand side of (17.35). Thus by (17.35), (17.40) holds and the new contract satisfies the participation constraint.

Finally, the incentive compatibility constraint is written as follows:

$$\hat{w}_{sf} + p_1(\hat{w}_{ss} - \hat{w}_{sf}) + [p_1 - (1 - p_1)\gamma]\alpha\hat{w}_{sf} \geq \frac{d}{\Delta_p}. \quad (17.41)$$

By (17.39) and $\hat{w}_{sf} = w_{sf} - w_{fs}$, the left-hand side of (17.41) is equal to

$$\frac{W}{p_1} + (p_1 - (1 - p_1)\gamma)\alpha(w_{sf} - w_{fs}). \quad (17.42)$$

Since $w_{fs} > 0$ or $w_{ff} > 0$ holds, it is easy to show that

$$\frac{W}{p_1} > (w_{sf} - w_{ff}) + p_1(w_{ss} + w_{ff} - w_{sf} - w_{fs}). \tag{17.43}$$

Therefore by (17.43), (17.42), and (17.34), the new contract satisfies the incentive compatibility constraint (17.41). The result is summarized in the following proposition.

Proposition 17.6 *For a contract* $(w, w) \in C^*$ *that satisfies* $w_{fs} > 0$ *or* $w_{ff} > 0$*, there exists a contract* $(\hat{w}, \hat{w}) \in C^*$ *such that* $\hat{w}_{fs} = \hat{w}_{ff} = 0$ *and the principal's expected payment is equal to that under* (w, w)*.*

Addendum: Revisiting Moral Hazard and Other-Regarding Preferences[24]

> I would conclude that if behavioral economists want their revolution to occur, they might be well served to focus on producing applied theory papers that economists in various fields will want to teach their students.
>
> Glenn Ellison[25]

When I was writing a paper that was eventually published as Itoh (2004), I had never heard nor seen term "behavioral contract theory." There were already several psychology-based individual decision making models[26]

that were developed following anomalies found in various experiments, and fields like "behavioral game theory" and "behavioral finance" were emerging. Partly inspired by Glenn Ellison's discussion, I attempted to extend the standard model of the principal-agent relationship with hidden action by applying a particular theory of "caring-about-others" preferences, i.e. inequity aversion. My paper was one of the earliest in this direction, and fortunately has been well cited, provided that it is published in *Japanese Economic Review*.

Today research in behavioral contract theory is rapidly growing, as evidenced by the recent survey titled "behavioral contract theory" (Kőszegi 2014), that is

[24]This addendum has been newly written for this book chapter.

[25]His discussion for the invited session "Behavioral Economics" at the eighth World Congress of the Econometric Society (Ellison 2003, p. 300).

[26]My favorite way to summarize these models is based on the presentation by Rabin (2002): people departure from standard assumptions by (i) caring about changes, such as reference dependence and loss aversion (Kahneman and Tversky 1979; Kőszegi and Rabin 2006, 2007; Tversky and Kahneman 1991); (ii) caring about others, such as inequity aversion (Bolton and Ockenfels 2000; Fehr and Schmidt 1999) and reciprocity (Rabin 1993); (iii) caring about now, such as present-biased preferences (Laibson 1997; Strotz 1955–1956); and more recently, (iv) caring about self-image (Bénabou and Tirole 2006).

forthcoming in *Journal of Economic Literature*. Recent theoretical work in the field incorporates not only inequity aversion but also other theories of individual decision making successfully into contract theory to address important issues on contract design (Bénabou and Tirole 2006; Ellingsen and Johannesson 2008; Englmaier and Leider 2012; Herold 2010; Herweg et al. 2010; Sliwka 2007), and I want to urge those interested in learning recent development in behavioral contract theory to read Kőszegi (2014). My modest purpose of this addendum is to revisit Itoh (2004) to supplement the part of Kőszegi (2014) that discusses inequity aversion and its applications to contract theory.

Kőszegi (2014) is partly motivated by results from laboratory experiments on gift exchange and discretionary bonus, such as Fehr et al. (2007), and shows nicely that inequity aversion itself works as an incentive device in the sense that the first-best action can be induced by either a fixed non-contingent wage or a voluntary bonus. On the other hand, Itoh (2004) is motivated to introduce inequity aversion into the standard principal-agent model with limited liability in which, in contrast to such experiments, the agent's binary action only stochastically determines output.

The more important difference comes from the fact that the potential application I have in mind is to large organizations, that leads to the following two features of my analysis. First, in the single agent case, the agent does not take into account the cost of action when comparing his income with the principal's. While it is reasonable to assume that the agent cares about his relative net income in laboratory experiments, employees in large organizations are not likely to compare their net incomes (wages minus the cost of action) to the income of the employer.

Second, the main focus of Itoh (2004) is on the analysis of the multi-agent case. Particularly in large organizations, employees tend to compare themselves with those with the same job titles, pay ranks, status, and so on. That is why I analyze a multi-agent setting in which each agent cares about his fellow agent rather than the principal.

Furthermore, comparing (the cost of) his action with (that of) the other agent is natural, and I show that the optimal "contracts" depend critically on whether or not the agents compare their actions, in the following sense. When the agents do not compare their actions, the main result is Proposition 17.4. Let me restate it in terms of α and $\beta = \alpha\gamma$ as Proposition 17.7 given below, since Proposition 17.4 of Itoh (2004) contains minor typos.

Proposition 17.7 *The optimal contract $w^* = (w^*_{ss}, w^*_{sf})$ is given as follows.*

Case 1: $w^* = (\hat{w}_{ss}, 0)$ *if* $\beta > \alpha p_1/(1 - p_1)$ *holds. It is an extreme team contract.*
Case 2a: $w^* = (0, \hat{w}_{sf})$ *if both* $\beta < \alpha p_1/(1 - p_1)$ *and* $\beta \leq 1 - \alpha(\ell_s/\ell_f)$ *hold. It is an extreme relative performance contract.*
Case 2b: $w^* = (\overline{w}_{ss}, \overline{w}_{sf})$ *if both* $\beta < \alpha p_1/(1 - p_1)$ *and* $\beta > 1 - \alpha(\ell_s/\ell_f)$ *hold.*

Figure 17.1 given below summarizes the optimal contract for the case of $p_1 < 0.5$. Case 1 applies to the grey region, Case 2a to the region with "extreme relative," and Case 2b covers both regions with "relative" and "team" where a relative

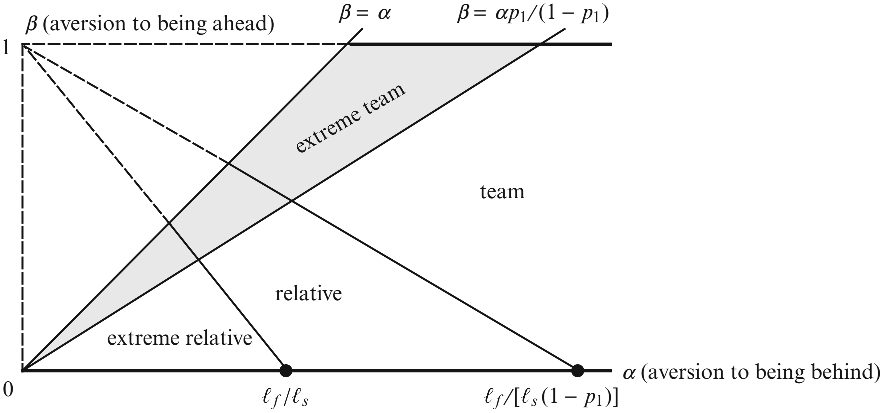


Fig. 17.1 The optimal contract in Proposition 17.7 (case $p_1 < 0.5$)

performance contract ($\overline{w}_{ss} < \overline{w}_{sf}$) or a team contract ($\overline{w}_{ss} > \overline{w}_{sf}$), respectively, is optimal.

Importantly, the principal's expected payments to the agents are the highest and do not depend on (α, β) in the grey region where an extreme team contract is optimal. When the agents compare their actions as well, the optimal contract is still an extreme team contract in the grey region (Proposition 17.5). However, the principal's expected payments are now the *lowest* in that region, and sometimes at the first-best level. Inequity-averse agents can thus benefit the principal only if they compare their actions.

I should note that my analysis is subject to the critical comments made by Kőszegi (2014). First, the choice of the reference group is still exogenous, and an important future research theme is to understand how each agent chooses his reference group. Second, although one of my interests is in the choice of the agents' preferences by the principal, "comparative statics with respect to variables typically studied in economic analysis" are still missing. And finally, my focus on inequity aversion is obviously restrictive for the study of teams. Each agent working in a team is motivated by team-based monetary incentives, private benefits, reciprocity, self-image from public, self-image from his fellow agents, and so on. How these forces interact has to be carefully studied for our further understanding of teams.

References

Baron JN (1988) The employment relation as a social relation. J Jpn Int Econ 2:492–525

Baron JN, Kreps DM (1999) Strategic human resources: framework for general managers. Wiley, New York

Bénabou R, Tirole J (2006) Incentives and prosocial behavior. Am Econ Rev 96:1652–1678

Bewley TF (1999) Why wages don't fall during a recession. Harvard University Press, Cambridge

Bolton GE, Ockenfels A (2000) ERC: a theory of equity, reciprocity, and competition. Am Econ Rev 90:166–193
Camerer CF (2003) Behavioral game theory. Princeton University Press, Princeton
Charness G, Haruvy E (2002) Altruism, equity and reciprocity in a gift-exchange experiment: an encompassing approach. Games Econ Behav 40:203–231
Charness G, Rabin M (2002) Understanding social preferences with simple tests. Q J Econ 117:817–869
Che YK, Yoo SW (2001) Optimal incentives for teams. Am Econ Rev 91:525–541
Cox JC, Friedman D (2002) A tractable model of reciprocity and fairness. Mimeo, University of Arizona
Dufwenberg M, Kirchsteiger G (2002) A theory of sequential reciprocity. Mimeo, Stockholm University
Ellingsen T, Johannesson M (2008) Pride and prejudice: the human side of incentive theory. Am Econ Rev 98:990–1008
Ellison G (2003) A discussion of the papers by Ernst Fehr and Klaus M. Schmidt and by Christopher Harris and David Laibson. In: Dewatripont M, Hansen LP, Turnovsky SJ (eds) Advances in economics and econometrics: theory and applications, eighth world congress, vol I. Cambridge University Press, Cambridge, pp 298–301 Seattle, USA
Englmaier F, Leider S (2012) Contractual and organizational structure with reciprocal agents. Am Econ J Microecon 4:146–183
Englmaier F, Wambach A (2002) Contracts and inequity aversion. Mimeo, University of Munich
Falk A, Fischbacher U (2000) A theory of reciprocity, working Paper No.6, Institute for Empirical Research in Economics, University of Zurich
Fehr E, Fischbacher U (2002) Why social preferences matter—the impact of non-selfish motives on competition, cooperation and incentives. Econ J 112:C1–C33
Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. Q J Econ 114:817–868
Fehr E, Schmidt KM (2003) Theories of fairness and reciprocity: evidence and economic applications. In: Dewatripont M, Hansen LP, Turnovsky SJ (eds) Advances in economics and econometrics: theory and applications, eight world congress, vol I, chap 6. Cambridge University Press, Cambridge, pp 208–257 Seattle, USA
Fehr E, Klein A, Schmidt KM (2001) Fairness, incentives and contractual incompleteness. Mimeo, University of Zurich
Fehr E, Klein A, Schmidt KM (2007) Fairness and contract design. Econometrica 75:121–154
Forsythe R, Horowitz JL, Savin NE, Sefton M (1994) Fairness in simple bargaining experiments. Games Econ Behav 6:347–369
Geanakoplos J, Pearce D, Stacchetti E (1989) Psychological games and sequential rationality. Games Econ Behav 1:60–79
Grund C, Sliwka D (2002) Envy and compassion in tournaments. Mimeo, University of Bonn
Güth W, Klose W, Königstein M, Schwalbach J (1998) An experimental study of a dynamic principal-agent relationship. Manag Decis Econ 19:327–341
Herold F (2010) Contractual incompleteness as a signal of trust. Games Econ Behav 68:180–191
Herweg F, Müller D, Weinschenk P (2010) Binary payment schemes: moral hazard and loss aversion. Am Econ Rev 100:2451–2477
Holmstrom B (1982) Moral hazard in teams. Bell J Econ 13:324–340
Holmstrom B, Milgrom P (1990) Regulating trade among agents. J Inst Theor Econ 146:85–105
Holmstrom B, Milgrom P (1991) Multitask principal-agent analyses: incentive contracts, asset ownership, and job design. J Law Econ Organ 7(special issue):24–52
Itoh H (1993) Coalitions, incentives, and risk sharing. J Econ Theory 60:410–427
Itoh H (2004) Moral hazard and other-regarding preferences. Jpn Econ Rev 55:18–45
Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica 47:263–291
Keser C, Willinger M (2000) Principals' principles when agents' actions are hidden. Int J Ind Organ 18:163–185

Kőszegi B (2014) Behavioral contract theory. J Econ Lit 52(4): 1075–1118
Kőszegi B, Rabin M (2006) A model of reference-dependent preferences. Q J Econ 121:1133–1165
Kőszegi B, Rabin M (2007) Reference-dependent risk attitudes. Am Econ Rev 97:1047–1073
Laibson D (1997) Golden eggs and hyperbolic discounting. Q J Econ 112:443–477
Lazear EP (1995) Personnel economics. MIT, Cambridge
Levine DK (1998) Modeling altruism and spitefulness in experiments. Rev Econ Dyn 1:593–622
Milgrom P, Roberts J (1992) Economics, organization and management. Prentice-Hall, Englewood Cliffs
Mookherjee D (1984) Optimal incentive schemes with many agents. Rev Econ Stud 51:433–446
Nalbantian HR, Schotter A (1997) Productivity under group incentives: an experimental study. Am Econ Rev 87:314–341
Neilson WS (2002) An axiomatic characterization of the Fehr-Schmidt model of inequity aversion. Mimeo, Texas A&M University
Neilson WS, Stowe J (2003) Incentive pay for other-regarding workers. Mimeo, Texas A&M University
Perrow C (1986) Complex organizations: a critical essay, 3rd edn. Random House, New York
Pfeffer J (1997) New directions for organization theory: problems and prospects. Oxford University Press, New York
Rabin M (1993) Incorporating fairness into game theory and economics. Am Econ Rev 83:1281–1302
Rabin M (2002) A perspective on psychology and economics. Eur Econ Rev 46:657–685
Rey Biel P (2003) Inequity aversion and team incentives. Mimeo, University College London
Rotemberg JJ (1994) Human relations in the workplace. J Politi Econ 102:684–717
Segal U, Sobel J (1999) Tit for tat: foundations of preferences for reciprocity in strategic settings, discussion Paper 99–10, Department of Economics, UC San Diego
Sliwka D (2007) Trust as a signal of a social norm and the hidden costs of incentive schemes. Am Econ Rev 97:999–1012
Sobel J (2001) Interdependent preferences and reciprocity. Mimeo, University of California, San Diego
Strotz RH (1955–1956) Myopia and inconsistency in dynamic utility maximization. Rev Econ Stud 23(3):165–180
Tirole J (1992) Collusion and the theory of organizations. In: Laffont JJ (ed) Advances in economic theory: sixth world congress, vol II, chap 3. Cambridge University Press, Cambridge, pp 151–206 Barcelona, Spain
Tversky A, Kahneman D (1991) Loss aversion in riskless choice: a preference-dependent model. Q J Econ 106:1039–1061

Chapter 18
Contracting with Self-Esteem Concerns

Junichiro Ishida

Abstract It is widely accepted in social psychology that the need to maintain and enhance self-esteem is a fundamental human motive. We incorporate this factor into an otherwise ordinary principal-agent framework and examine its impact on the optimal incentive scheme and the agent's behavior, especially focusing on the form of intrapersonal strategy known as self-handicapping. Incorporating self-esteem concerns into a contracting situation yields an implication that runs counter to conventional wisdom; that is, the standard tradeoff between risk and incentives may break down (i.e., more uncertainty reduces agency cost and hence results in stronger incentives) in the presence of self-esteem concerns. This is because uncertainty mitigates the need for self-handicapping. This result provides a possible reason for why we do not empirically observe this tradeoff in a robust manner. We present an intuitive condition for this anomaly to arise and provide a set of testable implications. The present framework also reveals why and how team production can be made more profitable by providing an explanation for the increasing popularity of team production. Finally, this simple logic is applied to identify additional implications for the hidden costs of external enforcers, such as evaluation and monitoring, which are discussed extensively in social psychology.

Keywords Self-esteem • Bayesian learning • Tradeoff between risk and incentives • Contract

JEL Classification Codes D81, D86

The original article first appeared in Journal of Economic Behavior & Organization 81: 329–340, 2012. A newly written addendum has been added to this book chapter.

J. Ishida (✉)
Institute of Social and Economic Research, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan
e-mail: jishida@osipp.osaka-u.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_18

1 Introduction

Most economic analyses typically assume that agents possess precise knowledge of their own attributes. As this assumption symbolically suggests, economists have paid relatively little attention to the potential consequences of learning about oneself. This aspect of economics is in clear contrast to psychology, where the issue of how people come to understand themselves in general, or of self-esteem in particular, has always been a topic of utmost concern. Although self-esteem is defined in various ways, including in terms of perceived ability and competence, physical attractiveness and interpersonal relations, it is now widely accepted in social psychology that the need to maintain and enhance self-esteem is a fundamental human motive (Leary and Downs 1995).

Recently, interest has been growing among economists on the issue of self-esteem concerns (e.g., Benabou and Tirole 2002). In this chapter, we explore the implications of self-esteem concerns in a contracting situation by following this line of research from a different perspective. To this end, we begin our analysis with two key presumptions: (i) the agent in question is a rational Bayesian learner, but with only vague self-knowledge to begin with; and (ii) the agent benefits directly from having a positive self-image, for various reasons. Under this setup, an outcome of some action, which forms a signal of the agent's own unknown attributes, has a direct effect on payoffs, aside from compensation from the agreed-upon contract. We find that the presence of self-esteem concerns plays a critical role especially when ability and effort are complementary, as assumed in many economic analyses. Under this condition, the agent actually stretches the distribution of outcome by exerting effort, thereby facilitating the process of learning about oneself. Conversely, by not exerting effort, and hence compressing the distribution of outcomes, failure becomes relatively insignificant as the agent can still remain vague about his or her own attributes. Put differently, even if the task turns out to be unsuccessful, the agent can still attribute the outcome to a lack of effort, rather than to a lack of ability.

Suppose that the agent is relatively averse to having a negative self-image or, more precisely, the utility gain from self-esteem is concave in level.[1] In such a case, and given the same mean, the agent is actually better off by not obtaining more information and hence remaining vague in terms of self-knowledge. This motive gives rise to a form of intrapersonal strategy, known in social psychology as self-handicapping, where individuals create impediments to successful performance. In some instances, self-handicapping takes a highly active form where performance-impairing drugs or alcohol are taken before an achievement activity such as an

[1]The shape of the utility function for self-esteem is certainly a matter of debate, and we acknowledge the possibility of some individual heterogeneity; for instance, some people may be risk loving in self-image. See Cowen and Glazer (2007) for a similar assumption. A body of psychological evidence on self-handicapping behavior appears to suggest, however, that it is reasonable to assume that people are, on average, risk averse in self-image.

exam[2]; in others, it takes a more passive form where individuals intentionally withhold effort in order to make failure meaningless (e.g., not studying hard before the exam). In either case, people resort to self-handicapping as a strategy to shield themselves from "facing the truth", which sometimes can be painful. In this chapter, we mostly focus on the role of the latter form of (passive) self-handicapping in a principal–agent framework.[3]

Incorporating self-esteem concerns into a contracting situation yields an implication that runs counter to conventional wisdom, enabling us to account for several stylized anomalies. We first show that the tradeoff between risk and incentives, which is a central tenet of contract theory, breaks down in the presence of self-esteem concerns; that is, more uncertainty may actually decrease agency cost and hence benefit the principal. In the standard moral hazard model, the cost of inducing any given level of effort generally increases with the extent of uncertainty faced by the agent. This general conclusion suggests that because agency cost increases with the uncertainty of the underlying environment, incentive pay would be subdued in risky situations. However, despite this clear prediction, identifying this tradeoff using actual data is a daunting task, as emphasized in a series of papers by Prendergast (2000, 2002a,b).[4]

In these works, Prendergast (2000, 2002a,b) provides several plausible reasons why we may not observe this tradeoff by focusing on the neglected impact of uncertainty on various aspects of monitoring behavior and/or the allocation of responsibility.[5] The basic stance is that there are some hidden aspects of uncertainty that tend to be overlooked, and these influence the optimal form of incentive structure. In contrast, the present chapter points out a channel through which the presence of uncertainty itself directly yields a positive incentive effect that weakens or even breaks down the standard tradeoff.[6] The main contention of our study is that uncertainty may yield a positive incentive effect because the uncertainty associated with the task obscures the agent's true worth, and consequently reduces the need for self-handicapping. As the need for self-handicapping naturally increases the agency

---

[2]For instance, Tucker, Vuchinich and Sobel (1981) find that college students who are unsure of performing well in an intellectual task voluntarily drink alcohol before the task.

[3]In what follows, we refer to self-handicapping simply as the act of withholding effort in an attempt to obscure one's own attributes.

[4]See Jin (2002) and Core and Guay (2002) for more recent empirical evidence.

[5]Several alternative explanations have also been suggested. For instance, Ackerberg and Botticini (2002) argue that the standard tradeoff may disappear if endogenous matching is explicitly considered. Alternatively, Raith (2003) emphasizes the impact of market competition on the provision of managerial incentives.

[6]Our stance here is that the presence of self-esteem concerns possibly provides one factor, perhaps among many others, that can break down the standard tradeoff. We should thus note that the purpose of this chapter is simply to complement, but certainly not to deny, existing views on this issue.

cost to induce any given level of effort, an increase in uncertainty may actually decrease the agency cost and hence result in stronger incentives. This is the case even when the agent in question is a rational Bayesian learner with no ability to manipulate information. Along the way, our analysis, and especially the reasoning detailed above, gives rise to a set of testable implications on the tradeoff between risk and incentives.

Although the main purpose of this chapter is to illustrate that greater uncertainty may benefit the principal, we also make the point that our main result is not a mere academic curiosity, as it yields several implications of practical importance for the optimal design of organizations. In this situation, given that the principal may benefit from uncertain environments and/or ambiguous outcomes, there indeed arises an incentive for her to artificially introduce noise, if it is technically feasible to do so, into the evaluation process. Building on this logic, the present analysis then reveals, as one of the implications, how and why team production can be made more profitable than individual production. In short, team production can be profitable because the outcome of team projects is inherently less reflective of each individual's attributes, whereas under individual production, the outcome more accurately reflects these attributes. As the need for self-handicapping is lower under team production, the cost of inducing effort can also be less, consequently making team production more profitable. We argue that this logic provides an explanation, more so from the psychological perspective, for the fact that firms are increasingly adopting team production as a means to motivate their workers.[7] We later employ similar reasoning to identify the implications for the hidden costs of external enforcers, such as evaluation and monitoring, which are discussed rather extensively in social psychology.

As noted above, a growing body of literature incorporates psychological factors in general, and the imperfectly known self in particular, into formal economic analyses. Two main strands of literature explore the consequences of learning about oneself. The first strand deals with interpersonal situations, which can be seen as a particular type of the informed-principal problem. For instance, Benabou and Tirole (2003) analyze how disclosing information about the agent (or the surrounding environment) might affect the agent's "intrinsic motivation".[8] Ishida (2006b) further extends this idea and analyzes optimal promotion rules when the agent is ambiguous about his or her own attributes and gains some information about him or herself from the principal's behavior.[9]

Conversely, the second strand of the literature deals with intrapersonal situations. In this body of work, Benabou and Tirole (2002) broadly analyze the economic consequences of self-esteem concerns by focusing mainly on the problems arising from

---

[7]See, for instance, Che and Yoo (2001) and the references therein for this trend in team production.

[8]Benabou and Tirole (2006) consider a different context where an agent possesses private information about his or her tastes and attempts to signal them in order to appear "prosocial". They then show how this reputation effect is contaminated by extrinsic rewards.

[9]See also Swank and Visser (2007) for an application of this idea.

time-inconsistent preferences. Using intrapersonal games, they analyze issues such as why self-confidence can be beneficial, and how the acquisition of information alters the behavior of future selves.[10] Benabou and Tirole (2004) combine imperfect recall with time-inconsistent preferences and develop a theory of personal rules arising from self-reputation concerns. Within a similar framework to the present study (unbiased Bayesian learning and self-esteem concerns), Koszegi (2006) shows how self-esteem concerns give rise to biased beliefs, even for rational Bayesian learners, and distort the choice of task. Similarly, Cowen and Glazer (2007) consider how the demand for esteem affects job choice and draw implications for various labor market concerns. The present chapter belongs to this second strand of the literature where the agent gains some information about him or herself through his or her own actions (self-learning through experimentation). As a point of departure from previous work, we focus on a contracting situation and analyze how self-esteem concerns affect the optimal incentive scheme, especially in relation to the tradeoff between risk and incentives, in order to derive some practical implications for organizational design.

The present setup is also closely related to models of career concerns, most notably Holmstrom (1999), in that it deals with dynamic signaling incentives.[11] The distinction between the present analysis and this strand of the literature is subtle and critical. In brief, in career concern models, signaling is directed to outside observers who have only a limited set of information about the sender. Career concerns then typically yield a positive incentive effect because the agent can gain nothing by withholding effort when outside observers expect him to do otherwise. This establishes a clear contrast with the present framework in which the focus is placed on "self-learning through experimentation" and the agent's information set, upon which the posterior belief is conditioned, naturally includes his own actions. As we show shortly, this apparently minor difference amounts to different implications where self-esteem concerns actually yield a negative incentive effect under some conditions, ultimately working to counteract the standard tradeoff between risk and incentives.

The chapter proceeds as follows. Section 2 illustrates the environment where we introduce self-esteem concerns into an otherwise ordinary principal–agent model. Section 3 derives the main results and shows that self-esteem concerns have an effect that tends to negate the standard tradeoff between risk and incentives. Section 4 extends the model and shows that the same logic also works for team production, thus providing an explanation for why team production can be so profitable. Finally, Sect. 5 makes some concluding remarks.

[10]See also Carrillo and Mariotti (2000).

[11]See also Dewatripont, Jewitt and Tirole (1999a,b).

2 The Model

We illustrate our contention using a simple moral hazard model with discrete effort choices and linear contracts.[12] The main purpose here is to make the point that the agency cost to induce any given level of effort may decrease with an increase in the extent of uncertainty of the underlying environment, and this lies contrary to the conventional wisdom.

Consider a model with a principal (female) and an agent (male). The agent is characterized by the ability type $\eta \sim N(\mu, \sigma_\eta^2)$, $\mu > 0$, which is initially not known to anyone, including the agent himself. The prior distribution of ability type is common knowledge. Under this setup, the mean μ (the prior belief about the agent's ability type) reflects the agent's initial self-esteem level while the variance σ_η^2 reflects its fragility. We say that the agent is more secure (fragile) about his own self-image when σ_η^2 is relatively small (large). When $\sigma_\eta^2 = 0$, the agent is certain about his own attributes, and the problem converges to the standard case of a perfectly known self.

We assume the principal allows the agent to engage in a task where the agent's effort level is denoted by $e \in \{0, 1\}$. The cost of exerting effort is given by $c(e)$, where $c(0) = 0$ and $c(1) = c$. The observable outcome $y \in R$ is a noisy signal of ability and effort, specified as

$$y = \eta e + \varepsilon, \tag{18.1}$$

where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ is a disturbance distributed independently of η. Among other things, this specification implies that ability and effort are complementary. The variance of the disturbance σ_ε^2 measures the extent of uncertainty faced by the agent in this economy.[13] Our ultimate concern is the impact of this uncertainty on the cost of inducing effort $e = 1$.

We assume that the outcome is the only contractible variable in this environment and restrict our attention to the class of linear contracts. More precisely, the principal offers

$$w = \beta y + \gamma, \tag{18.2}$$

where γ denotes some (constant) transfer payment. Given this, the contract can generically be written as (β, γ).

---

[12] Although it is in principle possible and conceptually straightforward to extend the model to continuous-effort choices, the derivation of optimal contracts in that situation can be quite computationally complicated. Given that our goal is to compare agency costs for inducing a given level of effort, the current specification of binary-effort choices is just sufficient to make our point in a relatively tractable manner.

[13] We use the terms risk and uncertainty interchangeably throughout the chapter and do not make any particular distinction between them.

Finally, we need to specify the agent's preferences, which mark the sole departure from the conventional setup. While the agent gains utility from material benefits (as typical), he also has vested interests in his own ability type. Let $\tilde{\mu}$ denote the agent's posterior belief about his own ability type, which is contingent on the effort level e and the outcome y. The agent's preferences are specified by the following constant absolute risk aversion (CARA) utility function:

$$u(w - c(e) + \alpha(\tilde{\mu} - \mu)) = -\exp\{-r[w - c(e) + \alpha(\tilde{\mu} - \mu)]\}, \; r > 0. \quad (18.3)$$

Throughout the analysis, we broadly refer to the last term $\alpha(\tilde{\mu} - \mu)$ as the self-esteem concern, and the parameter $\alpha \geq 0$ as a measure of its relative strength. When $\alpha = 0$, the self-esteem concern is totally absent, and the agent's utility depends solely on material payoffs, as conventionally assumed. The self-esteem concern arises directly when we simply like to consider ourselves as able, attractive, caring, and so on: in this case, self-esteem serves merely as another argument in the utility function along with the consumption of private goods.[14] On the other hand, the need for self-esteem arises indirectly if high self-esteem is instrumental in inducing the better performance and expected to lead to higher payoffs in future transactions.[15] However, in the present analysis, we do not make any particular distinction between these two cases.

The timing of the model is summarized as follows:

Stage 1: The principal offers a contract (β, γ). If the agent rejects the contract, he receives the reservation payoff, which is normalized to zero.[16]

Stage 2: The agent determines the effort level e.

Stage 3: The outcome y is realized and the wage is paid as specified by the agreed-upon contract.

---

[14]The most notable example in this respect is arguably sociometer theory (Leary 1999; Leary and Downs 1995), which provides a proper framework for understanding why we develop a taste for high self-esteem. In brief, this theory asserts that as mutual cooperation is indispensable in any human society, human beings inherently possess a psychological mechanism, often referred to as a sociometer, by which they measure the quality of their interpersonal relationships.

[15]People with high self-esteem tend to perform better in and persist longer at tasks (Bandura 1977). Many studies also indicate some positive correlation between self-esteem and academic or task performance, e.g., Davies and Brember (1999) and Judge and Bono (2001), although these results should be interpreted with caution because they do not answer the more challenging question of causality. See Maruyama et al. (1981) for a discussion.

[16]As for the principal's reservation payoff, we simply assume that it is sufficiently low, so that the principal always has an incentive to hire the agent in the first place.

3 The Optimal Contract with Self-Esteem Concerns

3.1 The Effect of Risk on the Agency Cost

The crux of the model is how the agent's behavior influences the extent of uncertainty that he faces. Let $\pi := w - c(e) + \alpha(\tilde{\mu} - \mu)$ denote the total net payoff. Then, the certainty-equivalent (CE) version of the agent's problem is defined as

$$\max_{e \in \{0,1\}} \quad \pi_{\mathrm{CE}} := \beta\mu e + \gamma - c(e) - \frac{r}{2} var(\pi \mid e).$$

Note that $E(\tilde{\mu} \mid e) = \mu$, regardless of the agent's effort choice e. The agent then exerts effort if and only if

$$\beta\mu \geq c + \frac{r}{2}[var(\pi \mid e = 1) - var(\pi \mid e = 0)]. \tag{18.4}$$

Provided that the agent's reservation payoff is normalized to zero, the participation constraint is given by

$$\beta\mu e + \gamma \geq c(e) + \frac{r}{2} var(\pi \mid e), \tag{18.5}$$

when the contract offered is designed to induce effort e. Define $w_e(\sigma_\eta^2, \sigma_\varepsilon^2) := c(e) + (r/2)var(\pi \mid e)$ as the agency cost of inducing effort e. Here, we focus on how w_1 is related to an increase in the extent of uncertainty σ_ε^2.[17]

To solve this problem explicitly, we obviously need to obtain the conditional variance of the payoff $var(\pi \mid e)$. For this, we can obtain the following result.

Lemma 1 *For $e \in \{0, 1\}$,*

$$var(\pi \mid e) = \Gamma_e^2(e^2\sigma_\eta^2 + \sigma_\varepsilon^2),$$

where

$$\Gamma_e^2 := \left(\beta + \frac{\alpha e \sigma_\eta^2}{\sigma_\varepsilon^2 + e^2\sigma_\eta^2}\right)^2.$$

Proof See Appendix.

Several remarks are in order. First, as $\sigma_\eta^2 \to 0$, the problem evidently converges to the standard case where $var(\pi \mid e) = \beta^2\sigma_\varepsilon^2$ regardless of e, and the incentive

[17] For this subsection, we assume that c is so small that it is always optimal for the principal to implement $e = 1$.

compatibility constraint is naturally reduced to $\beta\mu \geq c$. A decrease in α also yields a similar implication: as $\alpha \to 0$, i.e., as the self-esteem concern ceases to exist, the self-esteem fragility σ_η^2 is just another form of uncertainty, playing a similar role to σ_ε^2.[18] When both $\alpha > 0$ and $\sigma_\eta^2 > 0$, however, there arises some interaction between them that works to negate the standard tradeoff between risk and incentives under some conditions.

It follows from Lemma 1 that the incentive compatibility constraint can be written as

$$\beta\mu \geq c + \frac{r}{2}[\Gamma_1^2(\sigma_\eta^2 + \sigma_\varepsilon^2) - \Gamma_0^2\sigma_\varepsilon^2]. \tag{18.6}$$

Define β^* as the optimal incentive needed to induce $e = 1$, which satisfies (18.6) with equality. The next result establishes that there exists some nonempty interval of β that can satisfy (18.6) when the prior belief μ is sufficiently large.

Lemma 2 *Suppose that μ is large enough to satisfy*

$$(\mu - \alpha r\sigma_\eta^2)^2 > r\sigma_\eta^2(r\Phi + 2c),$$

where

$$\Phi := \frac{(\alpha\sigma_\eta)^2}{\sigma_\eta^2 + \sigma_\varepsilon^2}.$$

Then, there exists some nonempty interval $B := [\underline{\beta}, \overline{\beta}]$ such that (18.6) holds for all $\beta \in B$. The lowerbound, $\underline{\beta}$, is the slope of the optimal contract, and we shall explicitly denote its dependence on σ_η^2 and σ_ε^2 by denoting the optimal contract as $(\beta^*(\sigma_\eta^2, \sigma_\varepsilon^2), \gamma^*(\sigma_\eta^2, \sigma_\varepsilon^2))$, which is given by

$$\beta^*(\sigma_\eta^2, \sigma_\varepsilon^2) = \frac{\mu - \alpha r\sigma_\eta^2 - \sqrt{(\mu - \alpha r\sigma_\eta^2)^2 - r\sigma_\eta^2(r\Phi + 2c)}}{r\sigma_\eta^2},$$

$$\gamma^*(\sigma_\eta^2, \sigma_\varepsilon^2) = \frac{r}{2}\beta^{*2}\sigma_\varepsilon^2.$$

Proof See Appendix.

This lemma indicates that the agent needs to be sufficiently able to induce any positive level of effort (or, more precisely, $e = 1$ for any positive c). The reason is clear: given that the marginal value of effort is determined by η, the incentive compatibility constraint can never be satisfied when the expected value of η is close to zero.

[18]The only difference is that this uncertainty is complementary to the effort level so that more effort amplifies the extent of uncertainty. It is important to emphasize, however, that its impact is independent of σ_ε^2.

In what follows, we restrict our attention to the case where μ is large enough to satisfy the condition in the lemma, so that the optimal solution β^* exists.[19] With this well-defined solution, the agency cost of inducing $e = 1$ can be written as

$$w_1(\sigma_\eta^2, \sigma_\varepsilon^2) = c + \frac{r}{2}\Gamma_1^2(\sigma_\eta^2 + \sigma_\varepsilon^2) = c + \frac{r}{2}\left(\beta^*(\sigma_\eta^2, \sigma_\varepsilon^2) + \frac{\alpha\sigma_\eta^2}{\sigma_\varepsilon^2 + \sigma_\eta^2}\right)^2(\sigma_\eta^2 + \sigma_\varepsilon^2). \tag{18.7}$$

It is important to note that the optimal slope β^* depends on σ_ε^2 only when $\alpha\sigma_\eta^2 > 0$.[20] This fact leads to the following observation.

Proposition 1 *The agency cost w_1 increases with the extent of uncertainty σ_ε^2 if $\alpha = 0$ and/or $\sigma_\eta^2 = 0$.*

Proof If $\sigma_\eta^2 = 0$, the incentive compatibility constraint is simply reduced to $\beta\mu \geq c$ so that $\beta^* = c/\mu$. Note also that $\Gamma_1 = \beta^* = c/\mu$ under this condition. The proposition is then immediately obtained.

If $\alpha = 0$, the incentive compatibility constraint becomes $\beta\mu \geq c + \beta^2\sigma_\eta^2$. Once again, the optimal β^* is independent of σ_ε^2. Since $\Gamma_1 = \beta^*$ as above, the agency cost increases with σ_ε^2, provided that a well-defined solution β^* exists.

Q.E.D.

This conclusion, which generally and robustly holds true in the conventional setup, may be overturned because an increase in uncertainty obscures the self-evaluation process, thereby mitigating the need for self-handicapping. More precisely, we make the following claim, which constitutes the main contention of this chapter.

Proposition 2 *Suppose that $\alpha\sigma_\eta^2 > 0$. Then, the agency cost w_1 decreases with the extent of uncertainty σ_ε^2 if μ is sufficiently large.*

Proof See Appendix.

A straightforward way to interpret this result is that the standard tradeoff between risk and incentives tends to disappear when the agent is more self-confident. This interpretation should, however, be made with some caution because this result is to some extent an artifact of our model specification such that there are only two effort choices. Our preferred view is slightly different, although closely related. Fixing

[19] If μ is publicly observable, the principal may have no incentive to hire the agent whose prior expectation μ does not satisfy the condition in the lemma.

[20] When $\alpha\sigma_\eta^2 > 0$, the optimal slope β^* is decreasing in σ_ε^2. Although this appears to coincide with the prediction of the standard framework, this decrease has nothing to do with the tradeoff between risk and incentives. In the standard setup, the optimal slope is decreasing in σ_ε^2 because with an increase in the agency cost, the principal chooses to implement a lower effort level by offering a lower β. As can be seen from the case with $\alpha\sigma_\eta^2 = 0$, the optimal slope is independent of σ_ε^2 in the standard framework.

c constant, an increase in μ alternatively implies that the task at hand is getting easier or less costly, and weaker incentives are then needed to induce effort (the optimal contract converging to a fixed-wage scheme as $\mu \to \infty$). This means that the agent has fewer pecuniary benefits at stake, consequently rendering the self-esteem concern more significant in a relative sense. This implication thus leads to a potentially testable hypothesis: the tradeoff between risk and incentives becomes more ambiguous when less is at stake materially.[21] Conversely, it is also expected that the standard tradeoff becomes more visible for those with more lucrative contracts, such as CEOs and professional athletes, who supposedly face stronger material incentives.[22] Indeed, according to the evidence in Prendergast (2002a,b), the relationship between risk and incentives is at best ambiguous for CEOs whereas it is more clearly positive for sharecroppers.

In addition to this implication, the present model sheds light on two key components of the main result, each of which yields a potentially testable hypothesis. First, we assume at the outset that ability and effort are complementary, which is crucial in giving rise to the need for self-handicapping. In fact, the conclusion is totally reversed when they are substitutive, i.e., when only one of either ability or effort is sufficient to carry out a task. In that case, the agent actually exposes himself less to relevant information by exerting effort as greater effort can cover the lack of ability. Note that this need to conceal the lack of ability diminishes as the underlying environment becomes noisier and the outcome loses its informational content. The model thus predicts an important, possibly testable, connection between the complementarity between ability and effort on the one hand and the tradeoff between risk and incentives on the other, such that the standard tradeoff becomes less visible in situations where effort and ability are more complementary.

Second, the way we model self-esteem concerns implicitly assumes that the agent is risk averse in his own self-image. In other words, we implicitly assume by design that the agent is relatively averse to receiving bad news about himself (which may be referred to as "information aversion"). This condition is critical because the conclusion is once again totally overturned if the agent is risk loving in his own self-image. Suppose that this is the case, such that the agent is actually made better off by collecting more accurate information about himself, given the same mean. In this case, and outside monetary compensation, the agent has an

[21] If psychological factors are exogenously given and the magnitude of their effects is roughly fixed, these factors tend to be more accentuated when fewer material benefits are at stake. For instance, it is well known in social psychology that incentives (explicit rewards) may backfire as they often interfere with "intrinsic motivation". Subsequent studies reveal that this anomaly, which certainly contradicts conventional economic reasoning, is more likely to arise when material benefits are not sufficient (e.g., Gneezy and Rustichini 2000).

[22] With multiple effort levels (including the continuous-effort case), it is generally optimal to offer stronger incentives to agents with higher ability because an expected increase in effort is even larger. In these cases, therefore, the standard tradeoff is more likely to survive when the agent possesses higher ability (or, on the flip side, the task at hand is easier). This indicates that what is important is how much is materially at stake, rather than how able the agent is.

additional incentive to "discover himself" when he faces a task. This extra incentive is naturally stronger when the underlying situation involves less uncertainty as the agent is likely in that case to obtain a more accurate signal. The effect arising from self-esteem concerns then works in the same direction as the standard tradeoff: an increase in uncertainty diminishes this extra incentive, thereby increasing the agency cost. This reasoning implies that information aversion (or truth aversion) constitutes a necessary component for our account to be valid when it is combined with the complementarity between ability and effort.[23] Although this assumption of information aversion appears to be a reasonable specification for many cases, as exemplified by several episodes of self-handicapping behavior, this implication indicates that the question of how we value self-esteem (the sign of the *second* derivative), instead of just whether we care about it (the sign of the *first* derivative), can be part of the critical agenda on this issue.

3.2 On the Negative Relationship Between Risk and Incentives

In the previous subsection, we have seen that the agency cost of inducing effort could decrease with the extent of uncertainty when the agent cares about his self-images, which draws clear contrast to the conventional setup where the agency cost always increases with the extent of uncertainty. Up to this point, however, we have not discussed anything about the relationship between risk and incentives, because we have restricted our attention to an environment where the cost of effort c is so small that it is always optimal for the principal to induce $e = 1$.[24] Here, we relax this assumption to illustrate that the relationship between risk and incentives could indeed be positive when the second-best effort is not fixed.

Now suppose that c could take any value, so that it may be optimal for the principal to implement $e = 0$. In this case, given that the goal of the (risk-neutral) principal is to maximize the expected profit $E(y - w)$, it is optimal to implement $e = 0$ when

$$-w_0(\sigma_\eta^2, \sigma_\varepsilon^2) > \mu - w_1(\sigma_\eta^2, \sigma_\varepsilon^2). \tag{18.8}$$

[23]The standard tradeoff also breaks down when the agent is risk loving in his own self-image, and ability and effort are substitutive. We do not explore this case much as this combination seems less likely.

[24]The assumption allows us to sidestep the principal's problem of choosing the second-best effort and thus to focus on the effect of risk on the agency cost, which is the driving force to break down the negative relationship between risk and incentives. In the conventional setup where $\alpha\sigma_\eta^2 = 0$, β is totally independent of σ_ε^2 when the second-best effort is fixed.

Note that to implement $e = 0$, the contract needs to satisfy only the participation constraint, in which case the optimal slope is zero. Since $w_0(\sigma_\eta^2, 0) = 0$, we can rewrite (18.8) simply as

$$w_1(\sigma_\eta^2, \sigma_\varepsilon^2) > \mu. \tag{18.9}$$

Since β must be strictly larger than zero to induce any positive effort level, we can observe a positive relationship between risk σ_ε^2 and incentives β if there exists some $\bar{\sigma}_\varepsilon^2$ such that it is optimal to induce $e = 0$ for $\bar{\sigma}_\varepsilon^2 > \sigma_\varepsilon^2$ and $e = 1$ for $\sigma_\varepsilon^2 \geq \bar{\sigma}_\varepsilon^2$. In other words, the positive relationship arises if: (i) there exists some $\bar{\sigma}_\varepsilon^2$ such that $w_1(\sigma_\eta^2, \bar{\sigma}_\varepsilon^2) = \mu$; and (ii) $\partial w_1/\partial\sigma_\varepsilon^2 < 0$. Note that we never observe this positive relationship in the standard setup because the latter condition (ii) is never satisfied when $\alpha\sigma_\eta^2 = 0$ as we have seen in proposition 1. When $\alpha\sigma_\eta^2 > 0$, on the other hand, we can make the following claim.

Proposition 3 *Suppose that $\alpha\sigma_\eta^2 > 0$ and μ is sufficiently large. Then, there exist some $\bar{\sigma}_\varepsilon^2$ and c such that for any $\sigma_\varepsilon^{2\prime} > \bar{\sigma}_\varepsilon^2 > \sigma_\varepsilon^{2\prime\prime}$, $\beta^*(\sigma_\eta^2, \sigma_\varepsilon^{2\prime}) > \beta^*(\sigma_\eta^2, \sigma_\varepsilon^{2\prime\prime})$.*

Proof We need to show that there exists some $\bar{\sigma}_\varepsilon^2$ such that $w_1(\sigma_\eta^2, \bar{\sigma}_\varepsilon^2) = \mu$ in the range where $\partial w_1/\partial\sigma_\varepsilon^2 < 0$. In other words, we must find $\bar{\sigma}_\varepsilon^2$ such that

$$w_1(\sigma_\eta^2, \bar{\sigma}_\varepsilon^2) = \mu, \tag{18.10}$$

when μ is sufficiently large. Fixing $\mu - c = d$ for some $d \in R$ and letting $\mu \to \infty$ and $c \to \infty$, (18.10) becomes

$$r\sigma_\eta^2\left(\frac{\alpha\sigma_\eta^2}{\sigma_\varepsilon^2 + \sigma_\eta^2}\right)^2(\sigma_\eta^2 + \sigma_\varepsilon^2) = d. \tag{18.11}$$

We can always find d to satisfy this.

Q.E.D.

The result indicates that the optimal slope increases from zero to some positive level when the extent of uncertainty gets past the threshold $\bar{\sigma}_\varepsilon^2$, while it gradually decreases for $\sigma_\varepsilon^2 > \bar{\sigma}_\varepsilon^2$. The latter result is, of course, due to the fact that there are only two effort levels. With more effort levels (including continuous effort), non-monotonicity could occur at multiple points, so that we could observe positive correlation at multiple data points.

4 Implications for Organizational Design

An important implication of the model is that the principal may benefit from uncertain environments and/or ambiguous outcomes. This implies that the principal may have an incentive to deliberately introduce noise into the evaluation process if it is technologically feasible to do so. This yields several implications of practical importance for organizational design that are particularly important for inherently ego-threatening tasks, e.g., nonroutine and intellectually challenging tasks.

4.1 Why and How Can Team Production Be Profitable?

One recent suggestion is that many firms adopt team production as a means of motivating workers and improving productivity. To account for this success, psychologists along with managers and executives often emphasize the psychological side of team production. Among the most notable advocates are social psychologists such as Baumeister and Leary (1995), who emphasize the sense of social inclusion or acceptance as a crucial facet of life and hence the key to the effectiveness of team production. This line of reasoning also appears well received within the business community. For instance, a former director of management and organization development at General Foods Worldwide is quoted as saying, "Teams can withstand much more stress than individuals because teams reproduce a family structure... That sense of belonging—coupled with the additional energy team members provide for each other—results in more excitement and enthusiasm" (King 1989).

Despite these claims, however, the enthusiasm for teams remains a puzzle for most economists as standard theory asserts that team production invites a notorious free-rider problem because it becomes harder to distinguish between the contributions of the agents in a team.[25] After all, it is not entirely clear through which channel team production provides motivation for workers, especially to the extent where it can overcome the free-rider problem. Clearly, this is where economics diverges from other disciplines, especially social psychology.

We intend to offer an alternative view on the virtue of team production as a way to artificially introduce noise into the self-evaluation process. As failure can be attributed, at least partially, to teammates, team production inevitably introduces ambiguity into the self-evaluation process. As a practical interpretation, one can thus argue that "teams can withstand much more stress than individuals" because workers sense that they can share the blame when things do not turn out as favorably as expected. Given that the need for self-handicapping decreases under team production, the cost of inducing effort can also be lower, consequently making team production more profitable.

[25]Naturally, this has also attracted the attention of economists and several explanations have been offered to account for the rise in team production. To name a few, Itoh (1993) emphasizes the role of effort coordination that stems from mutual monitoring, while Che and Yoo (2001) show that team incentives may be optimal when the agents interact repeatedly over time. Ishida (2006a) builds on this framework and shows how team incentives work under relative performance evaluation.

For the sake of comparison, we consider exactly the same specification as in the previous section. The only difference is that the principal now hires two agents, denoted by $i = A, B$, and allows them to jointly perform some task. The only available measure of performance is the aggregate outcome y, which is given by

$$y = \frac{(\eta_A e_A + \varepsilon_A) + (\eta_B e_B + \varepsilon_B)}{2}, \tag{18.12}$$

where ε_i is i.i.d. The additive nature of the production function implies that there are no technological complementarities between the agents, a potential source of the benefit of team production.

Under this setup, we show that the principal may induce effort $e_i = 1$ from both agents at a lower cost when they work jointly. To see this, it is once again crucial to obtain the variance of the total payoff π_i. We obtain the following result by slightly modifying Lemma 11.

Lemma 3

$$\begin{aligned} var(\pi_i \mid e_i, e_j) &= \left(\frac{\beta e_i}{2} + \frac{\alpha e_i^2 \sigma_\eta^2}{2\sigma_\varepsilon^2 + e_i{}^2\sigma_\eta^2}\right)^2 \sigma_\eta^2 + 2\left(\frac{\beta}{2} + \frac{\alpha e_i \sigma_\eta^2}{2\sigma_\varepsilon^2 + e_i{}^2\sigma_\eta^2}\right)^2 \sigma_\varepsilon^2 \\ &\quad + \left(\frac{\beta e_j}{2}\right)^2 \sigma_\eta^2,\ i \neq j. \end{aligned}$$

Proof See Appendix.

Given this, the incentive compatibility constraint (for agent A) is

$$\frac{\beta\mu}{2} \geq c + \frac{r}{2}[var(\pi_A \mid e_A = 1, e_B) - var(\pi_A \mid e_A = 0, e_B)], \tag{18.13}$$

while the participation constraint is

$$\beta\mu\frac{e_A + e_B}{2} + \gamma \geq c(e_A) + \frac{r}{2}var(\pi_A \mid e_A, e_B). \tag{18.14}$$

Define β^{**} as the minimum incentive needed to induce effort $e_A = e_B = 1$ from both agents that must satisfy

$$\beta^{**} = \frac{2c}{\mu} + \frac{r}{\mu}\left[\left(\frac{\beta}{2} + \frac{\alpha\sigma_\eta^2}{2\sigma_\varepsilon^2 + \sigma_\eta^2}\right)^2 (2\sigma_\varepsilon^2 + \sigma_\eta^2) - 2\left(\frac{\beta}{2}\right)^2 \sigma_\varepsilon^2\right]. \tag{18.15}$$

The agency cost of inducing $e_A = e_B = 1$ under team production is then defined as

$$W_1(\sigma_\eta^2, \sigma_\varepsilon^2) = c + \frac{r}{2}\left[\left(\frac{\beta^{**}(\sigma_\eta^2, \sigma_\varepsilon^2)}{2} + \frac{\alpha\sigma_\eta^2}{2\sigma_\varepsilon^2 + \sigma_\eta^2}\right)^2 (2\sigma_\varepsilon^2 + \sigma_\eta^2) + \left(\frac{\beta^{**}(\sigma_\eta^2, \sigma_\varepsilon^2)}{2}\right)^2 \sigma_\eta^2\right]. \tag{18.16}$$

Comparing W_1 with w_1 obtained in the previous section, we can make the following claim, which is very similar in nature to Proposition 2.

Proposition 4 *The agency cost is lower under team production than under individual production, i.e.,* $W_1(\sigma_\eta^2, \sigma_\varepsilon^2) < w_1(\sigma_\eta^2, \sigma_\varepsilon^2)$*, for any given* $(c, \alpha, \sigma_\eta^2, \sigma_\varepsilon^2)$ *if* μ *is sufficiently large.*

Proof The agency cost decreases under team production if $var(\pi_i \mid e_A = 1, e_B = 1) < var(\pi \mid e = 1)$, i.e.,

$$\left(\frac{\beta^{**}}{2} + \frac{\alpha\sigma_\eta^2}{2\sigma_\varepsilon^2 + \sigma_\eta^2}\right)^2 (2\sigma_\varepsilon^2 + \sigma_\eta^2) + \left(\frac{\beta^{**}}{2}\right)^2 \sigma_\eta^2 < \left(\beta^* + \frac{\alpha\sigma_\eta^2}{\sigma_\varepsilon^2 + \sigma_\eta^2}\right)^2 (\sigma_\varepsilon^2 + \sigma_\eta^2). \tag{18.17}$$

Note that $\lim_{\mu\to\infty} \beta^* = 0$ and $\lim_{\mu\to\infty} \beta^{**} = 0$. As $\mu \to \infty$, therefore, (18.17) can be written as

$$\left(\frac{\alpha\sigma_\eta^2}{2\sigma_\varepsilon^2 + \sigma_\eta^2}\right)^2 (2\sigma_\varepsilon^2 + \sigma_\eta^2) < \left(\frac{\alpha\sigma_\eta^2}{\sigma_\varepsilon^2 + \sigma_\eta^2}\right)^2 (\sigma_\varepsilon^2 + \sigma_\eta^2), \tag{18.18}$$

which can be shown to hold for any given $(c, \alpha, \sigma_\eta^2, \sigma_\varepsilon^2)$.

Q.E.D.

4.2 Hidden Costs of External Enforcers

Adopting team production is a means of artificially introducing noise into the self-evaluation process so as to reduce the need for self-handicapping. There are, however, other possible ways, yielding further implications for organizational design. The crux of the model is that the agency cost tends to increase when the required task is more ego-threatening. This conversely implies that for the construction of efficient work organizations, it is important to create less ego-threatening environments where workers do not feel overly tested. Here, we relate this argument to the potential costs of external enforcers such as evaluation and monitoring. The types of managerial strategy to create less ego-threatening environments become even more important when objective performance measures are not readily available, so that it is easier for the principal to control or manipulate how much information is disseminated to the agent.

An immediate implication drawn from this logic is that the way workers are monitored and evaluated may have a critical bearing on their motivation. Standard agency theory in economics implies that monitoring and evaluation, which typically lead to more accurate performance measures, reduce agency cost (through the reduction in uncertainty) and are hence beneficial for the principal, given that their costs are sufficiently small.[26] This conclusion, however, is in clear contrast to the conventional wisdom in social psychology, where the hidden costs of external enforcers are discussed rather extensively.[27] In this regard, our analysis suggests a plausible route to bridge this gap: when self-esteem concerns are explicitly taken into account, workers may be better motivated if the supervisor relies less on monitoring and evaluation, thereby making, to some extent, the outcomes more ambiguous. In fact, experimental evidence indicates that creativity is enhanced when participants do not expect their products to be evaluated (Amabile et al. 1990). This line of evidence is particularly suggestive because creative activities are typically nonroutine and objective performance measures for these sorts of activities are not readily available.

As a final point, we should note that human motivation is a highly complicated subject, and we certainly do not claim that the present framework is capable of explaining it all. It is our view, and certainly that of many others, that there must be many factors simultaneously at work on this issue, and self-esteem concerns provide only a partial response to this problem. Nonetheless, we do argue that the present framework provides a step towards shedding light on one aspect of this issue. In this regard, on the one hand, information is undoubtedly indispensable for mitigating the problem of moral hazard or, more generally, making optimal decisions. On the other hand, our analysis also points out that more precise information may actually deplete workers' motivation. The optimal balance of these two sides depends on many external factors, which could lead to serious implications for organizational design. Work along this line of inquiry, investigating both theoretically and empirically when it is optimal to create a work environment so as to generate ambiguous outcomes, could offer important insights in this regard.

5 Concluding Remarks

This chapter examines the role of self-esteem concerns in contracting situations using a simple model of moral hazard. When the agent is relatively averse to having a negative self-image, an incentive arises to remain ignorant about himself, ultimately amounting to the need for self-handicapping. In this situation, uncertainty

[26]Obviously, the question of when and how subjective evaluation works effectively is an important issue worthy of rigorous investigation; see, e.g., Levin (2003) and MacLeod (2003). However, as this is clearly outside the scope of the present analysis, we do not deal with the strategic aspects of subjective evaluation.

[27]For instance, Cognitive Evaluation Theory (CET) suggests that external factors (including rewards, competition, surveillance and evaluation) tend to diminish intrinsic motivation. CET explains that these external factors undermine feelings of autonomy, thereby prompting a change in the perceived locus of causality from internal to external and inhibiting intrinsic motivation. Examples of the economic discussion of the hidden costs of rewards include Kreps (1997), Frey and Oberholzer-Gee (1997), Benabou and Tirole (2003, 2006) and Sliwka (2007).

may have a positive incentive effect as it introduces ambiguity into the self-evaluation process. This actually mitigates the need for self-handicapping and consequently reduces agency cost.

The main purpose of the chapter is simply to illustrate the point that the presence of self-esteem concerns yields an implication that runs counter to conventional wisdom. For this reason, several avenues are available to extend the present analysis. First, one important extension would be to examine the role of self-esteem concerns in more general settings, e.g., continuous effort choices and/or, more generally, nonlinear contracts. Of course, as is well known, it is difficult enough to characterize the optimal contract in moral hazard problems even without self-esteem concerns. It is nonetheless an important agenda to see how far one can push this logic and generalize the results obtained in the present chapter.

Second, this analysis assumed that the agent in question is a rational Bayesian learner who can objectively compute his posterior belief. In reality, it is well known that people often resort to different means to manipulate their own beliefs, ranging from self-serving bias and cognitive dissonance to downward comparison. For instance, self-serving bias is another inherent human tendency that is thought to be part of the self-defense mechanism. Although we emphasize the aspect that the main results hold *even* for rational Bayesian learners, it would be of some interest to see the interaction, if any, between self-esteem concerns and biased information processing.

Acknowledgements I am grateful to seminar participants at Contract Theory Workshop for helpful comments. Financial support by the Japanese Ministry of Education, Culture, Sports, Science and Technology (Grants-in-Aid for Young Scientists (B)) is gratefully acknowledged.

Appendix

Proof of Lemma 1 First, it is clear that the agent learns nothing about himself when $e = 0$, so that the posterior always equals the prior, i.e., $\tilde{\mu} = \mu$. When $e = 1$, the outcome y contains some information about the agent's ability type. For any given e, the theory of conjugate distributions implies that the posterior belief is obtained as

$$\tilde{\mu} = \frac{\frac{1}{\sigma_\eta^2}\mu + \frac{e^2}{\sigma_\varepsilon^2}\hat{\eta}}{\frac{1}{\sigma_\eta^2} + \frac{e^2}{\sigma_\varepsilon^2}} = \frac{\sigma_\varepsilon^2\mu + e^2\sigma_\eta^2\hat{\eta}}{\sigma_\varepsilon^2 + e^2\sigma_\eta^2} = \frac{\sigma_\varepsilon^2\mu + e^2\sigma_\eta^2\eta + e\sigma_\eta^2\varepsilon}{\sigma_\varepsilon^2 + e^2\sigma_\eta^2}, \tag{18.19}$$

where $\hat{\eta} = y/e = \eta + \varepsilon/e$ is the best estimate of η from the observation of y alone. See DeGroot (ch. 9, 1970) for details. It then follows that the agent's *ex post* payoff is

$$\pi = \beta(\eta e + \varepsilon) + \gamma - c(e) + \alpha\frac{e^2\sigma_\eta^2(\eta - \mu) + e\sigma_\eta^2\varepsilon}{\sigma_\varepsilon^2 + e^2\sigma_\eta^2}. \tag{18.20}$$

Given that α, β, γ and μ are all constant at this stage, we have

$$\begin{aligned} var(\pi \mid e) &= \Big(\beta e + \frac{\alpha e^2 \sigma_\eta^2}{\sigma_\varepsilon^2 + e^2\sigma_\eta^2}\Big)^2 \sigma_\eta^2 + \Big(\beta + \frac{\alpha e \sigma_\eta^2}{\sigma_\varepsilon^2 + e^2\sigma_\eta^2}\Big)^2 \sigma_\varepsilon^2 \\ &= \Big(\beta + \frac{\alpha e \sigma_\eta^2}{\sigma_\varepsilon^2 + e^2\sigma_\eta^2}\Big)^2 (e^2\sigma_\eta^2 + \sigma_\varepsilon^2). \end{aligned} \tag{18.21}$$

Q.E.D.

Proof of Lemma 2 It follows from (18.6) and Lemma 1 that the minimum incentive needed to induce $e = 1$, denoted as β^*, must solve

$$\begin{aligned} \beta^* &= \frac{c}{\mu} + \frac{r}{2\mu}\Big[\Big(\beta^* + \frac{\alpha\sigma_\eta^2}{\sigma_\varepsilon^2 + \sigma_\eta^2}\Big)^2 (\sigma_\eta^2 + \sigma_\varepsilon^2) - \beta^{*2}\sigma_\varepsilon^2\Big] \\ &= \frac{c}{\mu} + \frac{r}{2\mu}\Big[\sigma_\eta^2 \beta^{*2} + \frac{2\alpha\sigma_\eta^2(\sigma_\eta^2 + \sigma_\varepsilon^2)}{\sigma_\varepsilon^2 + \sigma_\eta^2}\beta^* + \frac{(\alpha\sigma_\eta^2)^2}{\sigma_\varepsilon^2 + \sigma_\eta^2}\Big]. \end{aligned} \tag{18.22}$$

Note that the right-hand side is always positive, meaning that the roots, if they exist, must be positive. When μ is sufficiently large, there exist two positive roots, of which the smaller constitutes the solution for our purpose. Applying the quadratic formula we obtain the optimal incentive as follows:

$$\beta^*(\sigma_\eta^2, \sigma_\varepsilon^2) = \frac{\mu - \alpha r \sigma_\eta^2 - \sqrt{(\mu - \alpha r\sigma_\eta^2)^2 - r\sigma_\eta^2(r\Phi + 2c)}}{r\sigma_\eta^2}, \tag{18.23}$$

where $\Phi := (\alpha\sigma_\eta)^2/(\sigma_\eta^2 + \sigma_\varepsilon^2)$. It is evident from this that two positive roots exist, and hence the solution is well defined, when

$$(\mu - \alpha r\sigma_\eta^2)^2 > r\sigma_\eta^2(r\Phi + 2c). \tag{18.24}$$

Finally, as the participation constraint must also hold with equality, we obtain

$$\gamma^*(\sigma_\eta^2, \sigma_\varepsilon^2) = \frac{r}{2} var(\pi \mid e = 0) = \frac{r}{2}\beta^{*2}\sigma_\varepsilon^2. \tag{18.25}$$

Q.E.D.

Proof of Proposition 2 The agency cost of inducing $e = 1$ decreases with the extent of uncertainty if $\partial w_1/\partial\sigma_\varepsilon^2 < 0$, which can be written as

$$\frac{\partial w_1}{\partial \sigma_\varepsilon^2} = r\Gamma_1\Big(\frac{\partial\beta^*}{\partial\sigma_\varepsilon^2} - \frac{\alpha\sigma_\eta^2}{(\sigma_\varepsilon^2 + \sigma_\eta^2)^2}\Big)(\sigma_\varepsilon^2 + \sigma_\eta^2) + \frac{r\Gamma_1^2}{2} < 0. \tag{18.26}$$

It is easy to see from (18.23) that $\partial\beta^*/\partial\sigma_\varepsilon^2 < 0$ as long as there exists a well-defined solution. This means that it suffices to show that

$$-\frac{\alpha\sigma_\eta^2}{\sigma_\varepsilon^2+\sigma_\eta^2}+\frac{\Gamma_1}{2}\le 0 \;\Leftrightarrow\; \beta^*-\frac{\alpha\sigma_\eta^2}{\sigma_\varepsilon^2+\sigma_\eta^2}\le 0. \tag{18.27}$$

This condition holds for any given $(c, \alpha, \sigma_\eta^2, \sigma_\varepsilon^2)$ when μ is sufficiently large as $\lim_{\mu\to\infty}\beta^* = 0$.

Q.E.D.

Proof of Lemma 3 The result obtained here largely parallels Lemma 1. The outcome under team production contains some information about the teammate's ability type, which functions as noise in the self-evaluation process. Again, applying the theory of conjugate distributions, the posterior belief is obtained as

$$\tilde{\mu}_i=\frac{\frac{1}{\sigma_\eta^2}\mu+\frac{e^2}{2\sigma_\varepsilon^2}\hat{\eta}_i}{\frac{1}{\sigma_\eta^2}+\frac{e_i{}^2}{2\sigma_\varepsilon^2}}=\frac{2\sigma_\varepsilon^2\mu+e_i{}^2\sigma_\eta^2\hat{\eta}_i}{2\sigma_\varepsilon^2+e_i{}^2\sigma_\eta^2}=\frac{2\sigma_\varepsilon^2\mu+e_i{}^2\sigma_\eta^2\eta_i+e_i\sigma_\eta^2(\varepsilon_i+\varepsilon_j)}{2\sigma_\varepsilon^2+e_i{}^2\sigma_\eta^2}. \tag{18.28}$$

It then follows that the agent's *ex post* payoff is

$$\pi_i=\beta\frac{(\eta_i e_i+\varepsilon_i)+(\eta_j e_j+\varepsilon_j)}{2}+\gamma-c(e_i)+\alpha\frac{e_i{}^2\sigma_\eta^2(\eta_i-\mu)+e_i\sigma_\eta^2(\varepsilon_i+\varepsilon_j)}{2\sigma_\varepsilon^2+e_i{}^2\sigma_\eta^2}. \tag{18.29}$$

Given that α, γ and μ at this stage are all constant, we have

$$var(\pi_i \mid e_i, e_j)=\left(\frac{\beta e_i}{2}+\frac{\alpha e_i^2\sigma_\eta^2}{2\sigma_\varepsilon^2+e_i{}^2\sigma_\eta^2}\right)^2\sigma_\eta^2+2\left(\frac{\beta}{2}+\frac{\alpha e_i\sigma_\eta^2}{2\sigma_\varepsilon^2+e_i{}^2\sigma_\eta^2}\right)^2\sigma_\varepsilon^2+\left(\frac{\beta e_j}{2}\right)^2\sigma_\eta^2. \tag{18.30}$$

Q.E.D.

Addendum[28]

This chapter examines how the presence of self-esteem concerns shapes the optimal form of contract, especially with regard to the tradeoff between risk and incentives, in an otherwise standard moral-hazard environment. To be more precise, the current analysis considers a setting where the agent is uncertain about his productivity and has self-esteem concerns in that he derives utility, either directly or indirectly, from having a positive self-image. In this setup, since the observed output level

[28] This addendum has been newly written for this book chapter.

forms a signal of the agent's unknown ability type, the agent must strategically choose his effort choice so as to control the amount of information to which he will subsequently be exposed.

As it turns out, the impact of the agent's self-esteem concerns on his effort choice is determined by his attitude towards the uncertainty about his type. For expositional purposes, we say that the agent is information-averse (information-loving) when his payoff is concave (convex) in the self-esteem level. When the agent is information-averse, he would be better off with less precise information about himself, given the same mean. This force then gives rise to a form of intrapersonal strategy, known as self-handicapping in social psychology, where the agent intentionally reduces the effort level to remain vague about himself. Building on this logic, we show that the standard tradeoff between risk and incentives may break down when the agent cares about self-esteem and moreover is information-averse, because more risk associated with the task mitigates the need for self-handicapping. This argument provides an explanation for why we do not empirically observe this tradeoff in a robust manner, despite the strong theoretical prediction.

As is clear from this argument, the analysis makes two crucial assumptions on the agent's information and payoff structures. First, the agent in question only has vague knowledge about himself to begin with and, more importantly, benefits directly from having a positive self-image. Second, the agent's payoff is concave in the self-esteem level, so that he is information-averse. The first assumption is effectively the assumption on the first derivative of the payoff function (whether the agent cares about his ability) whereas the second one is the assumption on the second derivative. Although the second assumption is more subtle than the first, it is nonetheless equally important to derive the main result: if the second assumption fails to hold, i.e., the agent is information-loving, the conclusion is totally reversed as he has an extra incentive to exert effort to "discover himself."

There is strong evidence in support of the first assumption: a large body of the literature in social psychology and other related disciplines suggests that our self-knowledge is far from perfect, and we indeed do care about our self-images to a considerable extent. In contrast, the validity of the second assumption is much less clear. There are few works, if any, which directly investigate the shape of the payoff function with respect to the self-esteem level. Although instances of self-handicapping behavior seem to imply that people are by and large information-averse, it is still nothing more than a conjecture which needs to be verified in a more rigorous manner. As a potential avenue for future research, it is of some interest to investigate, either empirically or experimentally, whether and under what conditions (e.g., for what type of personal trait) people exhibit information-aversion in order to better understand how the presence of self-esteem concerns affects our behavior in general.

References

Ackerberg DA, Botticini M (2002) Endogenous matching and the empirical determinants of contract form. J Political Econ 110:564–591

Amabile TM, Goldfarb P, Brackfield SC (1990) Social influences on creativity: evaluation, coaction, and surveillance. Creat Res J 3:6–21

Bandura A (1977) Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 84:191–215

Baumeister RF, Leary MR (1995) The need to belong: desire for interpersonal attachments as a fundamental human motivation. Psychol Bull 117:497–529

Benabou R, Tirole J (2002) Self-confidence and personal motivation. Q J Econ 117:871–915

Benabou R, Tirole J (2003) Intrinsic and extrinsic motivation. Rev Econ Stud 62:315–339

Benabou R, Tirole J (2004) Willpower and personal rules. J Political Econ 112:847–887

Benabou R, Tirole J (2006) Incentives and prosocial behavior. Am Econ Rev 96:1652–1678

Carrillo JD, Mariotti T (2000) Strategic ignorance as a self-disciplining device. Rev Econ Stud 67:529–544

Che Y-K, Yoo S-W (2001) Optimal incentives for teams. Am Econ Rev 91:525–541

Core JE, Guay WR (2002) The other side of the trade-off: the impact of risk on executive compensation: a revised comment. Mimeo, University of Pennsylvania, Philadelphia, PA

Cowen T, Glazer A (2007) Esteem and ignorance. J Econ Behav Organ 63:373–383

Davies J, Brember I (1999) Reading and mathematics attainments and self-esteem in years 2–6—an eight-year cross-sectional study. Educ Stud 25:145–157

DeGroot MH (1970) Optimal statistical decisions. McGraw-Hill, New York

Dewatripont M, Jewitt I, Tirole J (1999a) The economics of career concerns, part I: comparing information structures. Rev Econ Stud 66:183–98

Dewatripont M, Jewitt I, Tirole J (1999b) The economics of career concerns, part II: applications to missions and accountability of government agencies. Rev Econ Stud 66:199–217

Frey BS, Oberholzer-Gee F (1997) The cost of price incentives: an empirical analysis of motivation crowding out. Am Econ Rev 87:746–755

Gneezy U, Rustichini A (2000) Pay enough or don't pay at all. Q J Econ 115:791–810

Holmstrom B (1999) Managerial incentive problems: a dynamic perspective. Rev Econ Stud 66:169–182

Ishida J (2006a) Team incentives under relative performance evaluation. J Econ Manag Strateg 15:187–206

Ishida J (2006b) Optimal promotion policies with the looking-glass effect. J Labor Econ 24:857–77

Itoh H (1993) Coalitions, incentives, and risk sharing. J Econ Theory 60:410–427

Jin L (2002) CEO compensation, diversification and incentives. J Financ Econ 66:29–63

Judge TA, Bono JE (2001) Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: a meta-analysis. J Appl Psychol 86:80–92

King P (1989) What makes teamwork work? Psychology Today 23:16

Koszegi B (2006) Ego utility, overconfidence and task choice. J Eur Econ Assoc 4:673–707

Kreps DM (1997) Intrinsic motivation and extrinsic incentives. Am Econ Rev 87:359–364

Leary MR (1999) The social and psychological importance of self-esteem. In: Kowalski RM, Leary MR (eds) The social psychology of emotional and behavioral problems: interfaces of social and clinical psychology. American Psychological Association, Washington D.C.

Leary MR, Deborah LD (1995) Interpersonal functions of the self-esteem motive: the self-esteem system as a sociometer. In: Kernis MH (ed) Efficacy, agency, and self-esteem. Plenum, New York

Levin J (2003) Relational incentive contracts. Am Econ Rev 93:835–857

MacLeod WB (2003) Optimal contracting with subjective evaluation. Am Econ Rev 93:216–240

Maruyama G, Rubin RA, Kingsbury GG (1981) Self-esteem and educational achievement: independent constructs with a common cause? J Personal Soc Psychol 40:962–975

Prendergast C (2000) What trade-off of risk and incentives? Am Econ Rev 90:421–425
Prendergast C (2002a) The tenuous tradeoff between risk and incentives. J Political Econ 110:1071–1102
Prendergast C (2002b) Uncertainty and incentives. J Labor Econ 20:S115–S134
Raith M (2003) Competition, risk and managerial incentives. Am Econ Rev 93:1425–1436
Sliwka D (2007) Trust as a signal of a social norm and the hidden costs of incentive schemes. Am Econ Rev 97:999–1012
Swank OH, Visser B (2007) Motivating through delegating tasks or giving attention. J Law Econ Organ 23:731–742
Tucker JA, Vuchinich RE, Sobel MB (1981) Alcohol consumption as a self-handicapping strategy. J Abnorm Psychol 90:220–230

Chapter 19
Optimal Promotion Policies with the Looking-Glass Effect

Junichiro Ishida

Abstract This chapter considers a model where the agent is uncertain about his innate ability and instead makes an inference from others' (namely, the principal's) perception, as often emphasized in the psychology literature. When the principal has superior knowledge about the agent's productivity than the agent himself, the principal has an incentive to use promotions strategically to boost the agent's self-confidence. Within this framework the optimal promotion policy depends not only on the agent's current expected ability type but also on the history of his previous job assignments. We then use this fact to explain why we rarely observe demotions in organizations.

Keywords Promotion policies • Personal motivation • Self-confidence • Labor contracts • Peter principle.

JEL Classification Codes D21,M12.

1 Introduction

It is often assumed in economics that an agent typically has perfect knowledge of himself. It appears, however, that this view is not entirely shared by other disciplines. In social psychology, how we come to understand ourselves has always been one of the most critical issues: the self is often perceived as a result of the social process whereby the way we see ourselves is inevitably shaped by various factors

---

The original article first appeared in Journal of Labor Economics 24: 857–877, 2006. A newly written addendum has been added to this book chapter.

J. Ishida (✉)
Institute of Social and Economic Research, Osaka University,
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan
e-mail: jishida@osipp.osaka-u.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_19

surrounding us.[1] Along this line, Cooley (1902) argued that people obtain a sense of who they are by how others perceive them – the concept now known as the looking-glass self. This idea was further elaborated by Mead (1934) who emphasized that the sense of self is created by social interactions. According to those views, we shape our self-images as others see us.

In this chapter we incorporate this sociopsychological view on the self to a model of internal labor markets and examine its effect on managerial strategies. More specifically, we consider a situation where, although the principal and the agent share a common prior belief about the agent's innate ability, the principal gains more accurate information than the agent himself as the relationship progresses. The fact that the principal possesses more accurate information potentially has profound impacts on her behavior since it partially reveals what she knows.[2] This is particularly important when ability and effort are complementary to each other, as is often assumed in many economic analyses: in a situation like this, the principal has every incentive to boost the agent's self-confidence because more self-confidence in his own ability induces more effort from him.

As for the information structure, the stance we take in this chapter is as follows. While the agent's perception about himself has apparently been shaped by his past experiences prior to employment, we regard that most dominant components of those experiences, such as academic achievements, are reasonably publicly observable so that the principal and the agent roughly share a common prior belief concerning the agent's attributes. After the employment relationship begins, however, a divergence in the beliefs held by each party begins to surface. Possibly with more experiences observing other agents in the past or with more knowledge about surrounding environments, the principal generally has more resources to assess and evaluate the agent's ability from his achievements than the agent himself. Recognizing this fact, the principal's assessment of the agent's ability has profound effects on his own assessment. In this chapter, we broadly refer to any consequences that arise from this learning process as the looking-glass effect. Since promotion decisions are one of the most visible and also credible signal of the principal's assessment, the nature of job assignment becomes much more strategic in the presence of the looking-glass effect.

The present chapter is close in spirit to Benabou and Tirole (2003), who focus on the psychological effect of inter-personal strategies when the principal possesses some private information. They provide a general framework to analyze the interplay between extrinsic and intrinsic motivation in principal-agent situations

[1]The history of though on the self dates back at least to Aristotle. The contemporary study of the self was initiated by James (1890). James made a distinction between the self as knower (*I*, the process) and the self as known (*me*, the content). James argued that the known self (the *me*) consisted of three domains of self-experiences. One of the domains is the *social self*, which involves recognition one obtains from friends and significant others.

[2]The possibility that the principal's strategy reveals information, when she possesses superior information, is raised by Myerson (1983). The situation where the principal possesses private information is thoroughly analyzed by Maskin and Tirole (1990, 1992).

and show, among other results, that the principal delegates more frequently than under symmetric information. In this chapter we modify this idea to a model of job assignment and extend it to a multiperiod situation. The extension allows us to endogenize the degree of information asymmetry and to derive some dynamic implications consistent with several stylized facts of internal labor markets. To this end, we construct a two-period model of job assignment. There are two distinct jobs, called job M (management) and job L (labor). Suppose that job M is more ability-intensive than job L in that the productivity at job M depends more heavily on the agent's innate ability. Given that ability matters more at job M, the principal assigns the agent to the job only when she has sufficient confidence in his ability. Conversely speaking, this also implies that assigning the agent to job M is a credible signal of the principal's confidence in his ability. The job assignment in this context then has both direct and indirect effects. On one hand, the job assignment directly affects the agent's productivity due to the difference in the nature of job; one the other hand, it can also have an indirect effect, referred to as the looking-glass effect, through affecting the agent's self-confidence level and thus his effort level.

We then use this framework to account for two frequently cited observations in labor management as an important implication of the looking-glass effect. The first is the Peter Principle, which states that "in a hierarchy, every employee tends to rise to his level of incompetence (Peter and Hull 1969)." In their celebrated book, it is further claimed that "every post tends to be occupied by an employee who is incompetent to carry out its duties," because employees are promoted through positions where they have excelled until they reach a level of incompetence. While this phenomenon has been widely discussed in many fields,[3] a close look at this naturally raises another question: if an employee turns out to be incompetent at a new position, why not demote him to the previous position where he excelled? The Peter Principle is therefore closely related to the second observation of our interest, i.e., demotions (downward movements in the firm's hierarchy) are rarely observed in organizations.[4]

We argue that the present framework offers a potential explanation for these observations. The logic is as follows. In the presence of the looking-glass effect, a demotion undermines the agent's self-esteem and thus lowers the effort level while a promotion does the opposite. There then arises an incentive for the principal to manipulate the job assignment rule to exploit the value of her private information.

[3]Recent economic examples that deal with the Peter Principle include Bernhardt (1995), Fairburn and Malcomson (2001) and Lazear (2004).

[4]Unfortunately, due to limited data availability, there are not many empirical studies on promotion dynamics inside firms. Baker, Gibbs and Holmstrom (1994) are one of few studies which documents that demotions are rare in firms. On the contrary, a recent study by Hamilton and MacKinnon (2002) shows that demotions are used on a wider scale at the Canadian Pacific Railway. Gibbons and Waldman (1999) argue, however, that despite the lack of empirical evidence, it is uncontroversial that demotions are indeed rare in firms. On the theoretical side, Bernhardt (1995) presents a model of asymmetric learning based on Waldman's (1984) model of asymmetric learning to explain why demotions are rare in organizations.

As we will see shortly, the presence of the looking-glass effect in general makes the promotion rule more lenient for the agent in that the principal is forced to lower the minimum ability level necessary to be assigned to job M. Note, however, that the magnitude of the looking-glass effect depends on the difference in the accuracy of information held by each party. When the principal and the agent share a common belief, for instance, the job assignment reveals no relevant information and the looking-glass effect is totally absent as a consequence. This implies that the agent's productivity now depends possibly on the history of his previous job assignments. Since ability matters more at job M, more information about the agent's ability is revealed when he is assigned to it. Given this perception, a demotion (a movement from job M to job L) is much more demoralizing for the agent than no promotion (assigned to job L in both periods). With this effect, it may be in the principal's best interest to keep the agent, who turned out to be incompetent, at job M rather than to demote him to job L.

Besides this, the model also raises two additional implications. First, it provides a new perspective on the role of seniority. As stated, when the principal possesses more accurate information regarding the agent, she is forced to lower the promotion threshold. Provided that the principal and the agent share a common prior belief to begin with and a divergence in the beliefs gradually occurs, this implies that the promotion threshold declines as the employment relationship progresses, i.e., seniority matters. This draws clear contrast to the conventional learning approach to explain the role of seniority and delays in the timing of promotion: the conventional approach emphasizes that some attributes of a worker are not initially observable and a promotion to higher ranks occurs when the principal obtains favorable information about the agent. In contrast, in the present framework, an agent may be promoted even when the principal receives no favorable information about the agent afterwards.

Second, it also raises a case where more accurate information is not necessarily beneficial for the principal. In an ordinary circumstance, it is in general beneficial for the principal to collect more accurate information about the agent as it leads to more efficient job assignments. This is not necessarily true in the presence of the looking-glass effect. As the principal's information becomes more accurate, the job assignment reveals more information about the agent. The direction of this information-revelation effect is in general ambiguous and depends on the underlying context of the model. The overall effect can be negative when the optimal effort level is strongly concave in the agent's self-confidence level. In this case, the principal is actually made better off by strategically remaining (or pretending to be) ignorant about the agent. We briefly illustrate this situation and how this information-revelation effect affects the initial job assignment.

The chapter is an attempt to contribute to the understandings of the role of psychological factors in the workplace. Other than Benabou and Tirole (2003) mentioned above, several recent papers explicitly relate psychological factors to workers' incentives in various ways. To name a few, Kandel and Lazear (1992) consider the effect of peer pressure among team members. Rotemberg (1994) analyzes what factors motivate workers to be altruistic toward one another and the consequences of human relations in the workplace. Fang and Moscarini (2005) investigate the implication of worker overconfidence on the optimal wage-setting policies and provide conditions for the non-differentiation wage policy (pooling contract) to be optimal. Gervais and Goldstein (2003) examine the role of overconfidence in team production and show that overconfidence not only enhances team performance but may also be Pareto-improving. Itoh (2004) provides a framework which incorporates other-regarding preferences into the principal-agent setting. There is also a growing body of literature on labor market experiments: see Falk and Fehr (2003) for a brief overview on this.

The chapter is organized as follows. The next section outlines a basic environment which is a fairly standard model of job assignment. Section 3 provides a simple example of the model to illustrate the main result. Section 4 analyzes the general model and show how the presence of the looking-glass effect changes the optimal promotion policy. Finally, Sect. 5 offers some concluding remarks.

2 The Model

2.1 *Environment*

Consider a two-period model where a principal (female) hires an agent (male) to produce output. The agent's ability type, denoted by $\eta \in [0, 1]$, is distributed uniformly over $[0, 1]$. Let $\bar{\eta} \equiv E(\eta) = 0.5$. The ability type is initially unobservable to anyone, including the agent himself. Both parties are assumed to be risk neutral and maximize the sum of payoffs over two periods with no discounting. The relative length of period 1 is denoted by $\theta \in (0, 1)$, which signifies the timing of the arrival of information.

2.2 *Production*

In each period $t = 1, 2$, the agent decides whether to exert costly effort. Let $e_t \in R_+$ denote the effort level in period t. The cost of exerting effort is given by $\theta c(e_1)$ in period 1 and $(1 - \theta)c(e_2)$ in period 2. The cost function is assumed to be twice continuously differentiable with $c' > 0$ and $c'' > 0$. Moreover, to assure the existence of an interior solution, $\lim_{e \to 0} c'(e) = 0$ and $\lim_{e \to \infty} c'(e) = \infty$.

Given the effort level, the agent produces $y_t = a_t e_t$ units of output (per unit of time) in period t. In what follows we refer to a_t as the agent's productivity, which depends on the job assignment as well as the ability type. For the analysis we consider two distinct job types, called job M (management) and job L (labor). We regard that job M is superior to job L in the firm's hierarchy so that a demotion in this context means a movement from job M to job L. The principal must assign the agent to either task at the beginning of each period: let $x_t \in \{L, M\}$ denote the job assignment in period t. Given some job assignment x, the productivity is given by

$$a_t = \begin{cases} \eta \text{ with probability} p_x, \\ \tilde{a} \text{ with probability} 1 - p_x, \end{cases}$$

where $\tilde{a} \in [0, 1]$ is a random variable drawn from some distribution F. Let $\bar{a} \equiv E(\tilde{a}) = \int_0^1 a dF(a)$. We assume that $p_M > p_L$, i.e., the productivity at job M is more sensitive to the agent's ability type. In order to reduce the number of parameters, we further assume that $p_M = 1$ and $p_L = p \in (0, 1)$. Notice that a_2 is totally independent of a_1, i.e., the productivity in period 1 itself has no impact on the productivity in period 2.

It is important to note that this specification is simply a way, among other possibilities, to capture that job M is more ability-intensive than job L. A critical point of this is that the principal assigns the agent to job M when she is sufficiently confident in his ability. Conversely, assigning the agent to job M is a credible signal of the principal's confidence in his ability.

2.3 Information

Since the agent's ability type is not directly observable to anyone, both parties need to infer the true ability type from available information. In this model, the only relevant variable is the expected ability type, which we refer to as the belief. While the principal and the agent share a common prior belief, their posterior beliefs may differ from each other because they subsequently have access to different sets of information. Let μ_1^p and μ_1^a denote the prior belief (about the agent's ability type) held by the principal and the agent, respectively. Under the current setup, $\mu_1^p = \mu_1^a = \bar{\eta}$.

In the present analysis we focus on a situation where the principal gains more accurate information than the agent himself. We in particular assume that the principal can observe the agent's productivity in period 1, a_1, at the end of the period. The principal then updates her belief based upon this observation. Let μ_2^p, or more precisely $\mu_2^p(x_1, a_1)$, denote the principal's posterior belief conditional on her information set (x_1, a_1). While the agent cannot directly observe his productivity, he can apparently observe the job assignment upon which he updates his belief. Let μ_2^a, or more precisely $\mu_2^a(x_1, x_2)$, denote the agent's posterior belief conditional on the history of job assignments (x_1, x_2).

2.4 Compensation

Define $y \equiv \theta y_1 + (1 - \theta) y_2$ as the total output level, which is realized at the end of period 2. The total output level is observable to both parties. We assume that both parties bargain over the total output where the agent's bargaining power is $\lambda \in (0, 1)$, and the threat point for each party is normalized to zero. Given some realized total output level y, therefore, the gross benefit to the agent is λy while that to the principal is $(1-\lambda)y$.

3 A Simple Example

Before we elaborate on a general model, we first illustrate the gist of the model through a simple example. Consider a limiting case where $p \to 0$ so that job L reveals almost no information. Moreover, to obtain a closed-form solution, let $c(e) = e^2/2$.

We first show that, under the current setup, there exists a complementary relationship between ability and effort. Given the job assignment x and some belief μ_t^a, the agent's problem is defined as

$$\max_e \ \lambda A_x(\mu_t^a)e - \frac{e^2}{2},$$

where

$$A_x(\mu) \equiv p_x\mu + (1-p_x)\bar{a}.$$

The optimal solution, denoted by e^*, must satisfy

$$e^* = \lambda A_x(\mu_t^a). \tag{19.1}$$

Given this, we now characterize the optimal promotion rule in each period. In order to illustrate the impact of information asymmetry between the principal and the agent, we first analyze a benchmark case where there is no information asymmetry between them, i.e., the principal does not observe a_1. In this case, there is no reason to change the job assignment after period 1, and the problem is virtually reduced to a single-period model. The principal assigns the agent to job M if and only if

$$\lambda(1-\lambda)A_M(\mu_1^p)A_M(\mu_1^a) \geq \lambda(1-\lambda)A_L(\mu_1^p)A_L(\mu_1^a). \tag{19.2}$$

Since $\mu_1^p = \mu_1^a = \bar{\eta}$, we have

$$A_M(\bar{\eta})^2 \geq A_L(\bar{\eta})^2, \tag{19.3}$$

which is further simplified to

$$\bar{\eta} \geq \bar{a} \equiv \eta_{SI}^*, \tag{19.4}$$

where η_{SI}^* indicates the threshold ability type in the symmetric information case: that is, the principal assigns the agent to job M if and only if the principal's belief on the agent's ability type exceeds the promotion threshold η_{SI}^*.

We now shift our attention to the case where there exists some information asymmetry and the principal potentially possesses more accurate information regarding the agent's ability type than the agent himself. Suppose first that the agent is initially assigned to job L in period 1. In this case, when $p \to 0$, the agent's productivity is completely random and contains no useful information. This implies that $\mu_2^a = \mu_2^p = \bar{\eta}$ for any x_2 and the optimal promotion rule converges to that in the symmetric information case. Suppose, on the other hand, that the agent is assigned to job M in period 1. In this case, the principal can observe the agent's true ability type after the period, and the arrival of this new information has some impact on the promotion rule in period 2. To see this, note that $\mu_2^p = \eta$ and the agent knows that the principal knows his ability type. The job assignment in period 2 thus reveals some information and, in general, $\mu_2^a(M, M) \neq \mu_2^a(M, L)$. Define η_2^*, or $\eta_2^*(x_1)$, as the promotion threshold in period 2 in the asymmetric information case. Since the optimal effort level is given by $e^* = A_x(\mu_2^a)$, the threshold now must satisfy

$$A_M(\eta_2^*)A_M(\mu_2^a(M, M)) = A_L(\eta_2^*)A_L(\mu_2^a(M, L)) = \bar{a}^2. \tag{19.5}$$

Given that $\mu_2^a(M, M) = (\eta_2^* + 1)/2$, this condition becomes

$$\eta_2^*(\eta_2^* + 1) \geq 2\bar{a}^2, \tag{19.6}$$

from which we can conclude that $\eta_2^* < \bar{a} = \eta_{SI}^*$. That is, the fact that the principal possesses more accurate information forces her to lower the promotion threshold. As a practical implication, this shows that the agent who has been assigned to job M is less likely to be demoted to job L in subsequent periods. In the following section, we analyze this intuition in a more general setting and derive some implications consistent with stylized facts of internal labor markets.

4 The Analysis

4.1 The Optimal Effort Choice

As above, we start with the agent's problem. Given the job assignment x and some belief μ_t^a, the agent's problem is defined as

$$\max_e \ \lambda A_x(\mu_t^a)e - c(e).$$

Define e_t^* as the optimal effort level in period t. The first-order condition then implies that

$$\lambda A_x(\mu_t^a) = c'(e_t^*). \tag{19.7}$$

For expositional simplicity we express this solution as follows:

$$e_t^* = c'^{-1}(\lambda A_x(\mu_t^a)) \equiv \phi_x(\mu_t^a). \tag{19.8}$$

Note that $\phi_x(\mu_t^a)$ is strictly increasing in μ_t^a: that is, more confidence in his own ability induces more effort from the agent.

4.2 *The Symmetric Information Case*

We first characterize as a benchmark the optimal promotion rule in the symmetric information case. Even in this general setting, with no information asymmetry, there is no reason to change the job assignment, and the model is therefore reduced to a single-period model. The principal assigns the agent to job M if and only if

$$(1-\lambda)A_M(\mu_1^p)\phi_M(\mu_1^a) \geq (1-\lambda)A_L(\mu_1^p)\phi_L(\mu_1^a), \tag{19.9}$$

which can be written as

$$A_M(\bar{\eta})\phi_M(\bar{\eta}) \geq A_L(\bar{\eta})\phi_L(\bar{\eta}). \tag{19.10}$$

We first establish that the result obtained in the simple example applies for this general case as well.

Proposition 1 *With no information asymmetry, in both periods, the principal assigns the agent to job M if and only if the common prior belief $\bar{\eta}$ exceeds some fixed threshold $\eta_{SI}^* = \bar{a}$.*

Proof Notice first that $\phi_M(\mu) \geq \phi_L(\mu)$ if and only if $A_M(\mu) \geq A_L(\mu)$. It immediately follows from this that (19.10) holds if and only if $\bar{\eta} \geq \bar{a} \equiv \eta_{SI}^*$.

Q.E.D.

4.3 *The Second-Period Problem*

We now extend the analysis to the case where the principal, after a period, obtains additional information about the agent. Given this information there may arise an incentive for the principal to change the job assignment that she initially assigns to the agent. In doing so, however, the principal must take into account the fact that the job assignment in period 2 partially reveals the information held by the principal. The nature of job assignment becomes much more strategic in the presence of information asymmetry.

We first examine the second-period problem, taking the job assignment in period 1 as given. The optimal effort level can now be written as

$$e_2^* = \phi_{x_2}(\mu_2^a). \tag{19.11}$$

The principal makes an inference about the agent's ability type upon observing a_1. Based on this updated posterior belief, the principal must decide the job assignment in period 2, knowing that the agent in turn makes an inference about his own ability type from it. Throughout the analysis we restrict our attention to a situation where the optimal job assignment rule in period 2 takes the following cutoff form: the principal assigns the agent to job M if and only if $a_1 \geq a^*(x_1)$ for each $x_1 = M, L$.[5] In period 2, therefore, the principal assigns the agent to job M if and only if

$$A_M(\mu_2^p)\phi_M(\mu_2^a(x_1, M)) \geq A_L(\mu_2^p)\phi_L(\mu_2^a(x_1, L)), \tag{19.12}$$

for some given x_1. In particular, if there exists an interior solution, the threshold ability type $\eta_2^*(x_1) = \mu_2^p(x_1, a^*(x_1))$ must satisfy

$$A_M(\eta_2^*(x_1))\phi_M(\mu_2^a(x_1, M)) = A_L(\eta_2^*(x_1))\phi_L(\mu_2^a(x_1, L)). \tag{19.13}$$

In what follows we assume that there exists a unique interior solution that satisfies (19.13).[6]

Notice the difference from (19.10) where the principal and the agent share a common belief. In contrast, in the presence of information asymmetry, there is a potential divergence in the posterior belief held by each party, which forces the principal to lower the promotion threshold in period 2. We now present the following result, which is closely related to Proposition 5 of Benabou and Tirole (2003)

Proposition 2 *With information asymmetry, the principal assigns the agent to job M in period 2 if and only if her posterior belief exceeds some fixed threshold $\eta_2^*(x_1)$, which is strictly less than η_{SI}^* for $x_1 = L, M$.*

Proof See Appendix.

When the principal possesses superior information than the agent himself, the job assignment becomes more strategic in nature as it partially reveals to the agent what the principal observes. In this environment, assigning the agent to job L in period 2 is more demoralizing for the agent than doing so in period 1 since it shows the lack of confidence on the principal's part. The proposition indicates that the presence of the looking-glass effect generally makes the promotion rule more lenient for the agent.

---

[5] We later show an example where this type of job assignment rule may not maximize profit. We justify this restriction because the job assignment rule which involves randomization is too complicated for the agent to grasp completely. When the agent expects the job assignment rule to take a cutoff form, the optimal job assignment rule indeed takes a cutoff form.

[6] See Appendix B for a more precise condition for the uniqueness of the optimal promotion threshold.

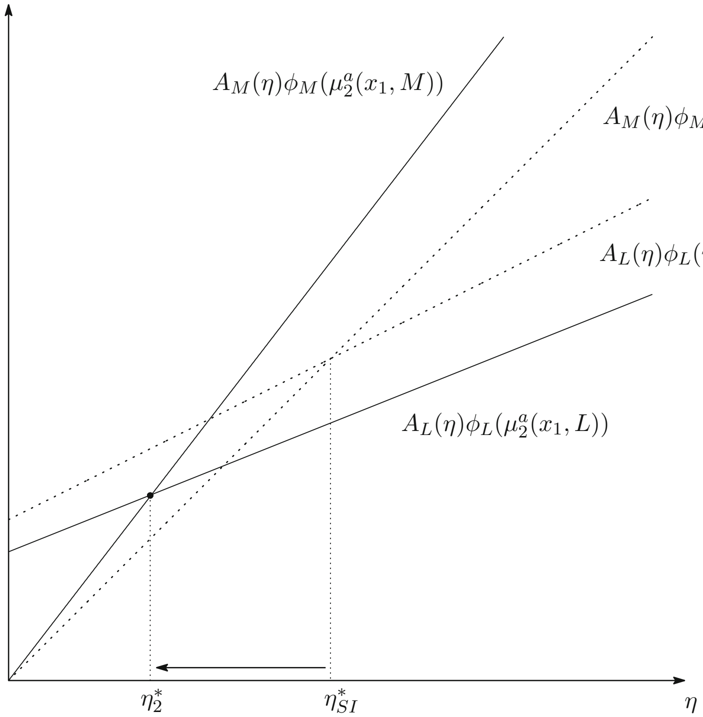


Fig. 19.1 The promotion threshold with the looking-glass effect

To see this more clearly, Fig. 19.1 depicts the impact of the looking-glass effect. In the absence of the looking-glass effect, the output level for an agent whose actual ability type is μ is $A_M(\mu)\phi_M(\bar{\eta})$ when assigned to job M while that is $A_L(\mu)\phi_L(\bar{\eta})$ when assigned to job L: the promotion threshold is then given by η_{SI}^* as depicted by two dotted lines. In the presence of the looking-glass effect, on the other hand, the output level is $A_M(\mu)\phi_M(\mu_2^a(x_1, M))$ when assigned to job M while that is $A_L(\mu)\phi_L(m_2^a(x_1, L))$ when assigned to job L. Since $\mu_2^a(x_1, M) > \bar{\eta} > \mu_2^a(x_1, L)$, the presence of the looking-glass effect unambiguously lowers the promotion threshold in period 2.

We now examine the impact of the initial job assignment upon the subsequent job assignment. To this end, we make several assumptions regarding the distribution of the disturbance F:

(A.1) F is twice continuously differentiable;
(A.2) $f(\tilde{a}) > 0$ for all $\tilde{a} \in [0, 1]$;
(A.3) $f'(\tilde{a}) \geq 0$ for $\tilde{a} \in [0, 0.5)$ and $f'(\tilde{a}) \leq 0$ for $\tilde{a} \in [0.5, 1]$;
(A.4) $f(\tilde{a})\tilde{a} < 2F(\tilde{a})$ for $\tilde{a} \in [0, 0.5)$ and $f(\tilde{a})(1 - \tilde{a}) < 2(1 - F(\tilde{a}))$ for $\tilde{a} \in [0.5, 1]$.

Those assumptions are purely technical and do not carry any significant meanings. There is indeed a large class of distributions that can satisfy those restrictions. Potential candidates are a uniform distribution over [0, 1] and its close variants. Also, any single-peaked distribution (which attains its maximum at 0.5) with sufficiently fat tails, i.e., sufficiently large $f(0)$ and $f(1)$, can satisfy those restrictions.[7] Under those assumptions we can now state the next result.

Proposition 3 *Under (A.1)–(A.4), the promotion threshold is lower when* $x_1 = M$ *than when* $x_1 = L$*, i.e.,* $\eta_2^*(M) < \eta_2^*(L)$*.*

Proof See Appendix.

The proposition indicates that, although the promotion rule becomes more lenient in the presence of information asymmetry, this tendency is more prevalent when the agent was previously assigned to job M. This is because ability matters more at job M, and hence it reveals more accurate information about the agent. In other words, a failure at an ability-intensive job is a sign of incompetence more than a failure at a job which requires only manual labor, and a demotion from job M to job L is hence more demoralizing than no promotion (being assigned to job L in both periods). This indicates that the agent's productivity now becomes history-dependent. We argue that this can be one of the reasons why we rarely observe demotions in organizations.

The driving force behind this result is the divergence in the accuracy of information held by the principal and the agent. When the agent is assigned to job M in period 1, the principal receives more accurate information regarding the agent's ability.[8] Since the divergence in the accuracy of information is large as a consequence, a stronger incentive arises for the principal to lower the promotion threshold to exploit the value of her private information. When the agent is assigned to job L in period 1, the principal's information is less perfect and the incentive to lower the promotion threshold is not as strong. To clarify this point, consider a limiting case where $p \to 0$. In this case, no relevant information is revealed when the agent is assigned to job L, and the subsequent job assignment in period 2 contains no informational value as a result. The model is then identical to the one without information asymmetry: the principal assigns the agent to job M in period 2 if and only if $\mu_2^p \geq \eta_{SI}^*$.

This result provides an implication which leads to what appears to be paradoxical promotion policies in internal organizations. Consider two agents, A and B, where agent A was assigned to job M in period 1 and agent B was not for some reasons. In period 2, new information arrives and the principal updates her belief accordingly. With the arrival of new information, there may arise a situation where agent B now

[7] It is important to note that those assumptions are sufficient but not necessary by any means for Proposition 3, which holds for a much wider class of distributions.

[8] In fact, since $p_M = 1$ in our model, the principal ends up with having perfect knowledge of the agent's ability after a period when the agent is assigned to job M in period 1.

appears to be more able than agent A, contrary to the initial expectation. Let μ_j, $j = A, B$, denote the principal's posterior belief on agent j. Even in this case, there is no change in the job assignment if $\eta_2^*(M) \leq \mu_A < \mu_B < \eta_2^*(L)$, despite the fact that the principal believes that agent B is more able than agent A. This result also implies a version of the Peter Principle where the principal keeps the agent at the same job even though he shows his incompetence at the job.

4.4 The First-Period Problem

We now turn our attention to the first-period problem. The job assignment in period 1 contains a problem absent in period 2 since assigning the agent to job M reveals more information about him. Let $G_x(a) \equiv p_x a + (1 - p_x)F(a)$ denote the probability (conditional on the job assignment) that the observed productivity is less than a; similarly, let $g_x(a) \equiv p_x + (1 - p_x)f(a)$ denote its corresponding density. The expected output in period 2 conditional on the initial job assignment (and the prior belief) is then given by

$$\begin{aligned} E(y_2 \mid x_1) &= \phi_L(\mu_2^a(x_1, L)) \int_0^{a^*(x_1)} A_L(\mu_2^p(x_1, a)) g_{x_1}(a) da \\ &\quad + \phi_M(\mu_2^a(x_1, M)) \int_{a^*(x_1)}^1 A_M(\mu_2^p(x_1, a)) g_{x_1}(a) da \\ &= G_{x_1}(a^*(x_1)) A_L(\mu_2^a(x_1, L)) \phi_L(\mu_2^a(x_1, L)) \\ &\quad + (1 - G_{x_1}(a^*(x_1))) A_M(\mu_2^a(x_1, M)) \phi_M(\mu_2^a(x_1, M)). \end{aligned} \tag{19.14}$$

Given that $\mu_1^p = \mu_1^a = \bar{\eta}$, the principal assigns the agent to job M in period 1 if and only if

$$\theta A_M(\bar{\eta})\phi_M(\bar{\eta}) + (1-\theta)E(y_2 \mid x_1 = M) \geq \theta A_L(\bar{\eta})\phi_L(\bar{\eta}) + (1-\theta)E(y_2 \mid x_1 = L), \tag{19.15}$$

which can also be written as

$$A_M(\bar{\eta})\phi_M(\bar{\eta}) - A_L(\bar{\eta})\phi_L(\bar{\eta}) \geq \frac{1-\theta}{\theta}(E(y_2 \mid x_1 = L) - E(y_2 \mid x_1 = M)). \tag{19.16}$$

One may conjecture that it is in general beneficial for the principal to collect more accurate information about the agent. As it turn our, though, this is not necessarily true in the presence of the looking-glass effect. Whether information revelation (through assigning the agent to job M) is beneficial from the principal's viewpoint is determined by the sign of $E(y_2 \mid x_1 = L) - E(y_2 \mid x_1 = M)$. When this term is negative, information revelation is on average beneficial and this tends to lower the promotion threshold in order to gain more accurate information about the agent. The direction of this information-revelation effect is in general ambiguous, however, depending on underlying characteristics of the environment. In this particular example, it depends on the shape of the cost function.

A graphic illustration may help clarify this point. For the sake of exposition we consider two extreme cases. Figure 19.2 depicts a case where the optimal effort level is strongly convex in the expected ability type so that $\phi_L(\mu_2^a(L, L)) \approx \phi_L(\mu_2^a(M, L))$. As the figure indicates, a demotion (a movement from job M to job L) is hardly demoralizing compared to no promotion (staying at job L in both periods) in this case: the principal can then benefit from more accurate information when $\phi_M(\mu_2^a(M, M))$ is sufficiently larger than $\phi_M(\mu_2^a(L, M))$. Figure 19.3, on the other hand, depicts a case where the effort level is strongly concave in the expected ability type so that $\phi_M(\mu_2^a(L, M)) \approx \phi_M(\mu_2^a(M, M))$. In this case, a demotion is highly demoralizing and the principal do not necessarily benefit from obtaining more accurate information.[9]

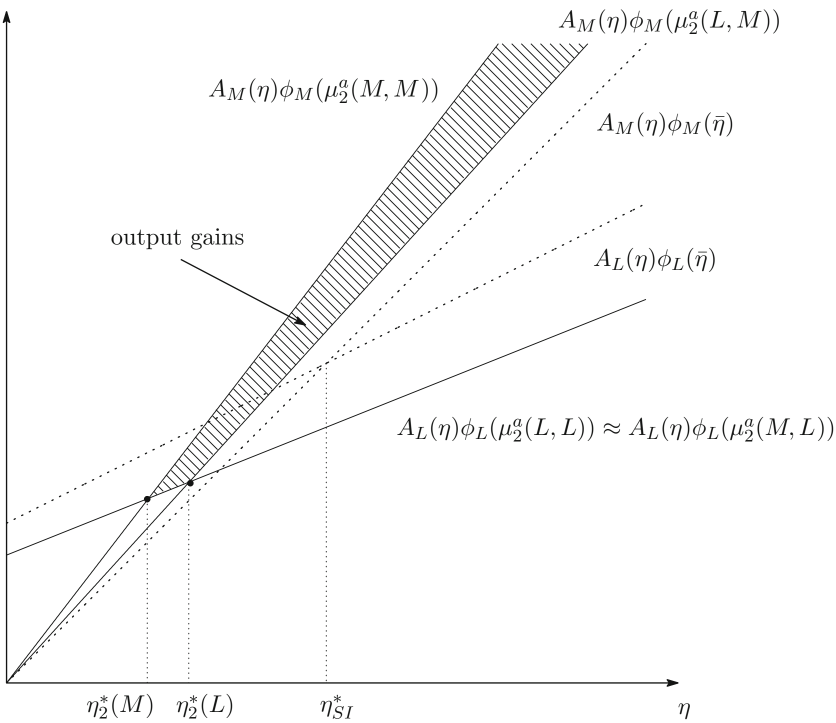


Fig. 19.2 Output gains due to more precise information

[9]This point indicates a possibility that the cutoff solution we have considered so far may not maximize the expected profit. When more accurate information does not lead to more profit, the principal can always ignore part of her information by randomizing the job assignment. As we stated earlier, however, we do not consider this type of mixed strategy since it is probably too complicated for the agent to grasp completely.

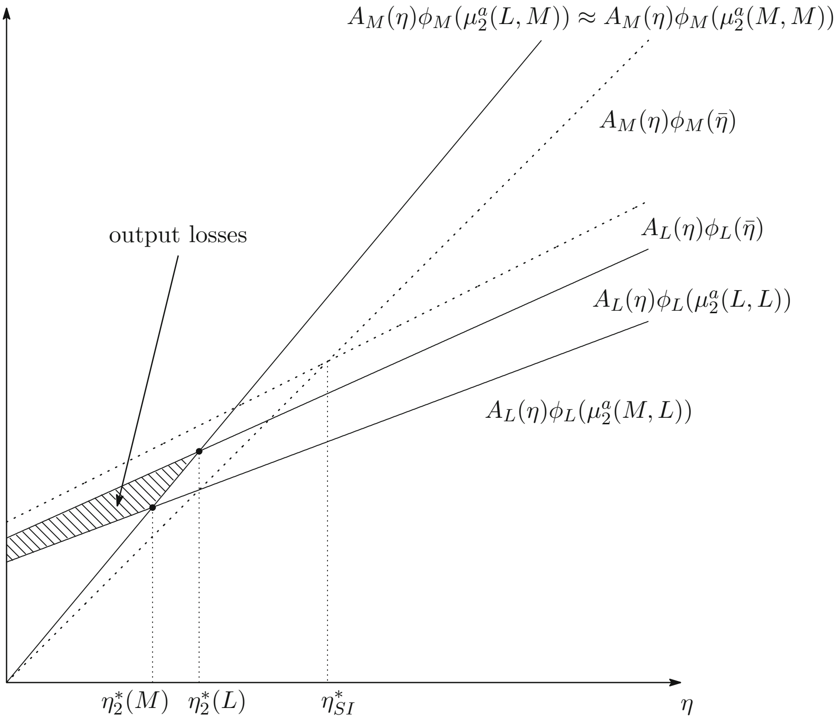


Fig. 19.3 Output losses due to more precise information

Note that $\eta_2^*(L) < \eta_{SI}^*$ by Proposition 2. It also follows from (19.16) that, as $\theta \to 1$, the promotion threshold in period 1, denoted by η_1^*, converges to η_{SI}^*. Now suppose that $\eta_2^*(L) < \bar{\eta} < \eta_{SI}^*$. In this case, as $\theta \to 1$, the principal initially assigns the agent to job L in period 1 but may choose to promote him in period 2 even when the principal receives no favorable information about the agent afterwards, i.e., $\mu_2^p(L, a_1) \approx \bar{\eta}$. This indicates that seniority may play a critical role when the principal learns about the agent more quickly than the agent himself. Notice that this mechanism draws clear contrast to the learning approach to explain delays in the timing of promotion. The learning approach emphasizes the fact that some attributes of a worker may be unobservable, and posits that a promotion to higher ranks occurs as more information is revealed over time. According to this view, a worker gets promoted when his employer receives favorable information about him. In the presence of the looking-glass effect, on the other hand, a promotion to higher ranks is triggered not necessarily by the arrival of favorable information about the agent per se; it is rather triggered by the very fact that the principal has more accurate information, regardless of its content.

5 Conclusion

This chapter considers a model of job assignment with the looking-glass effect. When a worker internalizes his superior's assessment of himself in the process of self-assessment, there arises a strategic aspect on the superior's part. In general, this strategic aspect of job assignment lowers the promotion threshold in that the least able agent to be promoted is less able with information asymmetry than without it. Moreover, in the presence of the looking-glass effect, the optimal promotion policy depends not only on the agent's expected productivity but also on the history of his previous job assignments: that is, the agent's productivity now becomes *de facto* history-dependent. This implies that it is not necessarily optimal to demote an agent to his previous job even when he turns out to be less productive at his new job, because a demotion undermines the agent's self-esteem more than no promotion. We argue that this effect partially explains why we rarely observe demotions in organizations.

Acknowledgements I am grateful to Roland Benabou and Muhamet Yildiz for helpful comments on an earlier version of the chapter. I also thank seminar participants at Contract Theory Workshop for helpful discussion. Financial support by the Japanese Ministry of Education, Culture, Sports, Science and Technology (Grants-in-Aid for Young Scientists (B)) is gratefully acknowledged.

Appendices

Appendix A: Proofs

Proof of Proposition 2: Define $\mu_2^a(x_1, x_2; a)$ as the agent's posterior belief when the promotion rule is such that the principal assigns the agent to job M if and only if $a_1 \geq a$. We can then show that

$$\mu_2^a(x_1, M; a) = \frac{p_{x_1} \int_a^1 \eta d\eta + (1 - p_{x_1})(1 - F(a))\bar{\eta}}{p_{x_1}(1 - a) + (1 - p_{x_1})(1 - F(a))}, \tag{19.17}$$

$$\mu_2^a(x_1, L; a) = \frac{p_{x_1} \int_0^a \eta d\eta + (1 - p_{x_1})F(a)\bar{\eta}}{p_{x_1} a + (1 - p_{x_1})F(a)}. \tag{19.18}$$

It is straightforward to see that $\mu_2^a(x_1, M; a) > \mu_2^a(x_1, L; a)$ for any a (as long as $p_{x_1} > 0$). Given this and $A_M(\eta_{SI}^*) = A_L(\eta_{SI}^*) = \bar{a}$, we can show that

$$A_M(\eta_{SI}^*)\phi_M(\mu_2^a(x_1, M; \eta_{SI}^*)) > A_L(\eta_{SI}^*)\phi_L(\mu_2^a(x_1, L; \eta_{SI}^*)). \tag{19.19}$$

Moreover, since $A_M(\mu) > A_L(\mu)$ for all $\mu > \eta^*_{SI}$,

$$A_M(\mu)\phi_M(\mu_2^a(x_1, M; \mu)) > A_L(\mu)\phi_L(\mu_2^a(x_1, L; \mu)), \tag{19.20}$$

for all $\mu > \bar{a}$. This proves that if there exists some η_2^* that satisfies (19.13), it must be that $\eta_2^* < \eta^*_{SI}$.

Q.E.D.

Proof of Proposition 3: Define $\omega(a)$ such that $\mu_2^p(M, \omega(a)) = \mu_2^p(L, a)$. Under the maintained assumptions, we first establish the following result.

Lemma 1 *Under assumptions (A.1)–(A.3), $\omega(a)$ is well-defined and strictly increasing in $a \in [0, 1]$.*

Proof Before we proceed, note first that

$$\mu_2^p(M, a) = a, \tag{19.21}$$

$$\mu_2^p(L, a) = \frac{pa + (1-p)f(a)\bar{\eta}}{p + (1-p)f(a)}. \tag{19.22}$$

It follows from these that $\omega(a) = \mu^p(L, a)$. To prove the lemma, it thus suffices to show that $\mu_2^p(L, a)$ is strictly increasing in a, i.e.,

$$\frac{d\mu_2^p(L, a)}{da} = \frac{p^2 + p(1-p)f(a) + p(1-p)f'(a)(\bar{\eta} - a)}{(p + (1-p)f(a))^2} > 0. \tag{19.23}$$

A sufficient condition for this is

$$p(1-p)f'(a)(\bar{\eta} - a) \geq 0, \tag{19.24}$$

which holds under (A.3).

Q.E.D.

The lemma indicates that ω is a well-defined function of a. We now show the next result which is critical for the proposition.

Lemma 2 *Under assumption (A.4), $\mu_2^a(M, L; \omega(a)) < \mu_2^a(L, L; a)$ and $\mu_2^a(M, M; \omega(a)) > \mu_2^a(L, M; a)$ for any given $a \in [0, 1]$.*

Proof We first show that $\mu_2^a(M, L; \omega(a)) < \mu_2^a(L, L; a)$. Suppose that, on the contrary to the claim, $\mu_2^a(M, L; \omega(a)) \geq \mu_2^a(L, L; a)$ for some $a \in [0, 1]$. Note that

$$\omega(a) = \frac{pa + (1-p)f(a)\bar{\eta}}{p + (1-p)f(a)}. \tag{19.25}$$

It also follows from (19.17) and (19.18) that $\mu_2^a(M, L; \omega(a)) \geq \mu_2^a(L, L; a)$ can be written as

$$\frac{\omega(a)}{2} \geq \frac{p \int_0^a \eta d\eta + (1-p)F(a)\bar{\eta}}{pa + (1-p)F(a)}. \tag{19.26}$$

We can then obtain from (19.25) and (19.26) that

$$\frac{pa + (1-p)f(a)\bar{\eta}}{p + (1-p)f(a)} \geq \frac{pa^2 + 2(1-p)F(a)\bar{\eta}}{pa + (1-p)F(a)}. \tag{19.27}$$

With some algebra, this can be written as

$$\begin{aligned} &(pa)^2 + p(1-p)F(a)a + p(1-p)f(a)a\bar{\eta} + (1-p)^2 f(a)F(a)\bar{\eta} \\ &\geq (pa)^2 + p(1-p)F(a) + p(1-p)f(a)a^2 + (1-p)^2 f(a)F(a), \end{aligned} \tag{19.28}$$

which is further simplified to

$$pf(a)a(\bar{\eta} - a) \geq F(a)(p(1-a) + (1-p)f(a)\bar{\eta}). \tag{19.29}$$

Since $f(a)\bar{\eta} > 0$, it is necessary that

$$f(a)a(\bar{\eta} - a) = f(a)a(0.5 - a) > F(a)(1-a). \tag{19.30}$$

This condition never holds, however, since $f(a)a < 2F(a)$ for $a \in [0, 0.5)$ under Assumption 4.

The second part of the lemma can be proved in a similar manner. Suppose that, on the contrary to the claim, $\mu_2^a(M, M; \omega(a)) \leq \mu_2^a(L, M; a)$ for some $a \in [0, 1]$. This condition can be written as

$$\frac{1 + \omega(a)}{2} \leq \frac{p(1-a)(1+a)/2 + (1-p)(1-F(a))\bar{\eta}}{p(1-a) + (1-p)(1-F(a))}. \tag{19.31}$$

It then follows from (19.25) and (19.31) that

$$\begin{aligned} \frac{pa + (1-p)f(a)\bar{\eta}}{p + (1-p)f(a)} &\leq \frac{p(1-a)(1+a) + 2(1-p)(1-F(a))\bar{\eta}}{p(1-a) + (1-p)(1-F(a))} - 1 \\ &= \frac{pa(1-a)}{p(1-a) + (1-p)(1-F(a))}. \end{aligned} \tag{19.32}$$

With some algebra we can obtain

$$\begin{aligned} &p^2 a(1-a) + p(1-p)(1-F(a))a + p(1-p)f(a)(1-a)\bar{\eta} \\ &+(1-p)^2 f(a)(1-F(a))\bar{\eta} \leq p^2 a(1-a) + p(1-p)f(a)a(1-a), \end{aligned} \tag{19.33}$$

which is further simplified to

$$(1 - F(a))(pa + (1 - p) f(a)\bar{\eta}) \leq pf(a)(1 - a)(a - \bar{\eta}). \tag{19.34}$$

Since $f(a)\bar{\eta} > 0$, it is necessary that

$$(1 - F(a))a < f(a)(1 - a)(a - \bar{\eta}) = f(a)(1 - a)(a - 0.5). \tag{19.35}$$

This condition never holds, however, if $f(a)(1 - a) < 2(1 - F(a))$ for $a \in [0.5, 1]$.

Q.E.D.

With those lemmas we can now prove the proposition. Suppose that the threshold ability type when $x_1 = L$ is given by $\eta_2^*(L)$. In this case, by definition, the principal assigns the agent to job M in period 2 if and only if $a_1 \geq a^*(L)$ where $\eta_2^*(L) = \omega(a^*(L))$. If there exists an interior solution, the threshold $\eta_2^*(L)$ must solve

$$A_M(\eta_2^*(L))\phi_M(\mu_2^a(L, M; a^*(L))) = A_L(\eta_2^*(L))\phi_L(\mu_2^a(L, L; a^*(L))). \tag{19.36}$$

Now consider an agent who was initially assigned to job M. By lemma 2, we can show that

$$A_M(\eta_2^*(L))\phi_M(\mu_2^a(M, M; \eta_2^*(L))) > A_L(\eta_2^*(L))\phi_L(\mu_2^a(M, L; \eta_2^*(L))), \tag{19.37}$$

which indicates that $\eta_2^*(M) < \eta_2^*(L)$.

Q.E.D.

Appendix B

The uniqueness of the optimal job assignment rule: Define $\mu_2^a(x_1, x_2; a)$ as in the proof of Proposition 2. There then exists at least one interior solution if

$$A_M(\mu_2^p(x_1, 1))\phi_M(\mu_2^a(x_1, M; 1)) > A_L(\mu_2^p(x_1, 1))\phi_L(\mu_2^a(x_1, L; 1)), \tag{19.38}$$

and

$$A_M(\mu_2^p(x_1, 0))\phi_M(\mu_2^a(x_1, M; 0)) < A_L(\mu_2^p(x_1, 0))\phi_L(\mu_2^a(x_1, L; 0)). \tag{19.39}$$

Suppose that these conditions hold. Then, the solution is unique if

$$\frac{d}{da}\Big[A_M(\mu_2^p(x_1, a))\phi_M(\mu_2^a(x_1, M; a)) - A_L(\mu_2^p(x_1, a))\phi_L(\mu_2^a(x_1, L; a))\Big] > 0. \tag{19.40}$$

Note that $A_M(\mu_2^p(x_1,a))\phi_M(\mu_2^a(x_1,M;a))$ is strictly increasing in a. Note also that as $p \to 0$, the agent's productivity is completely random and $A_L(\mu_2^p(x_1,a))\phi_L(\mu_2^a(x_1,L;a))$ thus becomes independent of a. This implies that (19.40) holds and there exists a unique interior solution when p is sufficiently small.

Addendum[10]

In economics, it is typically assumed that economic agents have superior, and often perfect, information about themselves. This is in clear contrast with social psychology where the issue of how we come to understand ourselves has always been a topic of utmost concern. A key concept in this regard is what is known as the looking-glass self, originally due to Cooley (1902), which states that people develop a sense of who they are by how others perceive them. Recently, there has been some attempts to incorporate this socio-psychological views into various economic problems; among the most notable along this line is Benabou and Tirole (2003).

The idea that people may learn about themselves though social interactions, especially others' evaluations, seem to bear particular importance in internal firm organizations where employees are constantly evaluated through means such as job assignments, promotions and pay raises. These actions are typically payoff-relevant and therefore reveal credible information about how competent they are in a given situation, which inevitably makes the whole process of employee evaluation more strategic in nature. The current chapter is one of the first works to investigate the impact of the looking-glass self on managerial strategies with particular focus on promotion decisions.

Researchers in this field now increasingly pay more attention to the possibility of the looking-glass self and its potential consequences on various aspects of human resource management. Just to name some, Nafziger (2011) considers a similar environment as the present chapter and augments it with endogenous wage setting to see the interaction between job assignment and incentive provision (through pecuniary compensation). She shows that it is more effective to motivate the agent by inefficient job assignment rather than rewarding him with a higher bonus, which illustrates why inefficient promotions are so ubiquitous in firm organizations. Crutzen et al. (2013) consider a cheap-talk model in which a manager faces a team of employees and must decide to what extent she differentiate those employees. In such a context, differentiation by means of comparative cheap talk can boost the self-image of the favored employee but undermine that of the unfavored one. They characterize conditions for the manager to refrain from differentiation, thereby providing an explanation for why managers are often reluctant to differentiate their employees, despite noticeable differences in productivity. Kamphorst and Swank (2013) consider a more complicated scenario in which the employee is concerned

[10]This addendum has been newly written for this book chapter.

with how the manager views his ability, aside from the possibility that the worker may also learn about himself from the manager's assessment of his ability. A key departure is that while the manager's high expectation enhances the employee's self-confidence as in the previous studies, her low expectation may challenge the employee to prove himself, due to his incentive to impress the manager. They then examine how the manager's attitude towards the employee affects his motivation and task choice.

Each of these works successfully sheds new light on well-documented anomalies of internal organizations which cannot be easily reconciled with a more traditional set of assumptions. Works along this line would provide further insights into the impact of the looking-glass on human resource management and, more broadly, improve our understanding of the role of psychological factors in the workplace.

References

Baker G, Gibbs M, Holmstrom B (1994) The internal economics of the firm: evidence from personnel data. Q J Econ 109(November):921–55
Benabou R, Tirole J (2003) Intrinsic and extrinsic motivation. Rev Econ Stud 70(July):489–520
Bernhardt D (1995) Strategic promotion and compensation. Rev Econ Stud 62(April):315–39
Cooley CH (1902) Human nature and the social order. Scribner, New York
Crutzen BSY, Swank OH, Visser B (2013) Confidence management: on interpersonal comparisons in teams. J Econ Manag Strateg 22 (Winter):744–767
Fairburn JA, Malcomson JM (2001) Performance, promotion, and the Peter principle. Rev Econ Stud 68(January):45–66
Falk A, Fehr E (2003) Why labour market experiments? Labour Econ 10(August):399–406
Fang H, Moscarini G (2005) Morale hazard. J Monet Econ 52(May):749–77
Gervais S, Goldstein I (2003) Overconfidence and team coordination. Mimeo, Duke University, Durham
Gibbons R, Waldman M (1999) A theory of wage and promotion dynamics inside firms. Q J Econ 114(November):1321–58
Hamilton B, MacKinnon M (2002) An empirical analysis of career dynamics and internal labor markets during the great depression. St. Louis, Mimeo: Washington University
Itoh H (2004) Moral hazard and other-regarding preferences. Jpn Econ Rev 55(March):18–45
James W (1890) The principles of psychology. [republished in 1950, Dover, New York]
Kamphorst JJA, Swank O (2013) When galatea cares anout her reputation: how having faith in your workers reduces their motivation to shine. Eur Econ Rev 60(May):91–104
Kandel E, Lazear EP (1992) Peer pressure and partnership. J Political Econ 100(August):801–817
Lazear EP (2004) The peter principle: a theory of decline. J Political Econ 112(February):S141–S163
Maskin E, Tirole J (1990) The principal-agent relationship with an informed principal, I: the case of private values. Econometrica 58(March):379–409
Maskin E, Tirole J (1992) The principal-agent relationship with an informed principal, II: common values. Econometrica 60(January):1–42
Mead GH (1934) Mind, self, and society. University of Chicago Press, Chicago
Myerson RB (1983) Mechanism design by an informed principal. Econometrica 51(November):1767–98
Nafziger J (2011) Motivational job assignments. Economica 78(October):676–696
Peter LJ, Hull R (1969) The Peter principle: why things always go wrong. Morrow, New York
Rotemberg JJ (1994) Human relations in the workplace. J Political Econ 102(August):684–717
Waldman M (1984) Job assignments, signaling, and efficiency. Rand J Econ 15(Summer):255–67

Part VII
Market Efficiency and Anomalies

Chapter 20
Is No News Good News? The Streaming News Effect on Investor Behavior Surrounding Analyst Stock Revision Announcement

Takahiro Azuma, Katsuhiko Okada, and Yukinobu Hamuro

> *Fundamentals might be good for the first third or first 50 or 60 percent of a move, but the last third of a great bull market is typically a blow-off, whereas the mania runs wild and prices go parabolic.*
>
> —Paul Tudor Jones

Abstract We investigate media influence on stock returns that are revised by sell-side analysts. Our main findings are twofold. First, post-announcement returns depend on whether the stock is covered by the media. Media-covered stocks demonstrate weaker post-announcement returns than their non-media-covered counterparts. Second, for media-covered event samples, we create a sentiment proxy using a unique news word count method and investigate whether pre-event sentiment affects post-event returns. Our results indicate that pre-event sentiment dictates short-run investor behavior and affects the post-announcement return in a significant manner.

Keywords News effect • Analysts' rate revision • Market sentiment

JEL Classification Codes G14

The original article first appeared in the International Review of Finance 14:29–51, 2014. A newly written addendum has been added to this book chapter

T. Azuma
JST ERATO Minato Discrete Structure Manipulation System Project, Osaka, Japan

K. Okada (✉) • Y. Hamuro
Institute of Business and Accounting, Kwansei Gakuin University, 1-155 Uegahara Ichibancho, Nishinomiya, Hyogo 662-8501, Japan
e-mail: katsuokada@kwansei.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_20

1 Introduction

In an efficient market, stock prices at any given time thoroughly reflect all available information. A priori, there is good reason to believe that stock markets are efficient, because such markets are paradigmatic examples of competition. Yet, rather than adjusting immediately to news surprises, stock prices tend to drift over time in the same direction as the initial surprise. When sell-side analysts change the ratings of stocks, short-run drift occurs. Previous research suggests two explanations for the existence and persistence of drift. First, the persistence of this anomaly may be due to high transaction costs (limits of arbitrage). Thus, mispricing persists only if market frictions are severe enough to prevent arbitrageurs from exploiting it. Barber et al. (2001), for example, present evidence that supports this view. They find significant drift in analysts' post-recommendation stock price returns; however, they conclude that their anomaly-based trading strategies do not reliably beat a market index after accounting for transaction costs. Alternatively, the drift may be a function of whether investors pay attention to the stock or the type of information investors receive about the stock. The second explanation comprises a behavioral view that investors face a formidable search problem. Barber and Odean (2008) predict that individual investors actively buy stocks on high attention days. They argue that professional investors as a whole (inclusive of market-makers) will exhibit a lower tendency to buy, rather than sell, on high-attention days and a reverse tendency on low-attention days. This will create a short-term overreaction followed by subsequent reversal.

The goal of this research is to deepen our understanding of what type of information flows drive event-related anomalies. Interest in the relations between media and the market has been growing among both researchers and practitioners (e.g. Klibanoff et al. 1998; Tetlock 2007; Tetlock et al. 2008). In the hedge fund industry, a London-based family office launched a Twitter-based investment fund that claims to invest in the stock market at the appropriate time through measuring market psychology. The origin of this investment is an academic paper published by researchers in computer science, Bollen et al. (2011). We contribute to this strand of research by examining the relation between post-event abnormal stock returns and the media. Specifically, we look at news coverage of stocks that face analysts' rating revision (obvious good/bad fundamental information about the stock) and how attention-grabbing and non-attention-grabbing stocks respond to the fundamental information. Our approach is similar to that of Fang and Peress (2009), who examine the cross-sectional difference in monthly returns depending on the news coverage, but differs from them in three ways. First, we examine not only headlines, but also massive and comprehensive amounts of news disseminated by the major financial information vendor in Japan. These data are more appropriate for our study than newspaper articles because they affect market participants directly on a real-time basis. Second, we look into the contents of the news and the impact of mass media sentiment. Specifically, we are interested in the frequency of mass media coverage and its effect on stock prices following firm events. To examine how mass media mood affects subsequent stock market returns, we categorize market news based on the number of positive and negative words appearing in the news articles.[1] Third, we focus on event-related abnormal returns to investigate how investors react to the arrival of new fundamental information in conjunction with the prevailing market news.

Our prediction is that upon arrival of upgrading news, attention-grabbing stocks would go up less than their no-attention-grabbing (non-media-covered) counterparts. Presumably, there are two effects at play. First, Bayesian updating investors would be less surprised upon the arrival of news when they have been exposed to any news in the past. Second, as pointed out by Barber and Odean (2008), attention-grabbing stocks are likely to be bought by individual investors and sold by professional traders. Because professional traders sell to individual investors above the fair value, attention-grabbing stocks are overvalued at the time of analysts' announcement, thus limited response to good news. Our prediction is symmetrical in the case of downgrades. Bayesian updating investors would be less surprised when the firm is mentioned in the news. Non-media-covered stocks are expected to go down more than their media-covered counterparts due to the surprise effect of the event news. By the same token, frequency of news coverage is a proxy for attention intensity; therefore, we expect that the greater the news coverage, the lower the magnitude of stock price response upon arrival of new fundamental information.

First, as a preliminary examination, we calculate the post-announcement abnormal returns of stocks whose ratings are revised by sell-side analysts. Using a standard event study framework, we find a significant abnormal price reaction even after the first tradable price on the day following the announcement. We also find significant abnormal returns using industry, size and book-to-market control firms as a benchmark. Consistent with prior research (Stickel 1995; Womack 1996), stocks upgraded by analysts demonstrate limited or small-scale post-announcement drift, while stocks that are downgraded indicate a prolonged downward drift.

Next, we collected a large amount of news electronically disseminated from the QUICK database. Our news sample includes articles from QUICK news, NQN news and Nihon Keizai Shimbun news between January 2008 and December 2012. A total of 773,386 news articles were obtained, consisting of 10,068,140 sentences and 56,358,567 words. Based on these news articles, we classified our sample firms (12,148 firm events) into two groups: media-covered (6,353 firm events) and non-media-covered (5,795 firm events). If a news report covers a firm during the 10 business days prior to its event date, we categorize the firm as media-covered and non-media-covered otherwise.

Consistent with our prediction, we find that media coverage significantly mitigates the post-announcement abnormal returns. Our results in the short-term post-announcement return analysis show that stocks mentioned in the media demonstrate less post-announcement return difference than firms with no media mention.

[1] Negative and positive terms are called 'polarity' words. The polarity of each word appeared in the news texts is determined based on our own created sentiment dictionary.

Upgraded stocks demonstrate a positive post-announcement return, on average, but with stronger magnitude for firms without media attention. For downgraded events, a negative post-announcement return also appears stronger for stocks without media attention than for their media-covered counterparts.

Third, to investigate whether the level of attention affects post-announcement return, we divide the media-covered samples into three groups: highly covered, medium covered and marginally covered. The post-announcement return for three levels of coverage is consistent with our predication in upgraded samples; lower coverage is associated with stronger post-announcement drift. For the downgraded samples, the difference is unclear.

Finally, we further categorize the media-covered samples (6,353 firm events) into three groups: positive, neutral and negative. When a stock is quoted in an article that contains more negative words than positive, as defined by our dictionary,[2] the stock is categorized as having negative sentiment. If the number of positive words and negative words offset each other in the article, it is categorized as neutral. Likewise, if the article contains more positive words than negative, it is categorized as positive.

Negativity and positivity are defined as the simple addition of each type of word's appearance in the news for the stock. Using this unique sentiment scoring method, we create a sentiment proxy and observed the post-announcement performance of three classes (positive, neutral and negative) of stocks based on the sentiment.

We find that downgraded firms show little difference in returns regardless of their sentiment class. Upgraded stocks, however, show a difference: stocks with positive sentiment demonstrate almost zero post-announcement return while neutral and negative sentiment stocks marginally show abnormal returns. Our findings using our original sentiment proxy suggest that when the contents of the news have more positive expression than negative (defined as carrying positive sentiment), the subsequent rise following upgrades is limited. Sentiment effects on downgrades remain unclear.

Our empirical findings are consistent with the view that a market with many Bayesian updating investors would provide a window of opportunities for trading unnoticed stocks. They are also consistent with the view presented by Barber and Odean (2008) that individual investors are trading overpriced attention-grabbing stocks and professional traders are the sellers of such stocks. For profit-seeking investors, when stocks are upgraded, it is wise to purchase non-media-covered stocks or stocks that are media-covered but to a lessor extent. Ceteris paribus, it is wise to avoid stocks that are heavily covered by the media.

The remainder of this chapter is organized as follows. Section 1 reviews the literature. Section 2 describes our data. Section 3 explains our methodology. Section 4 presents and discusses the main empirical results. In Section 5, does the robustness check of our results. Section 6 concludes the chapter.

---

[2] We created a sentiment dictionary that identifies each word as positive/negative. The dictionary contains approximately 1,500 positive words and 1,500 negative words. We read each sentence in an article and count how many positive/negative words are used in each sentence.

2 Literature Review

Klibanoff et al. (1998) show that country-specific news reported on the front page of *The New York Times* affects the pricing of closed-end country funds. The authors find that during weeks of front-page news, price movements are more closely related to fundamentals. Therefore, they argue that news events lead some investors to react more quickly. More recently, Tetlock (2007) analyzed the linguistic content of the mass media and reports that media pessimism predicts downward pressure and a subsequent reversal. Tetlock et al. (2008) further document that the fraction of negative words used in news stories predicts earnings and stock returns. These findings suggest that qualitative information embedded in news stories contributes to the efficiency of stock prices.

Among papers that examine broadly-defined media exposure, ours is the first that documents post-event returns and their relation with media coverage. Several recent papers document a positive relation between media and liquidity but fail to find significant return differentials. For example, Antweiler and Frank (2004) find that stock messages predict market volatility but their effect on returns is small. Grullon et al. (2004) document that firms with larger advertising expenditures have more liquid stocks, and Frieder and Subrahmanyam (2005) report that individuals are more likely to hold stocks with strong brand recognition. Fang and Peress (2009) actually succeed in finding return differentials using media coverage. They examine cross-sectional return patterns and find that media-covered stocks have lower returns than non–media-covered stock. Chan (2003) examines momentum and reversal patterns following large price moves with and without accompanying news and supports the same findings.

This chapter is closely related to those of Fang and Peress (2009) and Chan (2003) but differs in one important aspect: These authors focus on news coverage and headline news, respectively, but do not distinguish between news positivity and negativity. Since assessment of true value is difficult and investors overreact to private information and underreact to public information (Daniel et al. 1998), how a news article is written is as important as the factual information it conveys. We obtained data mainly from the major financial information vendor QUICK. To measure news sentiment, we enumerate negative and positive words in the relevant news articles that are electronically disseminated through QUICK. Another distinction is that Fang and Peress (2009) examine cross-sectional differences in returns with and without news coverage and Chan (2003) looks at market reactions to news in time (and the differences therein between winners and losers), whereas we examine post-event differences in returns.

This chapter is also related to that of Barber and Odean (2008), who show that individual investors are the net buyers of attention-grabbing stocks, such as stocks in the news. These authors argue that individuals face difficulties choosing stocks to buy from a large pool of candidates; thus, attention-grabbing stocks such as those in the news are more likely to enter their choice set. Our evidence implies that investors trade among attention-grabbing stocks but the direction of their investment decisions is affected by news sentiment.

3 Data

Our sample consists of companies subject to analyst recommendation revisions. The recommendation revisions are identified using Bloomberg's database. We use Bloomberg only to identify analysts' rating revisions because QUICK does not offer such data. The sample firms are listed on the Tokyo Stock Exchange (TSE) and the Japan Securities Dealers Association Quotation System (JASDAQ). The recommendation revisions encompass the period from 1 January 2008 to 31 December 2012. The Bloomberg database includes, among other items, revision dates, new ratings, identifiers for the brokerage house issuing the recommendation, and the identity of the analyst writing the report (if known). Recommendations are expressed by a rating of between one and five. A rating of one reflects a strong buy recommendation, two a buy, three a hold, four a sell and five a strong sell. This five-point scale is commonly used by analysts. If an analyst uses a different scale, we convert the analyst's rating to the five-point scale.

Another characteristic of our data is that the data made available to us are incomplete. Certain brokerage houses have entered into agreements that preclude their recommendations from being distributed by Bloomberg to anyone other than their clients. Consequently, although the recommendations of the largest and the most well-known brokers are included by Bloomberg, they are not part of our dataset. Our event data originally contain 15,796 observations for the period between 1 January 2008 and 31 December 2012. These data include cases of double counting, such as follows. Suppose on day t, Toyota is upgraded by an analyst X. A different analyst, Y, downgrades Toyota on the following day, $t + 1$. In this case, the post-event performance of Toyota is affected by the adjacent rating revisions in time. Therefore, we exclude event samples whose rating revision occurs multiple times in our event window. The remaining total event sample subject to analysis is 12,148 observations.

We also use the number of electronic news articles about a stock to proxy for the stock's overall media sentiment. To collect this information, we systematically searched the QUICK database for articles in our sample referring to the company name. The QUICK database distributes news data from three sources: *Nihon Keizai Shimbun*, QUICK and NQN. The news is all from the Nikkei Group but each source has its own characteristics. For example, the *Nihon Keizai Shimbun* news is an electronic version of the newspaper, with articles by Nikkei Inc. writers, while the QUICK news is market-focused and articles are by writers from QUICK Inc., a subsidiary of Nikkei Inc. The NQN news is the real-time distributed market news with articles by both Nikkei and QUICK writers.

We obtain the company name for each article from the article itself. A writer entering a story into the news systems, will often manually write the company name and occasionally its four-digit TSE code. The manual input of the company name leads to variations, such as *NTT-Docomo* or *Docomo*. We then match these company names with our code dictionary. When the article provides the company code, we tag the article with that code. We exclude from our analysis news articles about an industry without specific mention of a company. To capture news about a given

Table 20.1 The number of rating revision events occurred during our sample period by year. Large firms tend to be revised more than once during the calendar year. Event total indicates the number of revision events during the calendar year including firms that are subject to revision for more than once

Year	Market	Number of listed firms	Number of covered firms	Coverage ratio (%)	Number of firms that are revised more than once	Up ward revision	Down ward revision	Total
2008	TSE	2,170	973	45	532	1,172	1,848	3,280
	JASDAQ	915	147	16	58	94	166	
2009	TSE	2,282	999	44	468	1,588	1,105	2,775
	JASDAQ	882	149	17	25	43	39	
2010	TSE	2,691	1,484	55	405	1,442	1,052	2,551
	JASDAQ	996	597	60	12	30	27	
2011	TSE	2,083	1,514	73	342	939	909	1,888
	JASDAQ	961	688	72	6	20	20	
2012	TSE	2,094	1,458	70	300	699	928	1,654
	JASDAQ	919	595	65	7	11	16	
Total					2,155	6,038	6,110	12,148

company, we retain articles with at least one mention of the company. If an article mentions more than one company name, the article is counted multiple times, once for each company mentioned.[3] Table 20.1 displays the descriptive statistics of our samples.

We quantify the news media sentiment, that is, its negativity and positivity, for the selected articles. Converting qualitative text into a machine-readable form requires several preliminary steps, but we skip the details in this chapter because they are in the realm of computer science. To distinguish whether a story's informational content is positive or negative, one needs to prepare standards against which to classify words and events. Because different groups of people are affected by events differently and have various interpretations of the same events, conflicts can arise. For example, the term *dividend cuts* can be classified as negative by a prevailing dictionary-based algorithm. In contrast, it can be interpreted as positive by market analysts, who believe such conduct indicates the company is saving money and, therefore, is better able to repay its debts. To avoid such problems, we produce a dictionary of 3,056, terms classified by experts. We give each firm in our sample a time-series sentiment number if there was any news in the 10 calendar days prior to the analysts' recommendation revision event. Sentiment numbers are calculated based on the simple addition and subtraction of the news content about a firm. For example, if negative words outnumber positive words by two, the sentiment number for the firm is −2.

Table 20.2 describes the summary statistics of our sample in relation to the news articles and the sentiment score of each sample. We divide our sample firms

[3]The percentage of multiple-used articles in our sample is a mere 1.6 %.

Table 20.2 Panel A indicates the total number of events and articles covered. Panel B shows the description of news articles associated with the recommendation revisions. Panel C describes the distribution of firms covered by the media with its sentiment

Panel A: News articles with a company name

	No. of stocks	No. of events	No. of articles	No. of sentences	No. of words
Small	2,596	281	14,736	166,961	934,723
Medium	1,419	1,654	61,959	633,521	3,382,126
Large	822	10,213	696,691	9,267,658	52,041,718
Total	4,837	12,148	773,386	10,068,140	56,358,567

Panel B: News articles associated with recommendation revisions

		No. of firms subject to revision	No. of events	No. of articles within 10 days prior to the events	No. of positive words prior to the events	No. of negative words prior to the events
Downgrades	Small	86	186	405	31	73
	Medium	296	912	1,585	478	498
	Large	513	5,012	68,571	49,408	55,445
	Sub-total	895	6,110	70,561	49,917	56,016
Upgrades	Small	51	95	139	22	27
	Medium	252	742	1,305	370	307
	Large	511	5,201	68,303	49,642	51,855
	Sub-total	814	6,038	69,747	50,034	52,189
Total		1,709	12,148	140,308	99,951	108,205

Panel C: Media coverage and sentiment

		No. of media covered event*	No. of no-media covered event	No. of events with positive score	No. of events with neutral score	No. of events with negative score
Downgrades	Small	42	144	3	17	22
	Medium	354	558	85	188	101
	Large	4,076	936	993	1,593	1,490
	Sub-total	4,472	1,638	1,081	1,798	1,613
Upgrades	Small	20	75	4	10	6
	Medium	310	432	53	204	53
	Large	4,215	986	1,380	1,582	1,253
	Sub-total	4,545	1,493	1,437	1,796	1,312
Total		9,017	3,131	2,518	3,594	2,925

*No. of media covered event is based on all the news released, not limited to within 10 days prior to the events into three categories using market capitalization. Firms with market capitalization below 10 billion yen (US$111 million at the exchange rate of 90 yen per dollar) are categorized as small, those larger than 10 billion yen and less than 60 billion yen are categorized as medium, and those above 60 billion yen are categorized as large. Of 12,148 recommendation revisions, 10,213 are concentrated on large firms that represent merely one-sixth of all listed companies (of 4,873 listed firms, only 822 large companies are the subject of more than 80 % of news articles). As shown in Panel B of Table 20.2, out of 773,386 articles obtained from QUICK, 140,308 appeared during the 10 calendar days prior to the event. We calculate sentiment score based on news during that 10-day period. The score calculation is the simple addition of word polarity, with negative words scored as -1 and positive words as $+1$. A total of 99,951 positive words and 108,205 negative words appeared in the entire collection of news articles on our sample firms in the pre-announcement period. Panel C shows the composition of media-covered and non-media-covered sample firms. Out of 12,148 events, 3,131 were not media-covered in the whole pre-announcement period. The remaining 9,017 events had news coverage: 2,518 events have a positive score, 2,925 events have a negative score and 3,594 events have a neutral score.

4 Media Coverage and Stock Returns

This section focuses on the relation between media coverage and post-recommendation stock returns. We first examine the abnormal returns of recommendation revisions and then examine abnormal returns by subdividing the sample firms based on news sentiment.

4.1 Abnormal Returns of Stocks Revised by Sell-Side Analysts

Analysts deliberately plan most rating revisions and reiterations. These decisions are rarely made in haste. Although analysts act based on public information, the majority of the research suggests that market response to rating revisions is considerable. Stickel (1995) and Womack (1996) show that favorable (unfavorable) changes in individual analyst recommendations are accompanied by positive (negative) returns at their announcements. The authors document a post-recommendation stock price drift that lasts up to 1 month for upgrades and up to 6 months for downgrades. Although investors can exploit analyst information and generate abnormal profits, obtaining full information about analyst ratings ex ante is difficult. Generally, brokers only allow professional investors who have a trading account with them to fully access to their analysts' rating reports. In this sense, rating revision information is not completely in the public domain and individual investors are normally allowed to access partial or delayed information. An event study based only on news

Table 20.3 Average three-day cumulative abnormal return for firms that are upgraded and downgraded by analysts. Panel A describes results based on the benchmark return generated using the market model. Panel B shows the result based on the respective control firm. Control firm is chosen using industry, size and book-to-market criteria

	Total	Strong outperform	Outperform	Neutral	Underperform	Strong underperform
Panel A: Benchmark return based on market model						
Upgrade						
n	6,038	676	3,871	1,473	18	n/a
CAR	0.89 %	1.02 %	1.11 %	0.24 %	1.94 %	n/a
p-value	***0.000***	***0.000***	***0.000***	*0.102*	*0.305*	n/a
Downgrade						
n	6,110	n/a	576	4,001	1,205	328
CAR	−1.24 %	n/a	−0.79 %	−1.26 %	−1.54 %	−0.72 %
p-value	***0.000***	n/a	***0.000***	***0.000***	***0.000***	*0.059*
Panel B: Benchmark return based on industry, size and book-to-market adjusted control firm						
Upgrade						
n	6,038	676	3,871	1,473	18	n/a
CAR	0.81 %	1.23 %	1.00 %	0.12 %	2.20 %	n/a
p-value	***0.000***	***0.000***	***0.000***	*0.354*	***0.043***	n/a
Downgrade						
n	6,110	n/a	576	4,001	1,205	328
CAR	−1.09 %	n/a	−0.85 %	−1.09 %	−1.13 %	−1.32 %
p-value	***0.000***	n/a	***0.000***	***0.000***	***0.000***	***0.000***

available to the public would enable us to investigate whether the market discounts information in an efficient manner.

We define post-announcement drift as the return attainable by trading on the first tradable price after the rate revision announcement, which is the opening price of the first business day after the announcement. Table 20.3 indicates the average cumulative abnormal return (CAR) for a 3-day event window. The return is calculated from the opening price of the day following the announcement to the closing price of the third day. An abnormal return is defined as the sample return minus the benchmark return.

Panel A of Table 20.3 demonstrates the abnormal return based on the market model. We use the Tokyo Stock Exchange Tokyo Price Index (TOPIX) as a market portfolio proxy and the beta of each sample is estimated using 200 pre-event business days. In Table 20.3, the third column through seventh column indicate the destination of each upgrade and downgrade. For example, the 676 firms are upgraded from lower ratings to ‘strongly outperform’. These firms’ CAR is 1.02 % rejecting the null of a zero 3-day CAR. Note that rating revisions to ‘neutral,’ meaning the target stocks perform in line with the market index, significantly outperform or underperform depending on the path they follow. The stocks significantly outperform the market when they are announced to be upgraded to ‘neutral’ from ‘underperform,’ and underperform the market when they are announced to be downgraded to ‘neutral’ from ‘outperform’. This is consistent with Francis and Soffer (1997), who argue that investors.

It is expected that small capitalization stocks are more prone to analysts’ rating revisions than large capitalization firms. High book-to-market stocks tend to outperform the market when the value style is in sync with the market; therefore, size and value factors should be controlled. The industry can also be a determinant factor of returns, particularly when the market sector rotation is active. For example, a weak yen induces investors to invest in export-related industries. To control for these factors, we compare sample firm returns with the respective control firms based on industry, size and book-to-market ratio.

Panel B shows the abnormal returns using a control firm as a benchmark. The corresponding control firm is selected according to the following procedure. First, we select firms in the same industry as the sample, using the TSE’s middle industry classification code. Among stocks in the same industry, we select firms whose market capitalization falls between 70 % and 130 % of the sample. Finally, we pick a single stock whose book-to-market ratio is the closest to the firm’s. When there is no firm that satisfies these three selection criteria simultaneously, we drop the industry criterion and repeat the screening process. For 54 samples, we use only the size and book-to-market ratio criteria for selection.

The direction of the post-event period return in Table 20.3 is consistent with prior findings. Firms that are revised upward gain a positive abnormal return and those revised downward suffer from a negative abnormal return. A total of 6,038 stocks that are revised upward rise, on average, 0.89 % (0.81 % using control firms) more than expected. Symmetric results are found in downward revisions, with 6,110 firms losing −1.24 % (−1.09 % using control firms), on average, upon downgrade.

4.2 Media Coverage and Post-recommendation Returns

This subsection investigates whether the post-recommendation returns of firms that gain media attention are affected by the media coverage. In a long-run abnormal return analysis, Fang and Peress (2009) report that high-media-covered firms underperform their non-media-covered counterparts by 0.39 % per month (4.8 % per year). The authors also argue that non-covered stocks generate an alpha, while the high-covered stocks underperform the market index. They argue that the long-run performance difference caused by the media coverage is consistent with the hypothesis that investors demand a risk premium for stocks in oblivion. In contrast, according to Barber and Odean (2008), individuals face difficulties when choosing which stocks to buy from a large pool of candidates; thus, attention-grabbing stocks, such as those in the news, are more likely to enter their choice set. This buying pattern seems consistent with the media effect documented by Fang and Peress (2009) to the extent that individuals’ buying pressure temporarily pushes up the prices of attention-grabbing (in-the-news) stocks, but such pressure subsequently reverses. To investigate the cause of the media effect on stocks, we divide our sample based on the news occurrence in the 10 business days prior to the announcement. We categorize stocks as media-covered if there is any single news item about the stock in the 10 business days, and non-media-covered otherwise.[4] Then, we look at how stock prices behave with new fundamental information flow, such as analyst recommendation revisions.

The results of our short-run analysis of stock prices are consistent with the Bayesian updating investor hypothesis and individual investors' trading behavior documented by Barber and Odean (2008). We find that non-media-covered stocks generate stronger abnormal post-recommendation drift in both directions. Figure 20.1 represents the abnormal return that investors would receive by trading at the opening price after the announcement. The benchmark return is the respective control firm's return. If the announcement is made before the market close, investors are able to trade before the close; however, estimating the intraday tradable price adds complexity, so we assume all announcement revisions are made after the market is closed. As shown in Fig. 20.1, the non-media-covered subset of firms demonstrate stronger positive drift toward the new level than their media-covered counterparts. Symmetrically, the non-media-covered downgraded subset demonstrates more severe negative drift in the post-announcement period. There are presumably two effects at play here. One is that the Bayesian updating investors would be more surprised with the new fundamental information when there is no information about the stock before the announcement. Therefore, investors would react more sensitively to the new information; thus, non-media-covered stocks have a more severe post-announcement reaction. The other effect at play in our upgrading sample is presumably the attention effect. Stock price behavior surrounding rating upgrades is consistent with Barber and Odean (2008). If individual investors are trading media-covered stocks at overvalued prices, the stock may not go up as much upon the arrival of good news because it is already in a state of overvaluation.[5]

4.3 Media Coverage Frequency and Post-recommendation Returns

We define media coverage as the proxy for the degree of attention by individual investors. Barber and Odean (2008) find that individual investors have a higher

[4]We have experimented with different look back periods of −20 days, −15 days and −5 days. The number of sample stocks for covered and uncovered changes in each experiment but the implication of the result remains intact.

[5]Altıkılıç et al. (2013) argue that analysts' revisions are typically information-free and piggyback on news. Our evidence is consistent with their view. If some analysts are merely piggybacking their revisions on public information of news and events about the firm, revisions on media uncovered stocks are more likely to have information content and have stronger drift in returns.

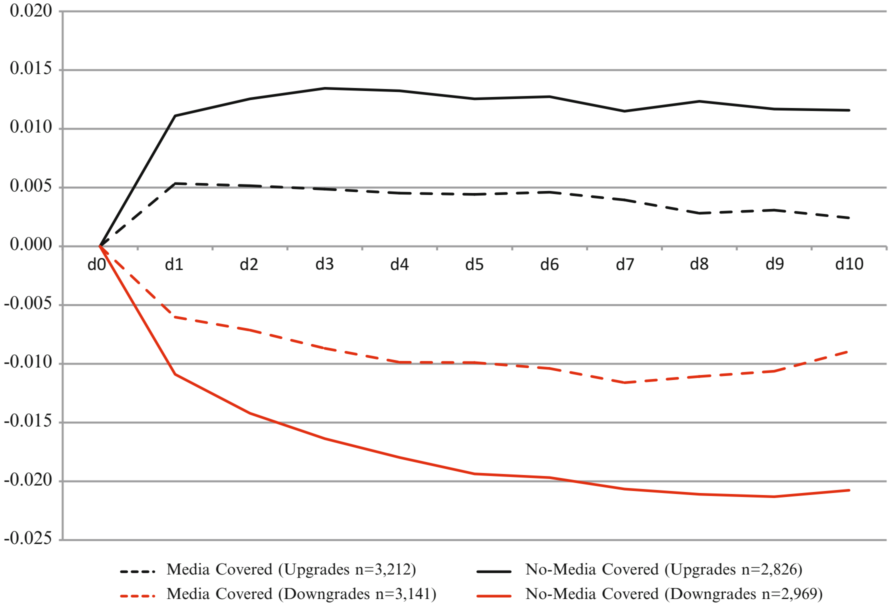


Fig. 20.1 Post-announcement performance by media coverage. Plot of cumulative abnormal return for the period of 10 business days after the announcement using the industry, size and book-to-market-based control firm. Cumulative return is calculated from the opening price of the following business day post-recommendation announcement (*dt*0). The *dotted line* indicates the cumulative abnormal return of stocks that are covered by the media, and the *solid line* stocks that are not media-covered. Among 6,038 upgraded stocks, 3,212 are media-covered and 2,826 not covered by the media. A total of 6,110 stocks are downgraded, with 2,969 media-covered and 3,141 not covered by the media

tendency to buy on high attention days. For every buyer there must be a seller. Professional investors as a whole exhibit a lower tendency to buy on high attention days. Therefore, stocks in high attention periods tend to be overvalued because professional investors would only agree to trade above the fair value. If this is the mechanism at play surrounding the revision, we should observe less post-announcement return in the stocks that are heavily covered by the media and more for those with lighter coverage.

To test this hypothesis, we have divided the media-covered samples into three subsets, heavily, medium and marginally covered, using the following criteria. The firms whose names are mentioned in more than 30 articles in the 10-day pre-event window are heavily covered; a mention in 10–30 articles is medium coverage; and those with fewer than 10 articles are marginally covered. If individual investors are buying on high attention days, it is likely that stocks heavily covered by the media are more overvalued than their marginally covered counterparts. We predict lower post-announcement drift for heavily covered firms.

Figure 20.2 represents the CAR that investors would receive when trading each subset of the sample firms at the opening price following the announcement.

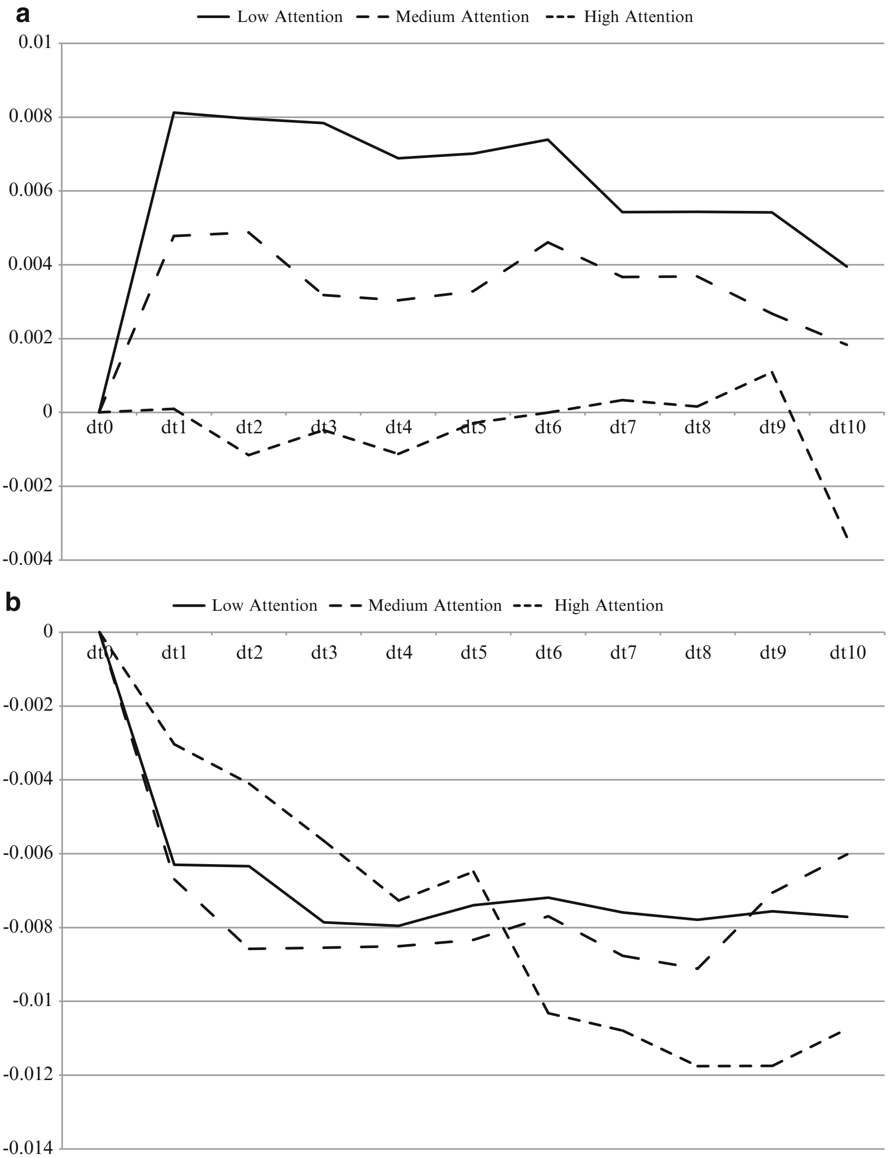


Fig. 20.2 Post-event cumulative abnormal return separately calculated for three subsets of samples. Low attention is defined as having less than 10 news articles disseminated during ten days prior to the event announcement date. Medium attention, 10–30 news articles; High attention more than 30 news articles

The results suggest that investors would be better off trading stocks that are less exposed to media. The greater the media coverage, the lower the post-announcement drift. Buying heavily media-covered stocks in accordance with the analysts' rating upgrades brings about zero abnormal return. Our evidence is consistent with the attention story for upgrading samples; however, downgrading samples generate an asymmetric result. There is no clear distinction among subsamples regardless of media coverage intensity. Conrad et al. (2006) argue that sell-side analysts are reluctant to downgrade in a timely manner. If this is the case, downgrading by analysts tends to occur when the outlook of the company becomes convincingly negative. This may be why downgraded samples drop in tandem regardless of the intensity of the media coverage.

4.4 Media Sentiment and Post-recommendation Returns

Measuring sentiment and its effect on stock market return has sparked the interest of many researchers in recent years. Our study is most closely related to Li (2006) and Davis et al. (2006), who analyze the tone of qualitative information using objective word counts from corporate annual reports and earnings press releases, respectively. Whereas Davis et al. (2006) examine the contemporaneous relationships between earnings, returns and qualitative information, Li (2006) focuses on the predictive ability of qualitative information, as do we. Li (2006) finds that the two words 'risk' and 'uncertain' in firms' annual reports predict low annual earnings and stock returns, which the author interprets as underreaction to 'risk sentiment.' Our study differs from Li (2006) in that we examine qualitative information in news stories at daily horizons rather than qualitative information in annual reports at annual horizons.

Some prior research analyzes qualitative information using more sophisticated subjective measures, rather than simple objective word counts. For example, Antweiler and Frank (2004) and Das and Chen (2006) design algorithms to reproduce humans' 'bullish,' 'neutral' or 'bearish' ratings of Internet chat room messages and news stories. Neither study finds any statistically significant return predictability in individual stocks. A study by Antweiler and Frank (2006), which uses an algorithm to identify news stories by their topic rather than their tone, does find some return predictability. For many of their topic classifications, Antweiler and Frank (2006) find significant return reversals in the 10-day period around the news, which they interpret as overreaction to news, regardless of the tone.

In this subsection, we concentrate our analysis on the fraction of words in *Nihon Keizai Shimbun*, *NQN* and *QUICK* stories about our sample firms and quantify the firm-specific sentiment based on the language used in the news. Merging the news stories and the financial information for a given firm requires matching firms' common names used in news stories. Although firms' common names usually resemble the firm names appearing in financial datasets, perfect matches are rare. To obtain the common names that we use as search strings for news stories, we begin with the company name variable in the Bloomberg data for all revised stocks during the relevant timeframe.

We obtain *Nihon Keizai Shimbun*, *NQN* and *QUICK* stories from QUICK terminal. For the period from January 2008 to December 2012 we collected 1,275,064 articles, or 68,740,386 words. We also include the date–time of submission (GMT + 0) and occasionally the contributor's name. Of 1,275,064 articles, 773,386 contain at least one company name. Because of the large number of firms and news stories, we implement an automated story retrieval system. For each target firm, the system constructs a query that specifies the sentiment of the stories to be retrieved. The system then submits the query and records the retrieved sentiment score. The sentiment score of the story is calculated as the number of positive words minus the negative words. To define the positivity and negativity of the text, we used a market expert, who constructed a dictionary of approximately 3,000 phrases. Each positive phrase is counted as +1 and negative as −1. The simple sum of these numbers per article defines the news sentiment score.

We illustrate the procedure for contents analysis for a sample firm, Mitsubishi Corporation (TSE code 8,058), a trading company whose market capitalization was US$36 billion in January of 2012. An analyst upgraded the firm on 18 January 2012. We look at news flow from 9 January 2012, which is 10 business days prior to the announcement. On 9 January, QUICK released a news story related to their airport management business. The story describes the government's new policy to sell the rights to manage domestic airports to the private sector. The sentiment score of this news is +6. When people read the news, it is generally agreed that the news is positive for the firm. On day $t-8$, there was no news. On day $t-7$, a story was issued about a copper mining company in Chile that sued the UK-based Anglo-American Co. Ltd. The sentiment score of the news is +1, but the contents are not necessarily positive. Our word count methodology has its own limitations in such a case because of its simplistic approach. However, in aggregate, it is unlikely to have many positive numbers if the firm's news story is bleak. The cumulative sentiment score over the 10-day period is +6 and Mitsubishi Corporation's subsequent 3-day CAR after the announcement is 5.03 %.

Table 20.4 reports typical examples of our sample stocks whose post-announcement returns are affected by sentiment in the 10 business days prior to the date.

Subsection 4.2 tests the hypothesis that investors trading behavior is influenced by whether investors are reminded of the stock through news. By comparing stocks covered by the media and stocks in oblivion, we find that the latter shows stronger post-announcement drift. When a stock is not mentioned by the media, the Bayesian updating investors do not have prior expectations of the stock; therefore, the arrival of new fundamental information moves the stock price by a larger margin. Our finding is also consistent with the view that individual investors are net buyers of attention-grabbing stocks at overvalued prices. Seasholes and Wu (2004) investigate the Shanghai Stock Exchange and find that individual investors are net buyers the day after a stock hits an upper price limit. Individual investors are attracted by the event of hitting a price limit (positive news) and individuals become the net buyers of stocks that catch their attention.

Table 20.4 Prior 10-day news appearance of four of our sample firms. The sentiment score of each day is the total number of positive word minus the negative word appearance of the day. We highlight some typical news for each stock of our choice

	CAR	Total	t-10	t-9	t-8	t-7	t-6	t-5	t-4	t-3	t-2	t-1
Mitsubishi Corporation upgraded on Jan 18, 2012	5.03 %		2012/1/8	1/9	1/10	1/11	1/12	1/13	1/14	1/15	1/16	1/17
Sentiment score		6	0	2	0	1	0	2	3	0	−2	0
Number of articles		15	0	1	0	4	3	1	1	1	4	0
Contents summary example	*Government announced domestic airport to be privatized. Mitsubisi Corp. has experience in it (t-9) *34bil dollar LNG project to start in Australia in 13th Jan. Mistubishi Corp has stake in it (t-5) *Mitsubishi Corp. and Hokuetsu Paper jointly open a state of the art plant to burn LNG (t-2)											
Mitsubishi Material Co. Ltd upgraded on May 29, 2009	7.99 %		2009/5/19	5/20	5/21	5/22	5/23	5/24	5/25	5/26	5/27	5/28
Sentiment score		4	0	0	−2	5	0	0	0	0	0	1
Number of articles		13	0	0	1	7	0	0	0	0	1	4
Contents summary example	*Mitsubishi Material to produce copper production with full capacity from August due to rising auto demand (t-7) *Lower coal price led five cement companies to have higher profit in the 2nd Q (t-2) *Non-ferrous metals industry is one of the best performers today (t-1)											
Sharp Corporation downgrade on Aug. 3, 2012	−2.32 %		2012/7/25	7/25	7/26	7/27	7/28	7/29	7/30	7/31	8/1	8/2
Sentiment score		−12	0	−8	0	−3	0	1	−2	1	−2	1
Number of articles		29	0	4	2	4	0	1	6	1	2	9
Contents summary example	*All electronic appliance companies are under pressure. Sharp made a new low (t-9) *Sharp developed smaller and lighter LCD panel for e-book. It is going to be in the market from Aug (t-5) *Sharp downward revised the earnings prospect. Also announced restructuring 5,000 workforce (t-1)											
Ricoh Co Ltd downgraded on April 27, 2012	−11.26 %		2012/4/17	4/18	4/19	4/20	4/21	4/22	4/23	4/24	4/25	4/26
Sentiment score		−11	0	0	0	−1	0	0	0	−3	−3	−4
Number of articles		−11	0	0	0	2	0	0	0	1	4	3
Contents summary example	*Nomura and its affiliated investor reduce their Ricoh's share holdings below 5 % (t-7) *Selling pressure mounts to the stocks that have high exposure to European economy including Ricoh (t-4) *Ricoh's earnings plunge to net loss, first time in its company history. 10,000 job cuts to follow (t-1)											

In this subsection, we test the hypothesis as follows. Among attention-grabbing (media-covered) stocks, the sentiment of the media determines the degree of overvaluation before the announcement. To test this hypothesis, we re-classify our media-covered sample firms into three subsets: firms with cumulative negative news scores on the day before an announcement, firms with positive news scores and their sample complement (neutral). Figure 20.3 illustrates the CAR up to 10 days after the event. When stocks are upgraded, firms with positive sentiment do not demonstrate positive abnormal returns, except as an initial reaction to the announcement. This is consistent with the hypothesis that positive news encourages individual investors to put in speculative bids; those bids are to be filled by professional traders at an overvalued level. Thus, the subsequent rise upon good news of such stocks is limited. The subset of the sample firms with neutral sentiment score and negative sentiment score do not have this effect. The difference between these two groups is statistically significant. We conjecture that positive sentiment in the news entices individual investors to trade at an overvalued price.

Again, we see an asymmetric result for the sentiment-based subsample analysis on downgrades. Regardless of the sentiment in the pre-event window, the stocks tend to drift downward in tandem. As discussed in subsection 4.3, downward stickiness in analysts' recommendation revisions (Conrad et al. 2006) may be making the sentiment factor trivial.

5 Robustness Checks

Post-earnings announcement drift, initial public offering (IPO) underperformance and delisting bias are well-documented return anomalies and, hence, we need to check that the media effect is not driven by them. These anomalies could lead to a spurious media effect if media coverage is more intense for firms announcing earnings, for IPO stocks or for stocks going through delisting. For example, if media coverage is biased toward bad earnings news, or if returns tend to drift more following bad earnings news compared to good earnings news, then, indeed, the non-media-covered perform better. A no-coverage premium would also result if high-coverage stocks are disproportionally represented by IPO stocks that subsequently underperform. Finally, if the media has a tendency to cover firms going through delisting for negative reasons (e.g. liquidation or takeover), then the delisting bias reported by Shumway (1997) could also lead to a spurious media effect.

To check that our results are not driven by post-earnings announcement drift, IPO underperformance or delisting stocks, we exclude all potentially earnings-related media coverage, all IPO stocks and all delisted stocks during our sample period. Our clean sample comprises 2,415 firm upgrade events and 2,560 firm downgrade events.

Table 20.5 indicates the post-event CAR up to 10 days into the post-announcement period. When stocks are upgraded, as described in Fig. 20.1, significant drift occurs. When we divide our sample into media-covered and non-

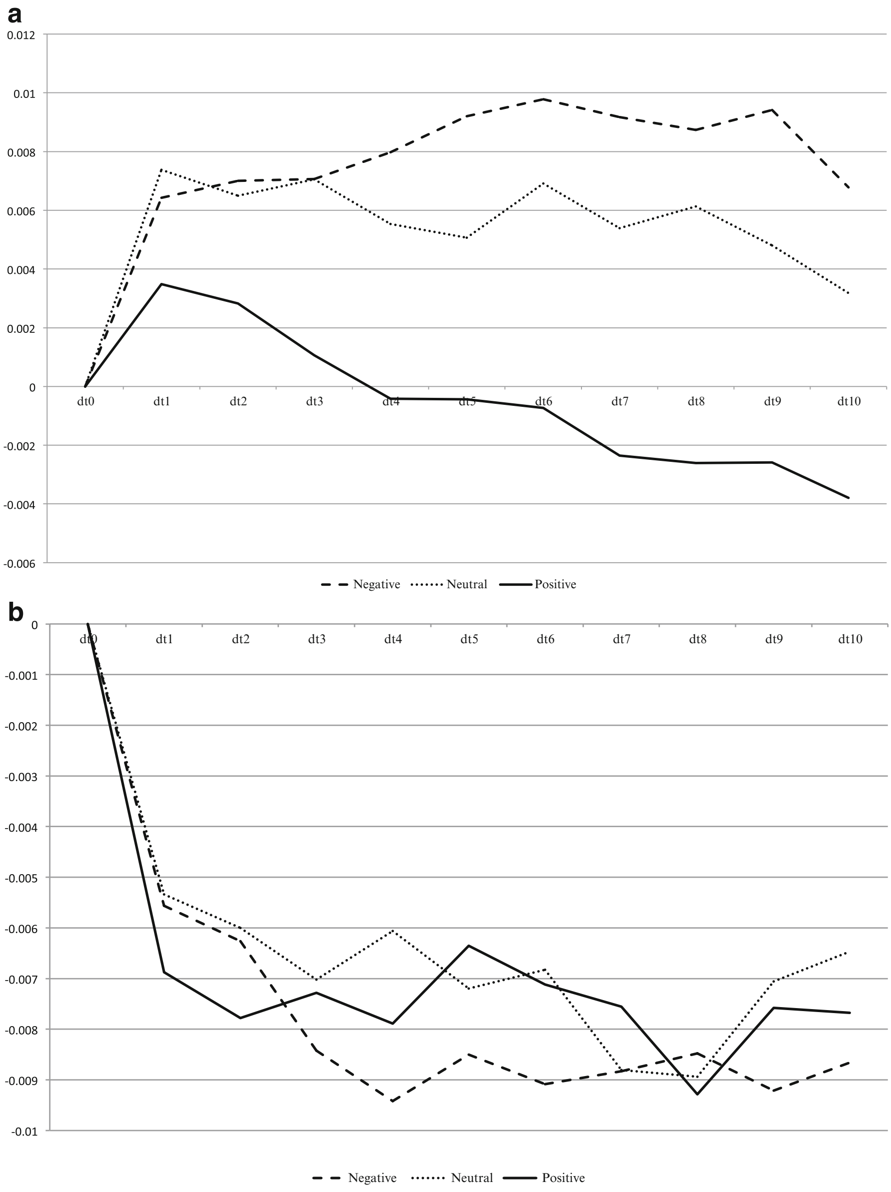


Fig. 20.3 Post-announcement performance by sentiment score. The 'negative', 'neutral' and 'positive' sub-sample sets are created based on the cumulative sentiment score over the 10-day period prior to the announcement date

Table 20.5 Post-event CAR up to 10 days into the post-announcement period

	No media cover	Media cover	Diff	*p-value*	No media cover	Media cover	Diff	*p-value*	Non-negative	Negative	Diff	*p-value*	Non-negative	Negative	Diff	*p-value*
Upgrades																
	Total sample (n = 6,038)				Clean sample (n = 4,605)				Total sample (n = 3,212)				Clean sample (n = 2,321)			
n	3,212	2,826			2,321	2,284			1,532	1,680			1,141	1,180		
CAR1	0.011	0.005	0.006	***0.000***	0.011	0.006	0.005	***0.000***	0.006	0.009	−0.003	***0.001***	0.006	0.009	−0.003	***0.030***
CAR2	0.013	0.005	0.007	***0.000***	0.013	0.006	0.007	***0.000***	0.005	0.010	−0.005	***0.000***	0.006	0.011	−0.005	***0.004***
CAR3	0.013	0.005	0.009	***0.000***	0.014	0.006	0.008	***0.000***	0.003	0.011	−0.008	***0.000***	0.004	0.012	−0.008	***0.000***
CAR4	0.013	0.005	0.009	***0.000***	0.015	0.006	0.009	***0.000***	0.002	0.011	−0.009	***0.000***	0.002	0.013	−0.010	***0.000***
CAR5	0.013	0.004	0.008	***0.000***	0.014	0.006	0.008	***0.000***	0.002	0.010	−0.009	***0.000***	0.002	0.013	−0.010	***0.000***
CAR6	0.013	0.005	0.008	***0.000***	0.014	0.007	0.008	***0.001***	0.002	0.011	−0.009	***0.000***	0.003	0.013	−0.011	***0.000***
CAR7	0.011	0.004	0.008	***0.000***	0.014	0.006	0.008	***0.002***	0.001	0.010	−0.008	***0.000***	0.002	0.013	−0.010	***0.000***
CAR8	0.012	0.003	0.010	***0.000***	0.015	0.006	0.009	***0.000***	0.000	0.010	−0.010	***0.000***	0.001	0.013	−0.012	***0.000***
CAR9	0.012	0.003	0.009	***0.000***	0.014	0.006	0.008	***0.002***	0.000	0.010	−0.010	***0.000***	0.001	0.013	−0.012	***0.000***
CAR10	0.012	0.002	0.009	***0.000***	0.014	0.005	0.008	***0.004***	−0.001	0.009	−0.011	***0.000***	0.001	0.012	−0.011	***0.000***
Downgrades																
	Total sample (n = 6,110)				Clean sample (n = 4,489)				Total sample (n = 3,141)				Clean sample (n = 2,189)			
n	3,141	2,969			2,189	2,300			1,200	1,941			914	1,275		
CAR1	−0.011	−0.006	−0.005	***0.000***	−0.011	−0.006	−0.005	***0.000***	−0.006	−0.009	0.003	***0.027***	−0.005	−0.009	0.004	***0.012***
CAR2	−0.014	−0.007	−0.007	***0.000***	−0.014	−0.007	−0.007	***0.000***	−0.007	−0.011	0.004	***0.007***	−0.006	−0.012	0.005	***0.008***
CAR3	−0.016	−0.009	−0.008	***0.000***	−0.016	−0.009	−0.007	***0.001***	−0.008	−0.013	0.005	***0.011***	−0.009	−0.014	0.005	***0.019***
CAR4	−0.018	−0.010	−0.008	***0.000***	−0.017	−0.010	−0.007	***0.002***	−0.008	−0.015	0.007	***0.001***	−0.008	−0.015	0.006	***0.009***
CAR5	−0.019	−0.010	−0.009	***0.000***	−0.019	−0.010	−0.009	***0.000***	−0.009	−0.016	0.007	***0.002***	−0.010	−0.015	0.006	***0.018***
CAR6	−0.020	−0.010	−0.009	***0.000***	−0.019	−0.010	−0.009	***0.001***	−0.009	−0.016	0.007	***0.008***	−0.010	−0.016	0.005	***0.031***
CAR7	−0.021	−0.012	−0.009	***0.000***	−0.020	−0.012	−0.008	***0.006***	−0.011	−0.017	0.006	***0.020***	−0.012	−0.016	0.004	*0.059*
CAR8	−0.021	−0.011	−0.010	***0.000***	−0.020	−0.010	−0.009	***0.003***	−0.010	−0.018	0.008	***0.005***	−0.011	−0.016	0.005	***0.043***
CAR9	−0.021	−0.011	−0.011	***0.000***	−0.020	−0.010	−0.010	***0.001***	−0.008	−0.018	0.010	***0.001***	−0.009	−0.016	0.008	***0.013***
CAR10	−0.021	−0.009	−0.012	***0.000***	−0.020	−0.008	−0.012	***0.001***	−0.005	−0.017	0.012	***0.000***	−0.005	−0.016	0.011	***0.002***

media-covered firms, we still find the same results. Table 20.5 indicates that the p-values are less than 1 % for CAR1 through CAR10 for both upgrades and downgrades. Statistical significance remains intact, even when we limit our analysis to the clean sample.

We subsequently conduct the same comparison for the sentiment score effect on post-event returns. We observe little difference between our total sample and the clean sample for either upgrades or downgrades.

6 Conclusion

We examine the effect of media coverage and media sentiment on investor behavior surrounding sell-side analysts' rating revisions. First, we find significantly stronger post-announcement drift when the stocks are not covered by the media. On average, stocks that are not featured in the media outperform the benchmark by over 1.35 % in the 3 days after the upgrading announcement and underperform by 1.64 % in the downgrading announcement. Our findings are consistent with the view that new fundamental information has stronger effect when Bayesian updating investors are not exposed to any news.

Second, we find significant return difference among media-covered stocks. For upgraded stocks, those with positive sentiment do not demonstrate positive announcement return. The stocks with high media exposure with positive sentiment are likely to be bought by naïve individual investors. Our result is consistent with the view that such attention-grabbing stocks are overvalued because professional investors sell to naïve investors at overvalued prices.

Finally, we show that the media effect is robust to the well-known post-earnings announcement drift, IPO underperformance and delisting bias anomaly. We provide test results for clean samples, excluding firms that are subject to these three biases, but the results remain intact. Interestingly, media coverage sentiment affects future returns (e.g. Tetlock 2007; Tetlock et al. 2008). The negative correlation between media sentiment score and post-event returns then suggests that that naïve investors, regardless of their fundamental news, long stocks when the media sentiment is positive and short stocks when negative. These observations suggest that the mass media's effect on security pricing stems from its ability to not only disseminate information broadly but also shape opinions or form consensus.

Acknowledgements The authors are thankful for the financial support by the Grant in Aid in Scientific Research (KAKENHI), JST-ERATO Minato Discrete Structure Manipulation System Project and Ishii Memorial Foundation for the Study of Securities Markets.

Addendum: News Effect on Pre-announcement Performance[6]

1. Attention Effect in the Pre-announcement Period

In this subsection, we investigate whether the pre-recommendation returns of firms are affected by the fact that they gained media attention. The relation between media attention and stock returns, if any, gives us a hint about the interaction between investor behavior and stock market returns. Fang and Peress (2009) find return differentials due to media coverage. They examine cross-sectional return patterns and find that media-covered stocks have lower returns than non–media-covered stocks. Chan (2003) examines momentum and reversal patterns following large price moves with and without accompanying news and supports the same findings. In this addendum, we examine pre-announcement price behavior with and without media attention.

Appendix Fig. 20.4 describes the pre-event abnormal returns based on our sub-sample sets. We divide our samples based on media coverage during the 10 business days prior to the announcement date ($dt-0$). As shown, the media-covered stocks to be upgraded (downgraded) generate greater positive (negative) abnormal returns in the run-up period than their non–media-covered counterparts. This can be interpreted to mean that streaming news conveys fundamental information about stocks and investors update their evaluations as news is disseminated. Interestingly, the media-covered stocks' return pattern reverses in the post-announcement period, as discussed in the main text.

One possible explanation of this reversal is that individual investors trade (long/short) attention-grabbing stocks and thus the stock price at the time of the announcement is overvalued (undervalued) (Barber and Odean (2008)). An alternative plausible explanation is that the media convey some fraction of the fundamental information; therefore, media-covered stocks are traded at a price that already partially discounts the good (bad) news. The latter interpretation is consistent with the work of Tetlock et al. (2008), who argue that the words contained in news stories are not redundant information but, instead, capture otherwise hard-to-quantify aspects of firms' fundamentals. Our conjecture is that both effects are behind the price move. The media not only convey future fundamental information but also affect investor behavior.

One of the characteristics of our data is those made available to us are incomplete. Certain brokerage houses have entered into agreements with their information vendors that preclude their recommendations from being distributed immediately after their release. Consequently, some of the analysts' recommendation information remains in the private domain for a few days before it becomes available to

[6]This addendum has been newly written for this book chapter.

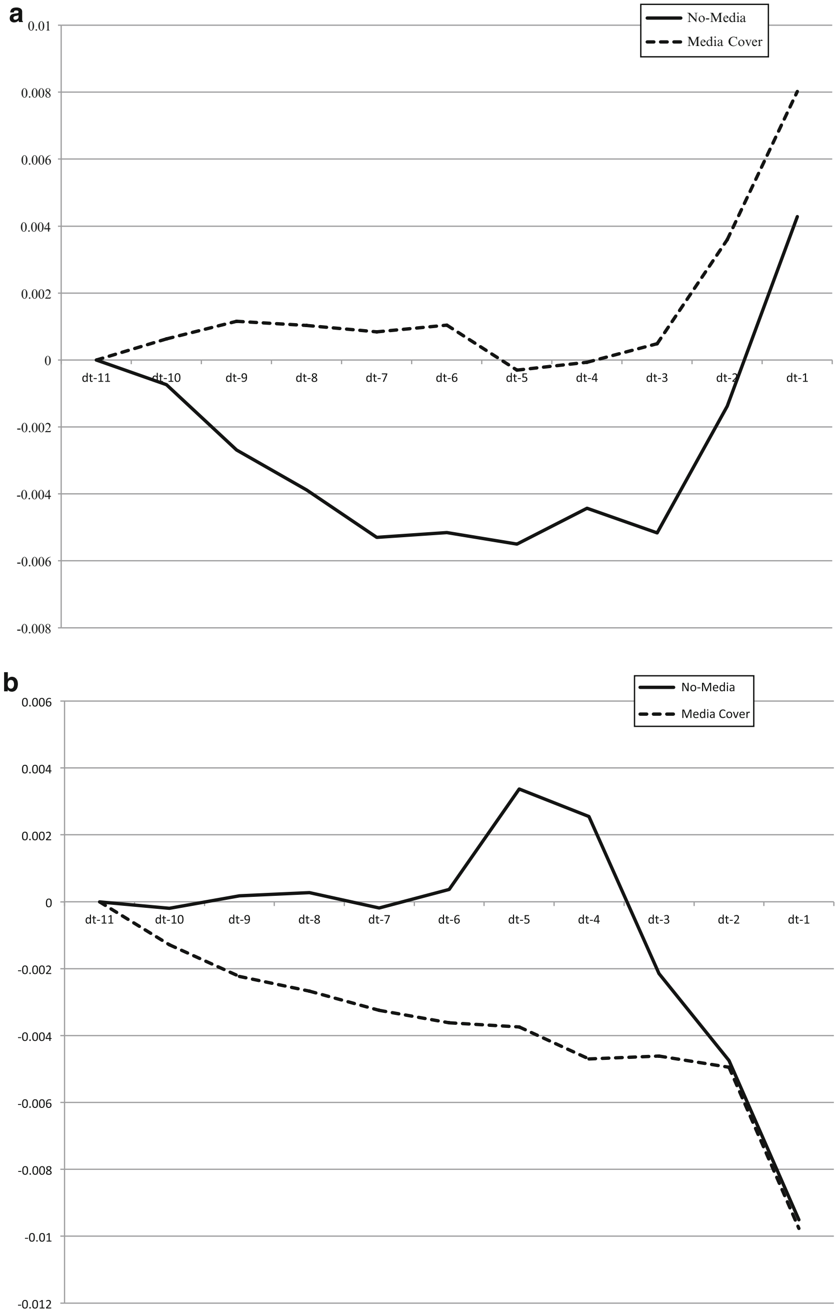


Fig. 20.4 The plot of the cumulative abnormal returns for (**a**) upgrades and (**b**) downgrades prior to the announcement. (*Notes*: The benchmark return is the return of the respective control firm, chosen based on industry, size, and book-to-market ratio. The market starts to react to analysts' upgrades starting three days prior to the announcement date. For downgrades, only media-covered stocks start to react three days prior)

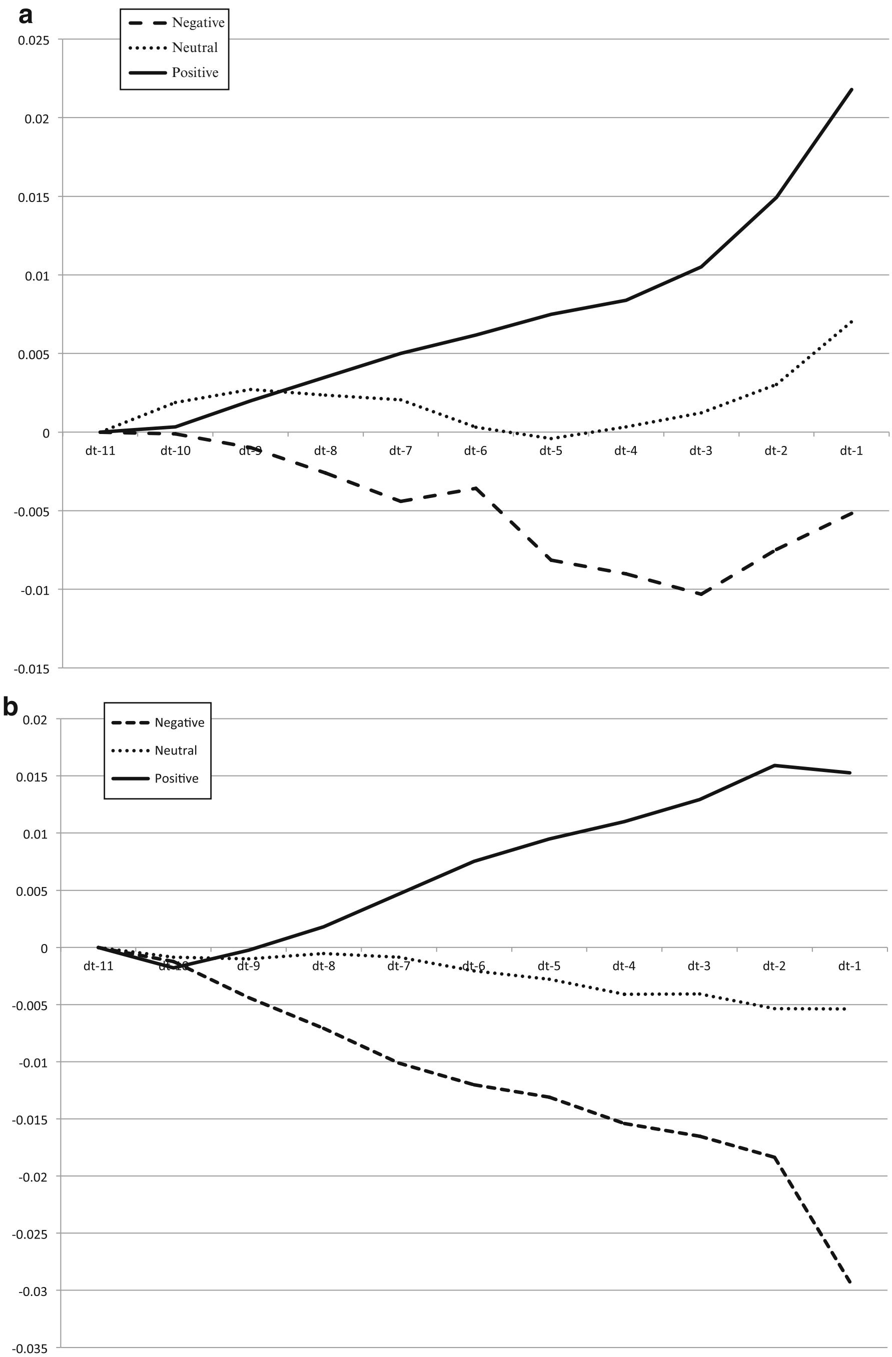


Fig. 20.5 The plot of cumulative abnormal returns for upgrades (**a**) upgrades and (**b**) downgrades prior to the announcement

the public. This enables certain privileged investors to act before others do. The abnormal return spike observed at $dt-2$ is primarily due to this time difference in information dissemination.

2. Sentiment Effect for Stocks in the Pre-announcement Period

In this subsection, we focus on media-covered stocks and investigate whether media tone – that is, whether negative, positive, or neutral – affects stock market returns in the pre-announcement period. As discussed in the main text, we use a simple word count method to proxy for the sentiment of streaming news. As shown in Fig. 20.5, positively reported stocks generate positive abnormal returns in the pre-announcement run-up period, while the neutral and negative subsets of the samples do not. This is primarily because investors obtain a fraction of otherwise hard-to-quantify fundamental information in positive news and discount it in the stock prices. In the run-up period, positively reported stocks rise two percentage points above their benchmarks. However, as shown in Fig. 20.2 in the main text, these stocks are the ones that demonstrate little positive abnormal return in the post-announcement period. This can be interpreted as the result of investors having discounted most of the good fundamental information. For downgraded stocks, the positively reported subset does perform well in the run-up period. This indicates that the fraction of fundamental information contained in the news is not relevant to the analysts' downward revision. The fact that the negatively reported subset performs poorly in the run-up period and deteriorates further beyond $dt-2$ is consistent with this irrelevancy conjecture.

References

Altınkılıç O, Balashov BV, Hansen RS (2013) Are analysts' forecasts informative to the general public? Manag Sci 59. doi:10.1287/mnsc.2013.1721

Antweiler W, Frank MZ (2004) Is all that talk just noise? The information content of internet stock message boards'. J Financ 59:1259–1293

Antweiler W, Frank MZ (2006) Do US stock markets typically overreact to corporate news? SSRN working paper. Available at http://ssrn.com/abstract=878091

Azuma T, Okada K, Hamuro Y (2014) Is no news good news? The streaming news effect on investor behavior surrounding analyst stock revision announcement. Int Rev Financ 14:29–51

Barber BM, Lehavy R, McNichols M, Source BT (2001) Can investors profit from the prophet? Security analyst recommendations and stock returns. J Financ 56:531–563

Barber BM, Odean T (2008) All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev Financ Stud 21:785–818

Bollen J, Mao H, Zeng X (2011) Twitter mood predicts the stock market. J Comput Sci 2:1–8

Chan WS (2003) Stock price reaction to news and no-news: drift and reversal after headlines. J Financ Econ 70:223–260

Conrad J, Cornwell B, Landsman WR, Rountree B (2006) How do analyst recommendations respond to major news? J Financ Quant Anal 41:25–29

Daniel KD, Hirshleifer D, Subrahmanyam A (1998) Investor psychology and security market and over-reactions'. J Financ 53:1839–1885
Fang L, Peress J (2009) Media coverage and the cross-section of stock returns. J Financ 64:2023–2052
Francis J, Soffer L (1997) The relative informativeness of analysts' stock recommendations and earnings forecast revisions. J Account Res 35:193–211
Frieder L, Subrahmanyam A (2005) Brand perceptions and the market for common stock. J Financ Quant Anal 40:57–85
Grullon G, Kanatas G, Weston JP (2004) Advertising, breadth of ownership, and liquidity. Rev Financ Stud 17:439–461
Klibanoff P, Lamont O, Wizman TA (1998) Investor reaction to salient news in closed-end country funds. J Financ 53:673–699
Seasholes MS, Wu G (2004) Profiting from predictability: smart traders, daily price limits, and investor attention. SSRN working paper. Available at http://ssrn.com/abstract=527182
Shumway T (1997) The delisting bias in CSRP data. J Financ 52:327–340
Stickel SE (1995) The anatomy of the performance of buy and sell recommendations. Financ Anal J 51:25–39
Tetlock PC (2007) Giving content to investor sentiment: the role of media in the stock market. J Financ 62:1139–1168
Tetlock PC, Saar-Tsechansky M, Macskassy S (2008) More than words: quantifying language to measure firms' fundamentals. J Financ 63:1437–1467
Womack KL (1996) Do brokerage analysts' recommendation have investment value? J Financ 51:137–167

Chapter 21
The Winner–Loser Effect in Japanese Stock Returns

Yoshio Iihara, Hideaki Kiyoshi Kato, and Toshifumi Tokunaga

Abstract This study examines the winner–loser effect using stocks listed on the Tokyo Stock Exchange (TSE) from 1975 to 1997. We uncover significant return reversals dominating the Japanese markets, especially over shorter periods such as 1 month. No momentum effect is observed, however. The 1- month return reversal remains significant even after adjusting for firm characteristics or risk. While the 1-month return reversal is not related to industry classification, it is partially a result of higher future returns to loser stocks with low trading volume. Our results show that investor overreaction may be a possible explanation for the 1-month return reversal in Japan.

Keywords Contrarian • Momentum • Predictability

1 Introduction

One of the striking empirical findings in recent financial research is the evidence of predictability in asset returns. Many articles in the recent literature find that mean stock returns are related to past stock price performance. Though accounting-related variables such as firm size, book-to-market equity, and cash flow to equity ratios are

The original article first appeared in the Japan and the World Economy 16: 471–485, 2004. A newly written addendum has been added to this book chapter.

Y. Iihara

H.K. Kato
Graduate School of Economics, Nagoya University, Nagoya, Aichi, Japan
e-mail: kato@soec.nagoya-u.ac.jp

T. Tokunaga (✉)
Musashi University, Tokyo, Japan
e-mail: t-tkng@cc.musashi.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_21

able to capture the cross-sectional variation in average returns, there is only a weak positive relation between average returns and beta using the Capital Asset Pricing Model (CAPM).

In their seminal papers, De Bondt and Thaler (1985, 1987) document return reversals over long horizons ranging from 3 to 5 years. Firms with poor past performance earn significantly higher returns in the subsequent period than those with above average past performance. This implies that a contrarian trading strategy performs well. In addition, winner–loser reversals seem to be related both to firm size and to the seasonal patterns of returns, especially January returns. Richards (1997) finds similar winner–loser return reversals in 16 national stock market indices after adjusting for risk.

On the other hand, several papers document that over medium-term horizons ranging from 6 to 12 months, stock returns exhibit momentum, that is, past winners continue to perform well and past losers continue to perform poorly in the following period. For example, Jegadeesh and Titman (1993) find that a strategy that buys past 6-month winners and shorts past 6-month losers earns approximately 1 % per month over the subsequent 6-month period. Rouwenhorst (1998) documents a similar return continuation in 12 European countries, which suggests that return continuation is a global phenomenon.

Although there is compelling empirical evidence that both contrarian and momentum strategies offer superior returns, the extant literature has failed to offer a conclusive explanation. There are two competing arguments explaining these anomalies with regard to market efficiency. The proponents of the efficient market view argue that the higher average returns from these strategies simply represent the reward from investing in risky stocks, which may not be captured by the CAPM. Fama and French (1993, 1996, 1998) propose a three-factor model in order to capture the cross-sectional variation in returns. Except for the continuation of medium-term returns, the anomalous patterns largely disappear in their three-factor model.

The opponents of the efficient market view take a behavioral approach. The superior return on these stocks is due to expectation errors made by investors. Investors overreact, and their excessive optimism or pessimism causes prices to be driven too high above, or too low below their fundamental values, and that the overreaction is corrected in a subsequent period. Similarly, investors under-react to information. They do not revise their own estimates in a timely fashion when they receive new information. As a result, asset prices do not fully reflect new information.

Daniel et al. (1998a, b) propose a theory of stock market over (or under) reaction based on two psychological biases—investor overconfidence about the precision of private information, and biased self-attribution—which causes asymmetric shifts in investor confidence as a function of their outcomes.[1] Investors tend

[1] Daniel and Titman (1999) also discuss investor overconfidence and market efficiency. Chan et al. (1996) find that medium-term return continuation can be explained in part by under-reaction to earnings information, but price momentum is not subsumed by earnings momentum.

to be overconfident about their estimates. The theory predicts negative long-lag autocorrelations, excess volatility, and positive short-lag autocorrelations.

Several studies attempt to investigate how return reversals and return continuation are related to other factors. For example, Moskowitz and Grinblatt (1999) have recently shown that a significant component of firm-specific momentum can be explained by industry momentum. Liew and Vassalou (1999) find that portfolios based upon firm size and book- to-market contain significant information about future economic growth, however, momentum-related portfolio returns do not seem to be related to future economic growth.

Several studies have focused on the Japanese stock markets concerning stock return regularities. Kato (1990) documents a long-term return reversal in which losers out- perform winners. However, winners do not perform badly in the subsequent period, unlike US firms. Furthermore, the January effect does not appear in these data. Bremer and Hiraki (1999) have recently documented a short-term return reversal using Japanese weekly stock returns. They find that loser stocks with high trading volume in the previous week tend to have larger return reversals in the following week. Chan et al. (1991) examine a cross-sectional relationship between portfolio returns and accounting- related variables such as earnings to price (E/P), firm size (F/S), book-to-market (B/M), and cash flow to price (C/P). They find that E/P, B/M, and C/P are positively related to returns, while F/S is a negative determinant. Kobayashi (1997) analyzes the relationship between B/M and the return reversals and documents that the B/M effect is independent of return reversals.

In this chapter, we conduct a comprehensive analysis of the winner–loser effect using stocks listed on the Tokyo Stock Exchange (TSE) during the period from 1975 to 1997. Our major findings are:

1. Return reversals dominate in Japanese stock markets, especially over short horizons such as 1 month.
2. No momentum effect is observed.
3. The 1-month return reversal is significant even after adjusting for firm characteristics and risk.
4. The 1-month return reversal is not related to industry classification and is weakly related to trading volume.
5. Investor overreaction may be the cause of the 1-month return reversal in Japan.

The next section documents anomalous patterns observed in Japanese stock returns by constructing portfolios based upon past performance. We uncover return reversals dominating Japanese stock markets, especially over short horizons. Our results are different from US findings, which document return reversals over long horizons ranging from 3 to 5 years. In addition, no momentum effect is observed in Japan though it is significant in many countries including the US.

In the third section, we investigate whether the Fama–French three-factor model or the characteristic model is able to explain return reversals in Japan. Our results show that these two models can explain most of the return reversals but neither model can successfully capture the 1-month return reversal in Japan. Like the momentum effect in the US, the 1-month return reversal in Japan may be an additional factor to be considered.

In the fourth section, we attempt to explain the 1-month return reversal focusing on three different factors, which are industry classification, trading volume and investor over- reaction. Though US studies document a relationship between industry classification and momentum effect, the 1-month return reversal in Japan is independent of industry classification. Our results regarding trading volume are also different from those of US studies. The 1-month return reversal is partially caused by higher future returns of loser stocks with low trading volume. Finally, focusing on each firm's fiscal year, we show that the 1-month return reversal is related to investor overreaction.

2 Do Patterns Exist in Japan?

Because of the popularity of technical analysis, both contrarian and momentum strategies have received a lot of attention from Japanese investors in past years. In this section, we attempt to identify specific patterns in Japanese stock returns using data covering longer periods with a variety of portfolio formation and holding periods. The data used in this study are from the database compiled by Pacific Basin Capital Market Research Center (PACAP) at the University of Rhode Island. This database contains a variety of information including monthly stock returns and accounting related values covering the period from 1975 to 1997.[2]

We form five equally weighted portfolios ranked on past performance following the approach of Jegadeesh and Titman (1993). The ranking variable used in this study is a stock's past compound raw return, extending back 1, 6, 12, 36, and 60 months prior to portfolio formation ($J = 1$, 6, 12, 36, 60 months). We also have five holding periods corresponding to each formation period ($K = 1$, 6, 12, 36, 60). As a result, we focus on 25 trading strategies with regard to length of formation and holding periods. We do not allow overlapping of formation periods when the portfolios are constructed.

Panel A of Table 21.1 reports mean monthly returns for the winner and loser portfolios as well as the zero-cost contrarian (loser minus winner) portfolio, for the 25 trading strategies. Significant return reversals are observed for all formation period portfolios. Loser portfolio returns exceed winner portfolio returns at all horizons. This is different from US studies, which show that at horizons of less than 1 year, a momentum effect is observed instead of return reversals.[3] Our results show that losers consistently outperform winners for all horizons. In other words, no momentum effects are observed in Japanese stock returns. In addition, the magnitude of the formation period returns decreases as the length of the formation period increases. The magnitude of holding period returns does not seem to change

[2] We exclude stocks with negative book equity. Since the PACAP data does not include consolidated financial statements, we use unconsolidated financial data to compute B/M.

[3] Richards (1997) and Rouwenhorst (1998) document a similar pattern in the world market.

Table 21.1 Mean monthly returns (percent) for winner, loser and contrarian portfolios

J	Formation period return	Holding period return				
		$K = 1$	$K = 6$	$K = 12$	$K = 36$	$K = 60$
Loser portfolios						
1	−8.9788 (−22.05)	1.7843 (3.59)	1.2699 (2.70)	1.1322 (2.47)	1.3815 (2.90)	1.7150 (3.60)
6	−2.7672 (−6.13)	1.7916 (3.40)	1.1951 (2.47)	1.1267 (2.43)	1.4247 (3.02)	1.7526 (3.75)
12	−1.6180 (−3.55)	1.5305 (3.00)	1.1452 (2.40)	1.1456 (2.50)	1.4694 (3.13)	1.7931 (3.83)
36	−0.2883 (−0.60)	1.7095 (3.54)	1.3653 (2.90)	1.3848 (2.99)	1.6627 (3.50)	1.9888 (4.13)
60	0.5674 (1.17)	1.8106 (3.74)	1.4838 (3.11)	1.4641 (3.10)	1.7230 (3.50)	2.0308 (4.02)
Winner portfolios						
1	14.0077 (24.43)	0.3898 (0.95)	0.7914 (1.90)	0.9366 (2.18)	1.2320 (2.65)	1.5885 (3.37)
6	6.0117 (12.35)	0.4295 (1.06)	0.8094 (1.86)	0.8963 (2.04)	1.1477 (2.42)	1.5069 (3.15)
12	4.4952 (9.51)	0.5773 (1.41)	0.7883 (1.78)	0.7892 (1.76)	1.0571 (2.19)	1.4239 (2.93)
36	3.1452 (6.64)	0.5154 (1.23)	0.6251 (1.40)	0.6179 (1.36)	0.9151 (1.86)	1.2970 (2.67)
60	3.1554 (6.25)	0.4608 (1.10)	0.5584 (1.26)	0.5894 (1.30)	0.8854 (1.83)	1.2739 (2.68)
Contrarian portfolios						
1	−22.9865 (−57.96)	1.3945 (4.70)	0.4786 (3.21)	0.1956 (1.94)	0.1495 (2.17)	0.1265 (1.88)
6	−8.7789 (−29.36)	1.3621 (3.59)	0.3857 (1.38)	0.2304 (1.07)	0.2770 (1.91)	0.2457 (1.89)
12	−6.1132 (−20.67)	0.9532 (2.62)	0.3569 (1.19)	0.3564 (1.37)	0.4123 (2.16)	0.3692 (2.13)
36	−3.4336 (−10.92)	1.1941 (3.65)	0.7402 (2.38)	0.7670 (2.61)	0.7476 (3.25)	0.6918 (2.91)
60	−2.5881 (−8.02)	1.3498 (4.64)	0.9254 (3.32)	0.8747 (3.27)	0.8375 (3.16)	0.7569 (2.70)

Winner and loser portfolios are formed based on J-month lagged returns and held for K-months. The stocks are ranked in ascending order on the basis of J-month lagged returns. An equally weighted portfolio of stocks in the lowest past return quintile is the loser portfolio and an equally weighted portfolio of stocks in the highest past return quintile is the winner portfolio. The contrarian portfolio is a zero investment portfolio, which is long in the loser and short in the winner. The average returns of these portfolios are presented in this table. t-statistics are in *parenthesis*

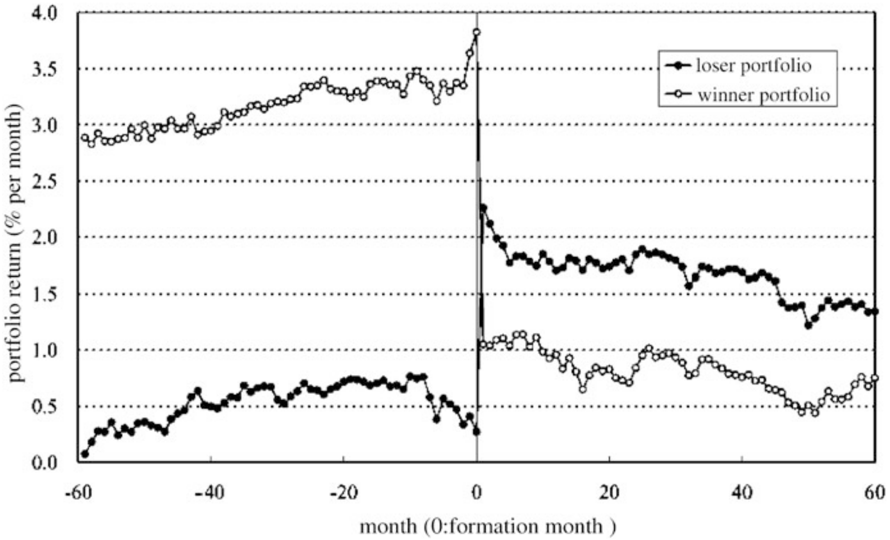


Fig. 21.1 Monthly returns (percent) of loser and winner portfolios in event time ($J = K = 60$)

across all horizons. As a result, return reversals are more pronounced over shorter periods.

In order to visualize the performance of winner and loser portfolios, we plot the evolution of the average returns of loser and winner portfolios before and after the time of formation ($J = 60$ months) as shown in Fig. 21.1.[4] At the time of formation, average portfolio returns take a big jump. The average returns of the winner portfolio fall from 3.7 % to 1.0 %. On the other hand, the average returns of the loser portfolio rise from 0.3 % to 2.3%. Significant return reversals occur for both portfolios.

After the time of formation, average returns of winner and loser portfolios change dramatically, but in opposite directions. This phenomenon may be attributed to the fact that more stocks in the winner (loser) portfolio are included in the loser (winner) portfolio in the following period. We calculate the portions of stocks included in the five ranking portfolios in the following period. The portions are not evenly distributed across the ranking portfolios. More stocks in the loser (winner) portfolio are included in the winner (loser) portfolio in the following period.[5] This implies that winner–loser reversals are not caused by a few outliers.

[4]We also plot the average returns before and after the formation for the portfolios ($J = 1$, 6, 12, 36). The patterns are similar and are not shown here.

[5]We conduct the same analysis for the $J = K = 1$ and the $J = K = 12$ portfolios. The results remain qualitatively unchanged.

In this section, we find strong return reversals in Japanese stock returns across a variety of formation and holding periods.[6] No momentum effect is observed, however. Our results are consistent with Daniel et al. (1998a) when no self-attribution bias is assumed. Interestingly, the return reversals are more pronounced for shorter formation periods, which is different from US findings. In addition, return reversals exist regardless of the length of holding period.

3 Risk/Firm Characteristics Adjustment

Loser–winner return reversals are clearly observed in Japan for a variety of formation periods and holding periods. On the other hand, momentum does not seem to exist in Japan, unlike the US. Our results suggest that the momentum effect observed in the US may not be a global phenomenon. Investor under-reaction, which may explain the momentum effect in the US, may not be valid in Japan. However, since we did not adjust for risk in the previous section, our results may not be entirely persuasive. Furthermore, the return reversals we observed may proxy for other variables such as B/M and F/S.

A number of studies document little cross-sectional relation between average stock returns and systematic risk measured either by market beta or consumption beta. Fama and French (1993) propose a three-factor model to explain the anomalous patterns of stock returns in the US. They find that their model largely captures average returns for US portfolios formed on F/S, B/M, E/P, C/P and other variables (past sales growth and long- term past returns) known to cause problems for the CAPM with the exception of medium- term past returns (momentum).

On the other hand, Daniel and Titman (1997) argue that Fama and French tests of their factor model lack power against an alternative hypothesis that they call the characteristic model. This model argues that the expected returns of assets are directly related to their characteristics for rational or irrational reasons. Using Japanese data, Daniel et al. (2001) reject the Fama–French three-factor model but fail to reject the characteristic model. Following the approach adapted by Daniel and Titman (1997), Davis et al. (2000), on the other hand, find the three-factor model provides a better description of the cross-sectional variation of stock returns by extending their sample period back to 1929.

Though the arguments are still inconclusive, both B/M and size are important factors to be considered. In this section, we investigate whether the characteristic model or the three-factor model can explain return reversals in Japan. First, we follow the procedure used in Daniel et al. (1997) for the characteristic model.

We form a set of 25 benchmark portfolios with similar stock characteristics, B/M and F/S. At the end of each June from 1975 to 1997, all TSE stocks in the sample

[6] We also conducted the same analysis by splitting the sample into two periods, the 1980s and the 1990s. We find that the return reversals are more pronounced during the 1990s.

are sorted into five equal groups from small to large based upon their F/S.[7] F/S is market capitalization at the end of June for each year. We also separately sort TSE stocks into five equal B/M groups from low to high. B/M is equal to the ratio of book value to market equity at the end of June for each year.[8] The 25 portfolios are created from the intersections of the five size and five book-to-market groups. Monthly equal-weighted returns for each of these 25 portfolios are calculated from June of year t to June of year $t+1$.

Using these 25 benchmark portfolio returns, we compute excess returns for the 25 past-performance-based trading strategies ($J = 1$–60 months, $K = 1$–60), as presented in Table 21.2.[9] Table 21.2 shows the performance of winner, loser and contrarian portfolios after adjusting for the characteristic premium. The characteristic model seems to explain most of the return reversals in Japan except for the shorter return reversals. Both winner and loser portfolios exhibit significant return reversals for the 1 month holding strategy ($J = 1$–60 months, $K = 1$). As a result, contrarian portfolios have significantly positive returns. The 1-month return reversal is larger as the formation period becomes shorter. Return reversals seem to disappear for longer holding strategies except the $J = 1$ month and $K = 6$ trading strategies. The 1-month return reversal may be another characteristic to be added to the characteristic model for Japan. This is somewhat different from the US evidence which finds evidence of a significant medium-term return continuation.

In order to test the robustness of our results, we apply the procedure used by Fama and French (1993) for their three-factor model. The following time series regression is estimated for each of the past-performance-based trading strategies.

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{i,HML}\,(R_{HML.t}) + \beta_{i,SMB}\,(R_{SMB.t}) + \beta_{i,Mkt,t}\,(R_{Mkt.t}) + \varepsilon_{i,t} \tag{21.1}$$

where $R_{i,t}$ is the return of past-performance-based portfolio i, and $R_{\mathrm{HML},t}$, $R_{\mathrm{SMB},t}$, $R_{\mathrm{Mkt},t}$ are respectively, the returns on the HML, SMB and Mkt factor portfolios at time t. $R_{\mathrm{f},t}$ is the risk-free rate at time t; $\beta_{i,j}$ is the factor loading of portfolio i on factor j. HML is a zero investment portfolio, which is long high B/M stocks and

[7]The majority of Japanese firms have their fiscal year ending in March (more than 80 percent of the firms in year of 2000), and essentially all companies publish their financial statements within 3 months after the end of their fiscal year. Accordingly, the portfolios are formed on the basis of the fundamental variables known to investors as of the end of June for firms with March and non-March fiscal year-ends. This ensures that our tests are predictive in nature.

[8]We took both book value of stock and number of shares issued from the balance sheet of the previous fiscal year. The data available to us are from parent-only financial statements. Though the consolidated financial statement has become more important over the last few years, parent-only statements had more influence on stock prices during our sample period.

[9]The excess return of a particular stock is computed by subtracting the benchmark portfolio's return from the stock's return.

Table 21.2 Excess returns (percent) adjusted by the characteristic model

	K				
J	1	6	12	36	60
Loser portfolios					
1	0.6343 (4.64)	0.1238 (1.85)	−0.0023 (−0.04)	−0.0066 (−0.15)	−0.0260 (−0.56)
6	0.4981 (2.90)	0.0204 (0.16)	−0.0672 (−0.73)	−0.0219 (−0.38)	−0.0344 (−0.64)
12	0.1875 (1.19)	−0.0990 (−0.78)	−0.0906 (−0.85)	−0.0301 (−0.44)	−0.0429 (−0.72)
36	0.2519 (2.04)	0.0102 (0.09)	0.0302 (0.30)	0.0874 (1.12)	0.0785 (1.04)
60	0.3183 (3.23)	0.1055 (1.14)	0.0967 (1.09)	0.1226 (1.32)	0.1117 (1.16)
Winner portfolios					
1	−0.7491 (−5.70)	−0.2241 (−3.65)	−0.0951 (−2.39)	−0.0732 (−2.23)	−0.0781 (−1.77)
6	−0.5283 (−3.56)	−0.0967 (−0.94)	−0.0066 (−0.08)	−0.0583 (−1.02)	−0.0767 (−1.27)
12	−0.3341 (−2.34)	−0.0482 (−0.42)	−0.0605 (−0.62)	−0.1008 (−1.42)	−0.1188 (−1.61)
36	−0.3278 (−3.29)	−0.1475 (−1.63)	−0.1502 (−1.64)	−0.1674 (−1.76)	−0.1845 (−1.64)
60	−0.3184 (−3.44)	−0.1842 (−2.02)	−0.1666 (−1.73)	−0.1790 (−1.52)	−0.1932 (−1.39)
Contrarian portfolios					
1	1.3833 (5.50)	0.3479 (2.89)	0.0927 (1.14)	0.0666 (1.34)	0.0520 (1.07)
6	1.0264 (3.34)	0.1171 (0.53)	−0.0606 (−0.37)	0.0364 (0.39)	0.0423 (0.52)
12	0.5216 (1.82)	−0.0508 (−0.22)	−0.0300 (−0.15)	0.0707 (0.59)	0.0759 (0.74)
36	0.5797 (2.76)	0.1577 (0.83)	0.1804 (1.00)	0.2548 (1.65)	0.2630 (1.60)
60	0.6366 (3.67)	0.2897 (1.75)	0.2633 (1.58)	0.3016 (1.59)	0.3050 (1.44)

Winner and loser portfolios are formed based on *J*-month lagged returns and held for *K*-months. The stocks are ranked in ascending order on the basis of J-month lagged returns. An equally weighted portfolio of stocks in the lowest past return quintile is the loser group and an equally weighted portfolio of stocks in the highest past return quintile is the winner. The contrarian portfolios are zero investment portfolios that are long in the loser and short in the winner portfolios. The excess return of a particular stock is computed by subtracting the benchmark portfolio's return from the stock's return. *t*-statistics are in *parenthesis* short low B/M stocks.[10] SMB is a zero investment portfolio, which is long small stocks and short large stocks.[11] Mkt is a zero investment portfolio, which is long the market portfolio and short the risk-free asset. We use the returns of an equally weighted portfolio of all stocks listed on the Tokyo Stock Exchange as a proxy for the market portfolio.

Since we have 25 different trading strategies based on past performance, we need to create time series observations of holding period returns for each strategy ($J = 1$–60 months and $K = 1$–60). However, we are not able to obtain a sufficient number of non- overlapping observations for the longer period trading strategies ($K = 36$ or 60) because of the limited length of our sample period. Therefore, we focus on 15 shorter period trading strategies ($J = 1$–60 months, $K = 1$–12). For each trading strategy, we form five portfolios based upon past performance. When J equals 1, we rebalance the performance-based portfolios every month. Accordingly, when K equals 6 (or 12), we rebalance the performance-based portfolios every 6 months (or every year). We estimate a time series regression for each of the five ranking portfolios for each strategy.

The intercepts and t-statistics of the Fama–French three-factor model are presented in Table 21.3.[12] The three-factor model seems to capture most of the return reversals in Japan. Most of the t-statistics are insignificant except for a few winner portfolios ($J = 1, 6$; $K = 1$) and a loser portfolio ($J = K = 1$). Our results show that the 1-month return reversal ($J = K = 1$) remains significant even after adjusting for risk using the three- factor model.

In this section, we investigated whether the characteristic model and the three-factor model, successfully explain return reversals in Japan. Most of the long-term return reversals disappear after adjusting for firm characteristic or risk. However, the short-term return reversals, especially for the $J = K = 1$ (1-month) trading strategy remain significant. We analyze this anomaly in more detail in the following section.

---

[10]At the end of June of year t, we form three groups using B/M for the fiscal year-end that falls between July of year $t-1$ and June of year t. B/M is equal to the ratio of book value to market equity at the end of June in year t.

[11]Similarly, we form two groups based upon F/S at the end of each June. F/S is market capitalization at the end of each June.

[12]For comparison purposes, we also consider a one-factor model (CAPM). Most of the t-statistics are significant for the $K = 1$ trading strategy, suggesting that the CAPM does not explain return reversals for a shorter holding period. However, the CAPM successfully explains return reversals for a longer holding period. This is similar to the results for the characteristic model in the previous section.

Table 21.3 Tests for alpha of Fama–French three–factor regressions

J	Rank 1 (loser)	Rank 2	Rank 3	Rank 4	Rank 5 (winner)
Holding period ($K = 1$)					
1	0.3900 (1.88)	0.0077 (0.05)	−0.1101 (−0.96)	−0.2831 (−2.24)	−0.7326 (−4.18)
6	0.2094 (0.84)	−0.0128 (−0.09)	−0.2472 (−2.15)	−0.2588 (−1.93)	−0.4043 (−2.13)
12	−0.2239 (−0.95)	−0.1785 (−1.28)	−0.1453 (−1.27)	−0.0664 (−0.49)	−0.0873 (−0.54)
36	−0.2462 (−1.40)	−0.1019 (−0.77)	−0.0336 (−0.27)	−0.1644 (−1.30)	−0.0903 (−0.70)
60	−0.0665 (−0.45)	−0.1024 (−0.77)	−0.1186 (−0.94)	−0.1609 (−1.26)	−0.1684 (−1.29)
Holding period ($K = 6$)					
1	3.0259 (1.00)	2.1832 (0.83)	1.2773 (0.52)	0.7542 (0.33)	−0.8726 (−0.29)
6	3.5457 (1.11)	2.6576 (0.89)	1.6602 (0.66)	0.4274 (0.17)	−1.8222 (−0.82)
12	2.5468 (0.81)	2.7954 (0.95)	2.2607 (0.85)	0.9001 (0.34)	−1.9534 (−1.03)
36	2.9521 (0.86)	2.7799 (0.91)	2.2900 (0.83)	0.3773 (0.12)	−1.7808 (−1.04)
60	4.2064 (1.18)	2.9113 (0.89)	1.7680 (0.69)	0.1424 (0.02)	−2.5476 (−1.43)
Holding period ($K = 12$)					
1	−0.9802 (−0.12)	−3.4364 (−0.49)	−4.3448 (−0.61)	−4.1756 (−0.55)	−3.6471 (−0.62)
6	−1.5401 (−0.13)	−3.2451 (−0.39)	−3.5423 (−0.53)	−3.8418 (−0.61)	−4.5366 (−0.76)
12	−2.4713 (−0.27)	−3.5075 (−0.43)	−3.6122 (−0.52)	−3.5514 (−0.57)	−3.5393 (−0.75)
36	−1.7609 (−0.15)	−2.1038 (−0.23)	−3.2134 (−0.43)	−5.3710 (−0.99)	−4.7516 (−0.93)
60	−0.4958 (−0.03)	−3.1191 (−0.34)	−3.3690 (−0.48)	−5.1701 (−0.99)	−5.2302 (−1.07)

$R_{i,t} - R_{f,t} = \alpha_i + \beta_{i,\mathrm{HML}}\left(R_{\mathrm{HML},t}\right) + \beta_{i,\mathrm{SMB}}\left(R_{\mathrm{SMB},t}\right) + \beta_{i,\mathrm{Mkt}}\left(R_{\mathrm{Mkt},t}\right) + \varepsilon_{i,t}$
$R_{i,t}$ is the monthly return of past-performance-based portfolio i in month t. $R_{f,t}$ is the 1 month risk-free rate. $R_{\mathrm{Mkt},t}$ is the monthly return of a zero investment portfolio, which is long the market portfolio and short the risk free asset. $R_{\mathrm{HML},t}$ is the monthly return of a zero investment portfolio, which is long high B/M stocks and short low B/M stocks. $R_{\mathrm{SMB},t}$ is the monthly return of a zero investment portfolio, which is long small stocks and short large stocks. t-statistics are reported in *parentheses*

4 Further Analysis

Most return reversals diminish after adjusting for risk or firm characteristics. However, neither the three-factor model nor the characteristic model (B/M and F/S) can successfully explain the 1-month return reversal ($J = K = 1$) in Japan. The 1-month return reversal can be considered risk or characteristic specific to Japan, similar to the momentum phenomenon in the US. In this section, we further examine this anomaly focusing on industry classification, trading volume and investor overreaction. We use excess returns adjusted by the characteristic model for the analysis in this section.

4.1 Industry Classification

Moskowitz and Grinblatt (1999) document a strong and prevalent momentum effect in industry components of stock returns that accounts for much of the individual

Table 21.4 Excess returns (percent) for performance-based portfolios ($J = K = 1$) by industry classification

	Formation period ($J = 1$)			Holding period ($K = 1$)		
	Manufacture	Finance	Other	Manufacture	Finance	Other
Loser	−9.0148	−6.3149	−8.8495	0.6163	0.6147	0.7751
	(−21.12)	(−14.68)	(−21.17)	(3.82)	(2.20)	(4.70)
Winner	14.2068	9.9341	13.6678	−0.7872	−0.2802	−1.0469
	(24.60)	(13.99)	(21.24)	(−5.58)	(−0.80)	(−5.19)
Contrarian	−23.2215	−16.2490	−22.5173	1.4035	0.8949	1.8221
	(−61.91)	(−26.23)	(−46.72)	(5.41)	(2.74)	(6.96)

In each month t, all stocks listed on the Tokyo Stock Exchange (TSE) are sorted into three groups (manufacturers, finance, or other) based on industrial classification. The stocks in the manufacturer, finance, and other groups are ranked in ascending order on the basis of 1 month lagged excess returns. An equally weighted portfolio of the stocks in the highest excess return quintile is defined to be the winner portfolio and an equally weighted portfolio of the stocks in the lowest excess return quintile is defined to be the loser portfolio. The contrarian portfolio is a zero investment portfolio, which is long the loser and short the winner portfolios. Excess returns are computed using the F/S and B/M benchmark portfolios. t-statistics are in parenthesis

stock momentum anomaly. Though momentum does not exist in Japan, the 1-month return reversal may be related to industrial classification. In order to examine this possibility, we sort 33 Tokyo Stock Exchange industry indices into five groups based upon past stock performance using the previous month's industry index returns. Each group contains six or seven industries. For each group, we compute an equally weighted return for each industry's index. We do not observe any particular patterns in stock returns after the formation period, however. Winners do not necessarily become losers (or winners) and losers do not become winners (or losers). We conclude that industry-based contrarian (or momentum) portfolios do not show superior performance. Although industry momentum exists in the US, we do not observe such patterns in Japan. In addition, we find no return reversals in industry-based portfolios.

Next, we examine the intra-industry effect. The 1-month return reversal may exist only in particular industries. In order to test this conjecture, we classify firms into three groups, manufacturers, financial firms and others. In each group, we conduct the same analysis as in the previous section by forming five performance-based portfolios every month ($J = K = 1$). One-month loser, winner and contrarian portfolios are created to examine excess returns in both the formation and the holding months. The results are presented in Table 21.4. Significant return reversals are observed for both winner and loser portfolios across all three industries. The contrarian portfolio returns are significantly positive across all three industries as well. Our analysis shows that the 1-month return reversal is not related to industry classification.

4.2 Trading Volume

Trading volume may be a proxy for the amount of information received by the market. Low trading volume may indicate that less information about a firm is available to investors. Because of limited information and low liquidity, a majority of investors, especially institutional investors, may stay away from these low trading volume stocks. These stocks are sometimes called neglected stocks. These neglected stocks are likely to be candidates for investor under or overreaction to new information. In other words, these stocks may be mispriced from time to time.

Conrad et al. (1994) show that return reversals are observed only for heavily traded stocks; less traded stocks exhibit return continuation. Lee and Swaminathan (2000) show that past trading volume provides an important link between momentum and value strategies. Firms with high (low) past turnover ratios exhibit many glamour (value) characteristics and earn lower (higher) future returns. In a related study, Brennan et al. (1998) use dollar trading volume as a proxy of liquidity and find low liquidity stocks earn higher returns than high liquidity stocks. Their results are consistent with the liquidity hypothesis.[13] Using Japanese weekly stock returns data, Bremer and Hiraki (1999) find that loser stocks with high trading volume in the previous week tend to have larger return reversals in the following week.

We examine the interaction between past returns and past trading volume in predicting future returns over a 1-month period. We use monthly turnover as a measure of trading volume.[14] We split the sample into three groups based upon trading volume and form five performance-based portfolios for each group. The results are presented in Table 21.5. The relationship between trading volume and the 1-month return reversal is not strong. The loser–winner reversal is more pronounced among low trading volume stocks, which is opposite to the result of Conrad et al. (1994). This is mainly caused by the higher future returns of low trading volume loser stocks. Our finding is different from that of Lee and Swaminathan (2000), which document higher future returns for both loser and winner stocks with high trading volume. Since low trading volume stocks do not always earn higher future returns, our results are inconsistent with the liquidity hypothesis. In addition, our results are somewhat different from Bremer and Hiraki (1999), which use weekly data.

Our results indicate that the 1-month return reversal is partially caused by the higher future returns of loser stocks with low trading volume. Low trading volume stocks tend to be neglected stocks which are candidates for investor overreaction. However, winner stocks with low trading volume do not have a similar pattern to loser stocks. In the following section, we examine investor overreaction in more detail by focusing on the turn of the fiscal year.

[13]According to the liquidity hypothesis, firms with relatively low trading volume are less liquid, and therefore, command a higher expected return.

[14]Turnover is defined as the ratio of the number of shares traded to the number of shares issued.

Table 21.5 Excess returns (percent) for performance-based portfolios ($J = K = 1$) by trading volume

	Formation period ($J = 1$)			Holding period ($K = 1$)		
	Low	Medium	High	Low	Medium	High
Loser	−8.4954	−8.7375	−9.0663	0.7571	0.7735	0.3104
	(−23.58)	(−20.85)	(−18.54)	(5.94)	(4.56)	(1.58)
Winner	9.0808	9.4236	20.5575	−0.6631	−0.5303	−0.8742
	(19.36)	(18.37)	(28.06)	(−3.34)	(−4.23)	(−4.81)
Contrarian	−17.5762	−18.1611	−29.6238	1.4202	1.3038	1.1846
	(−46.44)	(−56.02)	(−57.22)	(5.55)	(5.23)	(3.99)

In each month t, all stocks listed on the Tokyo Stock Exchange (TSE) are sorted into three groups based on trading volume (turnover ratio). The stocks in the low, medium and high volume groups are ranked in ascending order on the basis of 1 month lagged excess returns. An equally weighted portfolio of the stocks in the highest excess return quintile is defined as the winner portfolio and an equally weighted portfolio of the stocks in the lowest excess return quintile is defined as the loser portfolio. The contrarian portfolio is a zero investment portfolio, which is long the loser and short the winner portfolios. The excess returns are computed using the F/S and B/M benchmark portfolios. t-statistics are in *parenthesis*

4.3 Overreaction to New Information

In the previous section, we document the relationship between trading volume and the 1-month return reversal. Japanese investors may overreact to news and as a result, stock prices may overshoot temporarily and later come back to their fundamental values. In order to examine investor overreaction over the 1-month period, we focus on the fiscal year-end month because a large amount of information is released to the market at this time. For example, under the disclosure rules of the Tokyo Stock Exchange, Japanese firms must announce revisions to their financial forecasts if their actual audited results are likely to differ greatly from what they expected.[15] This kind of announcement is likely to take place toward the fiscal year-end because such accounting numbers become available at this time. Top management changes are also likely to occur during this period as a result; a change of corporate strategy may also be announced. Security analysts respond to the above information, revise their forecasts and release new forecasts toward the firm's fiscal year-end.

Since a majority of Japanese firms set their fiscal year-end in March, the performance of these stocks in March and April is worth examining.[16] If investors

[15]When actual sales differ by more than 10 % from the forecast, the firm must announce revisions. When actual operating income or actual net profit differs by more than 30 % from the forecasts, the firm must announce revisions. When the actual dividend differs more by than 20 % from the forecast, the firm must also announce the revision.

[16]More than 80 % of the firms listed on the Tokyo Stock Exchange ended their fiscal year in March, 2000.

Table 21.6 April excess returns (percent) for performance-based portfolios ($J = K = 1$) by fiscal year-end

	April excess returns		
Fiscal year-end	March	Non-March	H_0
Loser	0.5407 (1.10)	1.1611 (2.29)	(−0.88)
Winner	−1.3697 (−3.21)	0.5363 (1.06)	(−2.88)
Contrarian	1.9103 (2.27)	0.6249 (0.78)	(1.11)

Stocks are ranked in ascending order on the basis of 1 month returns at the end of March. The stocks included in winner (loser) portfolio are divided into two groups based on the company's fiscal year-end: March and non-March. An equally weighted winner (loser) portfolio is created for each group. The contrarian portfolio for each group is created by buying the loser and selling the winner. t-statistics are in *parenthesis*. t-statistics test whether the difference between April mean excess returns for March and non-March fiscal year-end firms equal zero

overreact to the new information, the bad (good) news firms are likely to be undervalued (overvalued) at the fiscal year-end month. We conduct our analysis focusing on the $J = K = 1$ trading strategy.

In order to test this conjecture, the stocks included in the winner (loser) portfolio are sorted into two groups based upon the firm's fiscal year-end: March fiscal year-end winner (loser) firms and non-March fiscal year-end winner (loser) firms. According to the conjecture, return reversals occur in April for the March fiscal year-end firms. The March fiscal year-end loser (winner) firms should experience higher (lower) returns in April than the non-March fiscal year-end loser (winner) firms. As a result, April contrarian portfolio returns using the March fiscal year-end winner and loser firms should be significantly positive and larger than those using non-March fiscal year-end winner and loser firms.[17]

The results are presented in Table 21.6. As predicted, April winner returns of March fiscal year-end firms are significantly negative and lower than those of non-March fiscal year-end firms. In addition, April contrarian portfolio returns of the March fiscal year-end firms are significantly positive. April contrarian portfolio returns of March fiscal year-end firms are higher than those of non-March fiscal year-end firms although the difference is not statistically significant. However, April loser returns of March fiscal year-end firms do not earn significantly higher returns than non-March fiscal year-end firms. Our results weakly support the conjecture that the 1-month return reversal is caused by investor overreaction around the turn of the fiscal year.

[17] Since the 1-month return reversal is more pronounced for low trading volume stocks, investor overreaction may also be related to trading volume. In order to test this conjecture, we separate our sample into two groups, high trading volume stocks and low trading volume stocks and conduct the same analysis. We did not observe any significant differences between low and high trading volume stock groups.

5 Conclusions

This study examines the predictability of Japanese stock returns focusing on past-performance-based trading strategies. Though stock returns exhibit momentum over short/medium-term horizons in the US, no such patterns are observed in Japan. Instead, return reversals are observed, especially for short-term portfolio formation strategies. The 1-month return reversal remains significant even after adjusting for risk or firm characteristics. Our results are different from US findings, which show short/medium-term momentum and long- term return reversals.

We further analyze the 1-month return reversal focusing on three factors, industry classification, trading volume and investor overreaction. Though the industry effect is related to momentum in the US, the 1-month return reversal in Japan is independent of industry classification. Trading volume is weakly related to the 1-month return reversal. Our results show that the 1-month return reversal is partially a result of higher future returns of loser stocks with low trading volume. This may be consistent with investor overreaction. Low trading volume stocks tend to be neglected stocks because less information is available to the investors about these securities. Investors are more likely to overreact to new information on these stocks.

We further examine investor overreaction focusing on stock returns around the turn of the fiscal year. Towards the end of the fiscal year, a variety of information is released to the market. If investors overreact to new information, stock prices are likely to be mispriced at the fiscal year-end month. We split the sample into March fiscal year-end firms and non- March fiscal year-end firms to examine this conjecture. Our results show that the 1-month return reversal is related to investor overreaction.

Overall, we find no return continuations but return reversals in Japan. Specifically, low trading volume losers in the previous month earn significantly higher returns in the subsequent month. Our results imply that the 1-month contrarian trading strategy concentrating on low trading volume stocks may be effective assuming that this pattern persists in the future. In addition, a 1-month contrarian portfolio using March fiscal year-end firms formed at the end of March may earn higher returns in April.

Acknowledgements We would like to thank Marc Bremer and the participants at the 2002 Korean Finance Association meeting and the 2002 Pacific Basin Finance Conference/Asian Pacific Finance Association meeting for their helpful comments. We are also very grateful to the editor, Ryuzo Sato and an anonymous referee for their insightful comments, which improved the paper. All remaining errors are our own. Kato thanks JSPS KAKENHI (13303009) for the financial support.

Addendum[18]

Ten Years After "The Winner–Loser Effect in Japanese Stock Returns": A Review and Recalculation

Around 2001, when we wrote this chapter, the short-term momentum effect was investigated intensively, while the discussion of the return reversal effect was temporarily settled in the U.S. At that time, our paper presented evidence that the short-term momentum effect did not exist in the Japanese stock market.

Discussion regarding the absence of the short-term momentum effect in the Japanese stock market was already underway even in the U.S. Daniel et al. (1998) proposed a model that simultaneously explains the return reversal effect and the short-term momentum effect based on a behavioral approach. They specifically referred to the absence of short term momentum in Japan. The main contribution of our paper is to confirm the absence of the short-term momentum effect using long-term time-series data in a way similar to the standard approach adopted by previous studies in the U.S.

We examine the possible explanations to the short term momentum which has been proposed by recent studies since our paper appeared. The key question is whether short term momentum is really absent in Japan using a longer period of data. A brief review and the results of recalculation follow below.

Discussion After Our Paper

There are two different approaches in the subsequent discussions of the long-term return reversal (De Bondt and Thaler 1985) and short-term momentum (Jegadeesh and Titman 1993).

The first is a rational approach. Fama and French (1996) who found that although the long-term reversal effect can be explained by their three-factor model, the short-term momentum effect cannot be explained (Grundy and Martin 2001; Korajczyk and Sadka 2004; Muga and Santamaria 2007).

The second is a behavioral approach, which is represented by the following three studies (Daniel et al. 1998; Barberis et al. 1998; Hong and Stein 1999).

Since then, researchers have presented several studies focusing on the discovery of overlooked factors and improvement of the model structure by adding a liquidity measure to Fama and French's three factors (Lee and Swaminathan 2000; Avramov et al. 2006; Sadka 2006), switching a model structure under the market conditions (Chordia and Shivakumar 2002; Cooper et al. 2004; Bhojraj and Swaminathan 2006), and using proxy variables to represent investor behavior (Grinblatt and Han 2005; Hvidkjaer 2006; Chui et al. 2010).

[18]This addendum has been newly written by Toshifumi Tokunaga for this book chapter.

A Revisit to the Momentum Effect in Japan

Though De Bondt and Thaler (1985) and Jegadeesh and Titman (1993) were exposed to many criticisms at their time of publication, they later reviewed the robustness of their results (De Bondt and Thaler 1987; Jegadeesh and Titman 2001).

The Fama–French alphas for momentum portfolio returns and market states in Japan from 1977 to 2005 are reported in Table 21.7, which is reconstructed from the results in Tokunaga (2008a, b, c, 2009). This table updates the data used in our paper 8 years ago. The results ("ALL") show that the significant short-term momentum effect still does not exist in the Japanese stock market.

However, once the momentum portfolio returns are classified according to market states, a very interesting result appears. When the stock prices for the past 36 months are trending upward, the risk-adjusted returns of the subsequent momentum portfolios are positive and statistically significant. These results are consistent with those found in the U.S. market (Cooper et al. 2004). On the other hand, when the stock prices for the past 36 months are trending downward, the risk-adjusted returns of the subsequent momentum portfolios are negative, but statistically insignificant. These results are consistent with those found in the U.S. market (Cooper et al. 2004), and are consistent with the viewpoint that momentum returns are not significant.

If the result as a whole is considered, why does the momentum effect appear in the U.S., but not in Japan? There is one interesting number. Although NUP/NALL, the ratio of the number of observations, in the U.S. (Table I in Cooper et al. (2004)) is approximately 85 %, it falls to approximately 65 % in Japan. This result might be a clue to solve the lack of momentum effect in Japan. Thus, trifles are possibly the cause. In order to make this idea less controversial, many empirical studies addressing the Japanese stock market would need to be published from now on.

Table 21.7 Fama–French alphas for momentum portfolio returns and market states in Japan (1977–2005)

	$K = 1$	6	12	36	60
ALL	−0.06	0.24	0.19	0.09	0.05
	(−0.18)	(1.07)	(1.10)	(0.94)	(0.65)
UP	0.63	0.71	0.55	0.32	0.20
	(1.46)	(2.25)	(2.24)	(2.38)	(1.81)
DOWN	−0.81	−0.25	−0.18	−0.17	−0.12
	(−1.78)	(−0.77)	(−0.73)	(−1.15)	(−1.00)

Winner and loser portfolios are formed based on 6-month returns (from $t-5$ to $t-1$) for all the listed firms on the Tokyo Stock Exchange First Section from 1977 to 2005 and held for K months (from $t+1$ to $t+K$) after skipping month t. The market states are called "UP" ("DOWN") when TOPIX returns over months from $t-36$ to $t-1$ are non-negative (negative). The momentum portfolio is a zero investment portfolio, which is long in the winner and short in the loser. Fama–French monthly alphas for momentum portfolio returns are presented in this table. t-statistics are reported in *parentheses*

References

Avramov D, Chordia T, Goyal A (2006) Liquidity and autocorrelations in individual stock returns. J Financ 61(5):2365–2394

Barberis N, Shleifer A, Vishny R (1998) A model of investor sentiment. J Financ Econ 49(3):307–343

Bhojraj S, Swaminathan B (2006) Macromomentum: returns predictability in international equity indices. J Bus 79(1):429–451

Bremer M, Hiraki T (1999) Volume and individual security returns on the Tokyo Stock Exchange. Pac Basin Financ J 7:351–370

Brennan M, Chorda T, Subrahmanyam A (1998) Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. J Financ Econ 49:345–373

Chan L, Hamao Y, Lakonishok J (1991) Fundamentals and stock returns in Japan. J Financ 46:1739–1789

Chan L, Jegadeesh N, Lakonishok J (1996) Momentum strategies. J Financ 51:1681–1713

Chordia T, Shivakumar L (2002) Momentum, business cycle, and time-varying expected returns. J Financ 57(2):985–1019

Chui ACW, Titman S, Wei KCJ (2010) Individualism and momentum around the world. J Financ 65(1):361–392

Conrad J, Hameed A, Niden C (1994) Volume and autocovariances in short-horizon individual security returns. J Financ 49:1305–1330

Cooper JM, Gutierrez RC Jr, Hameed A (2004) Market states and momentum. J Financ 59(3):1345–1365

Daniel K, Titman S (1997) Evidence on the characteristics of cross-sectional variation in stock returns. J Financ 52:1–33

Daniel K, Titman S (1999) Market efficiency in an irrational world. Financ Anal J 55:28–40

Daniel K, Grinblatt M, Titman S, Wermers R (1997) Measuring mutual fund performance with characteristic-based benchmarks. J Financ 52:1035–1058

Daniel K, Hirshleifer D, Subrahmanyam A (1998a) Investor psychology and security market under- and overreaction. J Financ 53(6):1839–1885

Daniel K, Hirshleifer D, Subrahmanyam A (1998b) Investor overconfidence, covariance risk, and predictors of securities returns. Unpublished manuscript

Daniel K, Titman S, Wei KCJ (2001) Explaining the cross-section of stock returns in Japan: factors or characteristics? J Financ 56:743–766

Davis J, Fama E, French K (2000) Characteristics, covariances, and average returns 1929–1997. J Financ 55:389–406

De Bondt WFM, Thaler R (1985) Does the stock market overreact? J Financ 40(3):793–808

De Bondt WFM, Thaler R (1987) Further evidence on investor overreaction and stock market seasonality. J Financ 42(3):557–581

Fama EF, French KR (1993) Common risk factors in the returns on stocks and bonds. J Financ Econ 33:3–56

Fama EF, French KR (1996) Multifactor explanations of asset pricing anomalies. J Financ 51(1):55–84

Fama E, French K (1998) Value versus growth: the international evidence. J Financ 53:1975–2000

Grinblatt M, Han B (2005) Prospect theory, mental accounting, and momentum. J Financ Econ 78(2):311–339

Grundy BD, Martin JS (2001) Understanding the nature of the risks and the source of the rewards to momentum investing. Rev Financ Stud 14(1):29–78

Hong H, Stein JC (1999) A unified theory of underreaction, momentum trading, and overreaction in asset markets. J Financ 54(6):2143–2184

Hvidkjaer S (2006) A trade-based analysis of momentum. Rev Financ Stud 19(2):457–491

Iihara Y, Kato H, Tokunaga T (2004) The winner-loser effect in Japanese stock returns. Jpn World Econ 16:471–485

Jegadeesh N, Titman S (1993) Returns to buying winners and selling losers: implications for stock market efficiency. J Financ 48(1):65–91
Jegadeesh N, Titman S (2001) Profitability of momentum strategies: an evaluation of alternative explanations. J Financ 56(2):699–720
Kato K (1990) Being a winner in the Tokyo Stock Market: the case for an anomaly fund. J Portf Manag 16:52–56
Kobayashi T (1997) Theoretical foundation of style management. 31–63 (in Japanese)
Korajczyk RA, Sadka R (2004) Are momentum profits robust to trading costs? J Financ 59(3):1039–1082
Lee CMC, Swaminathan B (2000) Price momentum and trading volume. J Financ 55(5):2017–2069
Liew J, Vassalou M (1999) Can book-to-market, size and momentum be risk factors that predict economic growth? Columbia University working paper
Moskowitz T, Grinblatt M (1999) Do industries explain momentum? J Financ 54:1249–1290
Muga L, Santamaria R (2007) "New economy" firms and momentum. J Behav Financ 8(2):109–120
Richards A (1997) Winner–loser reversals in national stock market indices: can they be explained? J Financ 52:2129–2144
Rouwenhorst KG (1998) International momentum strategies. J Financ 53:267–284

Chapter 22
Addition to the Nikkei 225 Index and Japanese Market Response: Temporary Demand Effect of Index Arbitrageurs

Katsuhiko Okada, Nobuyuki Isagawa, and Kenya Fujikawa

Abstract We examine the Japanese stock market response to additions to the Nikkei 225 Index from 1991 to 2002. Similar to the reactions in the U.S. markets, the stock prices of the added firms go up on the announcement date, continue to increase until the day before the effective change date, and then decrease on and just after the change date. The stock price increase in this run-up period is thus temporary, as it is canceled out by the decline that begins on the change date. We also find that the excess demand of index arbitrageurs for shares of newly added firms is the main source of the temporary stock price increase.

Keywords Nikkei 225 composite change • Index effect • Demand shock

JEL Classification Codes G14

1 Introduction

A number of empirical studies, such as those of Shleifer (1986), Harris and Gurel (1986), Lynch and Mendenhall (1997), and Wurgler and Zhuravskaya (2002), examine the stock price effect associated with a change in the composition of the S&P 500 Index. All of these studies report that additions to the S&P 500 index

---

The original article first appeared in the Pacific-Basin Finance Journal 14: 395–409, 2006. A newly written addendum has been added to this book chapter.

K. Okada (✉)
Institute of Business and Accounting, Kwansei Gakuin University, 1-155 Uegahara Ichibancho, Nishinomiya, Hyogo 662-8501, Japan
e-mail: katsuokada@kwansei.ac.jp

N. Isagawa • K. Fujikawa
Graduate School of Business Administration, Kobe University, 2-1 Rokkodai, Nada, Kobe, Hyogo 657-8501, Japan

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_22

increase the stock prices of the added firms. Some recent studies focus on non-U.S. stock indices. For example, Chakrabarti et al. (2005) study the price and volume effects of changes in the MSCI index, which is the most popular international stock index. Using changes in the MSCI Standard Country Indices for 29 countries, they find that stocks being added to the indices experience a sharp rise in price after the announcement date and a further rise during the period preceding the effective change date.

Despite the fact that the Japanese stock market has the largest market capitalization after that of the U.S., few studies focus on the changes in the composition of the Nikkei 225 Index, which is the most broadly quoted stock index in Japanese stock markets. Hanaeda and Serita (2003) examine the impact of the large composite change in the Nikkei 225 Index that occurred on April 24, 2000.[1] However, by restricting their examination to a single event, they may undermine the accuracy of their assessment of the price effect of the Nikkei 225 composite change. Furthermore, their focus on a single event raises concerns about the calendar clustering effect.

Using a relatively large sample of changes in the Nikkei 225 Index, we investigate the stock price behavior of firms around the time they are added to the index.[2] Our sample consists of 69 firms added to the Nikkei 225 Index from 1991 to 2002.

We find that the stock prices of firms to be added rise significantly in response to the announcement of their addition. They continue to increase during the run-up period (between the day after the announcement and the day before the change date), and then decrease on and just after the effective date of the change. Such stock price behavior in Japanese markets is very similar to that in the U.S. (Lynch and Mendenhall 1997) and international markets (Chakrabarti et al. 2005). While the positive announcement effect is permanent, the further increase in the run-up period is temporary, as it is offset by the decline following the change date.

There are two main lines of reasoning to account for the positive market response. One interpretation, which we refer to as the "information hypothesis," is that the stock price of the newly added firm increases because the news of the addition may convey sound and previously unavailable information about the firm's prospects. Consistent with the information hypothesis, Denis et al. (2003) find that firms experience unexpectedly significant improvements in realized earnings following their addition to the S&P 500 Index. Jain (1987) and Dhillon and Johnson (1991) also report empirical findings in support of the information hypothesis.

The other interpretation, which we refer to as the "demand shock hypothesis," sheds light on the demand aspect of the index additions. The demand shock hypothesis suggests that excess demand from index tracking investors (e.g., index arbitrageurs and index fund managers) is the major source of the stock price increase that occurs upon addition of the firm to the index. Harris and Gurel (1986) and

[1]The composite change on April 24, 2000, was a very large event in the sense that 30 firms were added to, and 30 others deleted from, the Nikkei 225 Index simultaneously.

[2]Although Liu (2000) examines the effects of changes in the Nikkei 500 Index on stock price and trading, this index is not particularly popular in Japanese markets.

Shleifer (1986) are the first to demonstrate such a demand shock effect. While Harris and Gurel (1986) identify a temporary demand shock on the stock price of the added firms (the price pressure hypothesis), Shleifer (1986) finds a permanent demand effect (the downward-sloping demand curve hypothesis).[3]

Lynch and Mendenhall (1997) and Wurgler and Zhuravskaya (2002) apply the demand shock hypothesis to index additions from a somewhat different angle. Lynch and Mendenhall (1997) investigate the stock price behavior surrounding the effective date of the change as well as the announcement date. This separation of the two event dates clarifies the stock price behavior during the run-up period. They find that the stock prices of the firms added to the S&P 500 Index increase during the run-up period and decrease on the change date. They indicate that such stock price behavior is consistent with the demand shock scenario due to heavy index fund trading around the time of the change date.

Wurgler and Zhuravskaya (2002) develop a theoretical model and examine whether both the demand shock and the arbitrage risk affect the announcement effects for firms added to the S&P 500 Index. In an efficient market, the excess demand due to the portfolio alignment by the index tracking investors would be cancelled out by the rational arbitrageurs' counter-trading. In the real world, however, there exists an arbitrage risk, and as a result, risk-averse arbitrageurs would be expected to trade less aggressively than would be necessary to offset the demand shock.[4] Consistent with the predictions of the demand shock hypothesis, Wurgler and Zhuravskaya (2002) find that the magnitude of the announcement day returns of the added firms has significant positive relationships with both demand shock and arbitrage risk.

One advantage of our study is that it investigates how the demand shock caused by the index tracking investors affects the stock price behavior of firms newly added to the Nikkei 225 Index. As suggested by Lynch and Mendenhall (1997), index tracking investors tend to purchase the added stocks as closely as possible to the change date in order to minimize the tracking error. Such a trading strategy by index tracking investors therefore affects the volume and the stock price behavior following the announcement of an addition to an index.

For each stock added to the Nikkei 225 Index, we measure the size of the index arbitrageurs' demand shock and the shock's arbitrage risk. Our focus is on the index arbitrageurs rather than the index funds, for two reasons. First, the cash-future arbitrage balance on the Nikkei 225 was relatively large during our sample period. For the index arbitrageurs, conducting cash-future arbitrage trading is less risky in the arithmetically averaged Nikkei 225 Index than in a capital-weighted index like the S&P 500. Second, we did not obtain any meaningful data regarding the amount of Nikkei 225-type index funds.

[3]Chen et al. (2004) find an asymmetric price response between the firms added to and those deleted from the S&P 500 Index. They suggest that investor awareness contributes to the asymmetric price effects.

[4]Shleifer (2000) stresses that, in the real world, rational arbitrage seems to be costly for several reasons. Empirically, Pontiff (1996) finds that market frictions constrain the rational traders' arbitrage process, such that closed-end fund prices can deviate from their fundamental values.

We find that the temporary abnormal returns in the run-up period are related to the demand shock in a significantly positive manner. Index arbitrageurs' rebalancing actively affects the stock price increase of the added firms during the run-up window, within which the information effect is less likely to be contained. We also find a positive but insignificant relationship between arbitrage risk and the run-up period stock price increase.

The remainder of this chapter is organized as follows, In Sect. 2, we discuss some of the details associated with the Nikkei 225 Index composite change and the demand effect of index arbitrage trading. In Sect. 3, we describe our sample and the methodology of our analysis. In Sect. 4, we present the results for the stock price behavior surrounding the composite change of the Nikkei 225 Index. In Sect. 5, we present the results associated with the effects of the demand shock due to the index arbitrageurs' rebalancing. In Sect. 6, we conclude this chapter.

2 Changes in the Nikkei 225 Index and the Demand Effect of Index Arbitrage Trading

In Japanese stock markets, there are two popular and widely used markets indices: the Nikkei 225 Index and TOPIX. While TOPIX is a value-weighted stock index comprising approximately 1,600 stocks, the Nikkei 225 Index is a simple price-weighted arithmetic average of a selection of 225 actively traded stocks on the first section of the Tokyo Stock Exchange (TSE). Historically, the *Nihon Keizai Shimbun* obtained a license from the Dow Jones Co., Ltd., to use the name, and the index was published as the Nikkei Sow Average. In 1985, the *Nihon Keizai Shimbun* changed the name of the index to the Nikkei Average Index, which is publicly known as the Nikkei 225 Index in the marketplace.

The *Nihon Keizai Shimbun* has periodically reviewed the composition of the Nikkei 225 Index so that the index effectively reflects the current structure of Japanese industry. For example, in our sample period from 1991 to 2002, there were 38 changes in the Nikkei 225 Index. In many cases, announcements of the review results were made in September but became effective at the beginning of October. There were, on average, approximately five business days between the announcement date and the actual change date. Changes could also be made at any other time, provided a stock was found to be ineligible for the Stock Average.

The criteria used for changing the index composition included a component stock's trading volume and its market capitalization in the preceding 3 years. A constituent stock would be deleted when its market liquidity, as measured by its trading volume, fell from the top half of the stocks on the first section of the TSE. Other reasons for deletion included mergers, bankruptcies, and delistings. To fill the vacancies left by the deleted stocks, candidates for addition must be selected. The criteria for addition included sector distribution (to keep the index well balanced among industries) and market liquidity (i.e., within an industry, selecting firms with higher market liquidity). All proposed changes were announced in Japanese newspapers.

Table 22.1 Cash position in Nikkei Index arbitrage and open interest in futures

	Panel A: cash position in Nikkei Index arbitrage			Panel B: futures open interest	
Year/month	Amount in millions of yen	Total number of shares (in 1000s of shares)	AB (in 1000s of shares)	Nikkei 225 (in millions of yen)	TOPIX (in millions of yen)
1991.9	1,138,675	1,014,648	3,608	3,648,901	569,723
1992.3	1,265,369	1,421,869	5,056	3,383,074	361,411
1992.9	741,335	741,334	2,636	2,248,312	379,065
1993.3	1,128,710	1,270,342	4,517	2,620,472	727,003
1993.9	1,205,743	1,254,016	4,459	2,217,639	983,198
1994.3	1,191,017	1,250,590	4,447	1,535,182	1,128,231
1994.9	737,596	771,877	2,744	2,156,518	987,669
1995.3	820,157	1,065,811	3,790	1,895,562	774,650
1995.9	1,882,532	2,181,199	7,755	2,719,919	1,393,275
1996.3	2,826,069	2,784,809	9,902	5,472,683	2,326,006
1996.9	2,796,620	2,757,016	9,803	4,265,861	1,771,132
1997.3	1,702,533	1,949,953	6,933	4,218,575	1,526,804
1997.9	1,369,777	1,547,996	5,504	2,991,487	1,486,377
1998.3	1,623,328	1,977,549	7,031	3,278,693	1,537,876
1998.9	538,367	752,001	2,674	2,303,057	1,363,779
1999.3	944,304	1,111,438	3,952	3,169,250	1,202,098
1999.9	801,750	827,508	2,942	2,699,586	1,499,462
2000.3	2,285,801	2,062,232	7,332	2,207,799	2,382,721
2000.9	2,929,952	2,186,433	7,774	1,986,202	1,854,624
2001.3	3,365,349	2,805,554	9,975	2,150,878	2,237,304
2001.9	1,507,301	1,599,583	5,687	1,374,125	1,787,659
2002.3	1,623,321	1,598,890	5,685	1,696,793	2,362,432
2002.9	938,608	980,718	3,487	1,341,782	1,785,741

Note: The amount in millions of yen (the second column) and the total number of shares (the third column) held in the cash position of index arbitrage are reported by the TSE. AB (the fourth column) is the estimated number of shares per firm in the Nikkei 225 arbitrage position, which is calculated by multiplying number of shares by 0.8 and dividing it by 225. Open interest in Nikkei 225 futures is reported by Osaka Securities Exchange, and that in TOPIX is reported by TSE. Since a large amount of Nikkei index-arbitrage was conducted against the SGX (SIMEX) Nikkei 225 futures contract, open interest in Nikkei 225 futures was smaller than the amount in the cash position in the latter part of the sample period

The futures markets on the Nikkei 225 Index have been the most actively traded of the sock index futures in Japan (Panel B of Table 22.1 and Japan Securities Research Institute (2002)). Since there is little official data as to what percentage of the arbitrage balance is in the Nikkei 225, we have traced newspaper articles in our sample period. We have found several articles with securities brokers' comments on the ratio between the index arbitrage balance in the Nikkei 225 and that in the TOPIX. The consensus estimate for the ratio of index arbitrage balance in the Nikkei 225 appears to be time-variant but generally ranges from 60 % to 80 %. Our recent interviews with several institutional arbitrageurs in the Japanese stock markets have suggested that roughly 80 % of the stock index arbitrage was carried out on the Nikkei 225 Index. This is one of the most prominent features distinguishing the Nikkei 225 Index from the S&P 500.

In particular, for the index arbitrageurs, conducting cash-future arbitrage trading in an arithmetically averaged index like the Nikkei 225 is less risky than in a capital-weighted index like the S&P 500. Since a capital-weighted index has different weightings for its component stocks, it is necessary for arbitrageurs to adjust their stock portfolios as the price fluctuates. On the other hand, the Nikkei 225 Index can be entirely hedged with the futures contract, regardless of the movement in the price of the component stocks. This safety feature for arbitrageurs promotes the cash-future arbitrage balance on the Nikkei 225 Index. Its balance usually exceeds one trillion yen (approximately ten billion US dollars), and sometimes exceeds two trillion yen.

The typical trading behavior of the index arbitrageurs facing a change in the composition of the Nikkei 225 Index is as follows. On the first trading day following an announcement, the stocks to be added to the index are likely to be bought up by numerous bidders. However, the index arbitrageurs do not rebalance their portfolios until one day prior to the change date (CD − 1), because it is of the utmost importance for them not to make tracking errors. Thus, the arbitrageurs avoid taking any risk between the announcement date and the change date. By selling stocks deleted from the index and purchasing stocks added to the index at the closing price on CD − 1, the index arbitrageurs can immunize their cash stock portfolios against tracking errors. In reality, the arbitrageurs start to rebalance their portfolios a few minutes before the closing bell on CD − 1.[5]

Managers of Nikkei 225-type index funds are also averse to ignoring the index, and thus could be expected to engage in behavior similar to that of the index arbitrageurs. Although these fund managers do not have to take any position in the index futures markets, they are evaluated based on how well they track the Nikkei 225 Index. Then, from the viewpoint of analyzing the demand effects on the stock prices of firms added to the Nikkei 225 Index, there is little difference in the rebalancing behavior between the index arbitrageurs and index fund managers.

Unfortunately, the index fund balance by category classification (Nikkei 225 type or TOPIX type) is not publicly disclosed, making it difficult to estimate the millions of Yen held in Nikkei 225-type index funds. Interviews with index fund managers of major asset management companies in Tokyo indicate that the majority of the index funds are of the TOPIX-type. Therefore this chapter does not analyze the demand effect due to index fund managers' behavior.

Table 22.1 provides information on the cash position of the stock index arbitrages and the open interest of index futures in Japanese markets during our sample period.

[5]There is one scenario in which index arbitrageurs would fail to track the index. Since the TSE has a matching rule for the closing, stocks may not close for the day if there is a seller-buyer imbalance. In such cases, index arbitrageurs are unable to trade relevant stocks, and therefore their portfolios are subject to subsequent price fluctuation. In order to avoid this risk, index arbitrageurs sometimes attempt to complete their rebalancing before the closing bell.

Since the composite changes in the Nikkei 225 Index occurred mostly in September and March, we use data collected at the end of September and the end of the March each year. Panel A of Table 22.1 shows both the amount in millions of Yen (in the second column) and the total number of shares (in the third column) in an all arbitrage positions based on data provided by the TSE. Assuming that 80 % of the index arbitrage is on the Nikkei 225 Index, we estimate the number of shares per firm held by the Nikkei 225 index arbitrageurs, AB (in the fourth column). Dividing this number by the total number of outstanding shares of the firm (OUT), we obtain a measure of the demand effect that is due to the rebalancing by the index arbitrageurs, DS. The measure of the demand shock, DS = AB/OUT, is the percentage of the outstanding shares of the added firm that are purchases by the index arbitrageurs when a change in the Nikkei 225 Index composition occurs.

We also use another measure of the demand shock, DSF, which takes into account the cross-held shares among Japanese firms. The measure DSF = AB/FF is the percentage of the added firm's free float (FF) to be purchased by index arbitrageurs. We collected data on the free float firms from the *Toyo Keizai Data Bank*. In our sample, the average percentage of the total number of outstanding shares that were free-floating shares was 36.2 %. On average, less than 40 % of outstanding shares were available for trading because of the cross-shareholdings. The firm with the lowest percentage of free-floating shares (11.9 %) in our sample is Mizuho Trust and Banking Co. (Yasuda Trust). This is a traditional trust and banking institution whose largest shareholder is Mizuho Bank. Conversely, the firm with the highest percentage of free-float shares (83.8 %) is Trend Micro, which has a relatively short history; it has grown rapidly along with the growth of the internet.

Table 22.2 shows the summery statistics for the two measures of demand shock. In our sample firms, DSF was 5.50 % on average, which is much larger than average DS of 23.8 %. Table 22.2 also presents the statistics on the average percentage trading volume for sample firms over pre-announcement periods. On average, 0.25 % of the firm's outstanding shares were traded per day over the 100-day period prior to the announcement (Average V/OUT (AD – 100, AD – 1)) in Table 22.2.[6] Over the 30-day period prior to the announcement, 0.29 % of the firm's outstanding shares were traded per day (Average V/OUT (AD – 30, AD – 1)). Comparing DS and DFS with these values reveals the impact of the index arbitrageurs' trading on the stock behavior. Even if we compare DS or DSF with the average percentage of the firm's trading volume to free float over the pre-announcement periods (Average V/FF (AD – 100, AD – 1) and Average V/FF (AD – 30, AD – 1) in Table 22.2, it seems reasonable to suppose that index arbitrageurs' rebalancing activity affects the stock price of added firms surrounding the change date.

[6]Throughout this chapter, day AD $-$ t (day AD $+$ t) denotes t days before (after) the announcement date (AD), and day CD $-$ t (day CD $+$ t) denotes t days before (after) the effective change of the addition (CD).

Table 22.2 Comparison of demand shock to trading volume in the pre-announcement periods

	Sample mean (n = 69)	Standard deviation	Max	Min
DS = AB/OUT	2.38 %	0.0190	10.5 %	0.08 %
DSF = AB/FF	5.50 %	0.0494	23.0 %	0.50 %
Average V/OUT (AD − 100, AD − 1)	0.25 %	0.00045	0.43 %	0.04 %
Average V/OUT (AD − 30, AD − 1)	0.29 %	0.00031	0.85 %	0.03 %
Average V/FF (AD − 100, AD − 1)	0.71 %	0.01522	12.6 %	0.10 %
Average V/FF (AD − 30, AD − 1)	0.82 %	0.02461	30.0 %	0.09 %

Note: DS = AB/OUT and DFS = AB/FF are measures of demand shock due to the rebalancing by index arbitrageurs as explained in the text and in the note in Table 22.1. Average V/OUT (AD − 100, AD − 1) is the average percentage of the added firm's trading volume to outstanding shares over the 100-day period of (AD − 100, AD − 1). Average V/OUT (AD − 30, AD − 1) is the same value over the 30-day period of (AD − 30, AD − 1). Average V/FF (AD − 100, AD − 1) is the average percentage of the firm's trading volume to free float over the 100-day period of (AD − 100, AD − 1). Average V/FF (AD − 30, AD − 1) is the same value over the 30-day period of (AD − 30, AD − 1)

3 Data and Methodology

Our original sample consists of 82 firms newly added to the Nikkei 225 Index between September 1991 and October 2002. We obtain information regarding the changes in index composition from *Nihon Keizai Shimbun*, which is the most popular financial newspaper in Japan. Since the *Nihon Keizai Shimbun* has a policy of announcing composite changes after the stock market is closed, we define the announcement date as the first business day after the release of such information.

From the 82 original samples, we exclude some firms for two reasons. First, we exclude ten firms due to a lack if the historical returns that are necessary two estimate the market model coefficients and the arbitrage risk measures. Second, because our study relies on a separation in time between the announcement date and the effective date of the change, we exclude three firms that were added immediately after it was announced that they would be added.

Among the remaining 69 firms, 30 were added to the index on the same day. April 24, 2000. In order to overcome bias due to the clustering effect, we use two methods. One is a regression model with dummy variables for the event period, and the other is the equally weighted portfolio approach.[7] In the dummy variable approach, we added a dummy variable equal to one for firms added to the index on April 24, 2000, and zero otherwise. The results are similar under both methods (see Table 22.4). Since the reduction of 30 firms to one portfolio will cause a loss of information, for the most part in this chapter we present the results from the regression model with dummy variables.

[7] For the clustering effects, see Peterson (1989) and Henderson (1990).

We calculate the daily stock returns by using the closing price taken from the *Nomura Aurora Database*, which is widely used in Japan. The market model residual is considered to be a firm's abnormal return on any given day.[8] This residual is calculated as the difference between the actual return and the predicted return based upon the market model parameter estimates and the market return for that day. The parameters of the market model are estimated over a 200-day period between AD − 230 and AD − 31, and TOPIX is used as the market index.

A cumulative abnormal return over the window from day t_1 to day t_2 is calculated by cumulating the daily abnormal returns over that window geometrically. The average abnormal return on the day t is denoted by AAR (t), and the average cumulative abnormal return over the window of (t_1, t_2) is denoted by CAR(t_1, t_2).

The volume data are also obtained from the *Nomura Aurora Database* for each firm added to the Nikkei 225 Index. In order to determine whether or not the trading activity increases in the added firms on any given day in the event window, we calculate the mean volume ratio proposed by Harris and Gurel (1986). The volume ratio for firm *i* on day *t*, is calculated by

$$VR_{it} = \frac{V_{it}}{V_{mt}} \cdot \frac{V_m}{V_i}. \tag{22.1}$$

Whether V_{it} is the trading volume for firm *i* on day *t*, V_{mt} is the total trading volume for the firms included in the market portfolio of TOPIX on day *t*, and V_{i} and V_{m} are the average trading volumes over the 200-day estimation period of (AD − 230, AD − 31) for firm *i* and for the total TOPIX index, respectively. The sample average of the volume ration on the day *t* is denoted as $\mathrm{MVR_t}$. If an addition to the Nikkei 225 Index has no effect on the trading volume, then MVR should not deviate widely from the expected value surrounding the announcement date and change date.

4 Stock Price Behavior Surrounding Addition to the Nikkei 225 Index

Panel A of Table 22.3 presents the daily average abnormal return (AAR) and the mean volume ration (MVR) relative to the announcement date, and panel B of Table 22.3 presents those values relative to the effective change date. For each of the five event days after the announcement date (from AD + 1 to AD + 5), only those firms for which the change date has not yet occurred were used to calculate AAR and MVR. Similarly, for each of the five event days prior to the change date (from CD − 5 to CD − 1), only those firms for which the announcement date occurred on a prior day were used to calculate AAR to MVR.

[8] We also attempted to calculate the abnormal returns of firms using the market-adjusted return model, and observed no significant difference in the results.

Table 22.3 Daily AAR and MVR surrounding addition to the Nikkei 225 Index

Days relative to AD	Number of samples	AAR (%)	t-value	MVR
Panel A: AAR and MVR relative to the announcement day (AD)				
AD − 10	69	0.03	0.01	1.06
AD − 9	69	0.60	1.57	1.04
AD − 8	69	−0.09	−0.23	1.06
AD − 7	69	0.15	0.38	1.04
AD − 6	69	−0.32	−0.83	1.09
AD − 5	69	0.74	1.92*	1.03
AD − 4	69	−0.09	0.24	1.06
AD − 3	69	1.06	2.75***	1.26
AD − 2	69	0.14	0.35	1.21
AD − 1	69	0.59	1.53	1.38
AD	69	5.16	13.42***	4.45
AD + 1	68	0.55	1.42	4.21
AD + 2	65	2.59	6.74***	4.88
AD + 3	51	1.35	3.51***	3.48
AD + 4	50	1.01	2.63***	3.44
AD + 5	16	0.26	0.68	2.37
Panel B: AAR and MVR relative to the change day (CD)				
CD − 5	16	0.31	0.79	2.18
CD − 4	50	0.83	2.16**	2.19
CD − 3	51	0.31	0.81	3.25
CD − 2	65	1.67	4.35***	4.40
CD − 1	68	4.19	10.92***	8.21
CD	69	−3.50	−9.11***	4.23
CD + 1	69	−1.37	−3.57***	2.30
CD + 2	69	−0.28	−0.73	2.20
CD + 3	69	0.25	0.65	3.40
CD + 4	69	−0.53	−1.38	3.72
CD + 5	69	−0.76	−1.98*	2.51
CD + 6	69	0.21	0.56	1.83
CD + 7	69	−0.53	−1.39	1.60
CD + 8	69	−0.03	−0.08	1.77
CD + 9	69	0.47	1.21	1.81
CD + 10	69	−0.38	−0.99	1.46

Note: AAR is the average abnormal return and MVR is the mean volume ratio for the specified day in the evening period. t-value is the statistic used to test the null hypothesis that AAR = 0
*Significant at the 10 % level, **Significant at the 5 % level, ***Significant at the 1 % level

The announcement of the addition of a firm to the Nikkei 225 Index pushes up the firm's stock price by a substantial margin. An abnormal return of 5.16 % with a t-value of 13.42 is observed on the announcement day. The t-value used to test the significance of each AAR is based on the standard deviation of AAR over the 200-

day estimation period. As shown in Table 22.3, the average stock price of added firms continues to increase gradually until CD − 1. The added firms experience a significant positive abnormal return of 4.19 %, with a t-value of 10.92 on CD − 1.

The index addition is announced in a well-publicized manner through various media. If the market responds to this news immediately, it is unlikely that any significant stock price reaction will be observed after the announcement date. However, the average stock price of the added firms is found to continue rising after the announcement date, and the magnitude of the increase is conspicuously large on CD − 1. One possible explanation for the result would be that the index arbitrageurs tend to purchase added stocks on CD − 1 in order to minimize the risk of making a tracking error, as discussed in Sect. 2.

Suppose that such heavy trading by the index arbitrageurs prior to the effective change date shifts the stock prices temporarily away from their fair values. We then would expect to observe a price reversal following the change date as the index arbitrageurs' trading disappears. Consistent with this expectation, Table 22.3 shows that the prices of the added stocks decline significantly on the change date (CD) and on the next day (CD − 1) as well. The average abnormal return on CD is −3.50 % with a t-value of −9.11, and that on CD + 1 with a t-value of −3.57.

We also expect that the trading volumes of the added firms would increase on CD − 1 for the same reason. The daily average trading volume ratio (MVR) over the event window is presented in the last column of Table 22.3. As shown in the Table, MVR increase abruptly on the announcement day, and continues to be greater than 1 after announcement date. The largest MVR is observed on CD − 1, which is consistent with the hypothesis that index arbitrageurs rebalance their portfolios on that day. After the change date, the MVR of the added firms is found to decrease gradually.

Next, we examine the average cumulative abnormal returns (CAR) over the event window. Like Lynch and Mendenhall (1997) and Chakrabarti et al. (2005), we examine CAR over five event windows, i.e., the run-up window of (AD + 1, CD − 1), the release window of (CD, CD + 7), the post-announcement windows of (AD + 1, CD + 7) and (AD + 1, CD + 10), and the total permanent window of (AD, CD + 10). We determine CD + 7 to be the release-ending day based on the same criterion used by Lynch and Mendenhall (1997).

Panel A of Table 22.4 shows the results of AAR on the announcement date and CAR over five event windows for our 69 samples (OLS with dummy variables). Panel B of Table 22.4 shows those results for 40 samples (portfolio approach). We report the results of the 40-sample case in order to show that there is little difference between the two methods in terms of quantitative results. In the following, we primarily report the results of the 69-sample case.

If all of the added firms have windows of identical length, then the t-value used to test the significance of CAR(t_1, t_2) is given by

$$t = CAR\left(t_1, t_2\right) \ / \sigma \sqrt{(t_2 - t_1)}. \tag{22.2}$$

Table 22.4 Cumulative abnormal returns over four event windows

Event window	Event days	Average days	CAR (%)	CAR (%)
Panel A: 69-sample case				
Announcement day	AD	1	5.16	13.42***
Run-up window	(AD + 1, CD − 1)	4.52	5.70	6.99***
Release window	(CD, CD + 7)	8	−6.51	−5.99***
Post-AD window 1	(AD + 1, CD + 7)	12.52	−0.81	−0.595
Post-AD window 2	(AD + 1, CD + 10)	15.52	−0.75	−0.50
Total permanent	(AD, CD + 10)	16.52	4.40	2.82***
Panel B: 40-sample case (39 individual firms and one portfolio)				
Announcement day	AD	1	4.94	12.41***
Run-up window	(AD + 1, CD − 1)	4.9	5.91	6.70***
Release window	(CD, CD + 7)	8	−5.77	−5.13***
Post-AD window 1	(AD + 1, CD + 7)	12.9	0.13	0.09
Post-AD window 2	(AD + 1, CD + 10)	15.9	0.39	0.25
Total permanent	(AD, CD + 10)	16.9	5.33	3.26***

Note: This table represents the results of average abnormal returns (AAR) on the announcement date and cumulative abnormal returns (CAR) for five different event windows for two samples
***Significant at the 1 % level

However, in the run-up window, the window's length varies across firms. We therefore use the average value of $(t_2 - t_1)$ of all 69 samples in Eq. (22.2)

On average, the added firms experience a significant cumulative abnormal return of 5.70 % with a *t*-value of 6.99 over the run-up window of (AD + 1, CD − 1). Note that most of this CAR is the abnormal return on CD − 1 of 4.19 %. In the release window of (CD, CD + 7), the average cumulative abnormal return of the added firms is −6.51 %, which is significant at a 1 % level. The stock price rise in the run-up window is erased by the decline during the release period. Naturally, in the post-announcement windows of (AD + 1, CD + 7) and (AD + 1, CD + 10), the average cumulative abnormal return is almost zero.

Figure 22.1 plots the CAR for the added firms surrounding the announcement date and the effective change date. In the figure, the interval between AD + 1 and CD − 1 is shown as 5 days, because the actual average number of days in the interval is 5.5, inclusive of AD + 1 and CD − 1. The CAR is displayed as if each daily AAR over this interval is the interval's CAR divided by five.

Figure 22.1 shows that the average stock price of the sample firms increases in response to the announcement that the firms will be added to the Nikkei 225 Index. The average stock price continues to rise over the run-up window, then experiences a substantial rise on the day before the change date (CD − 1) and a large drop on the change date (CD). The average stock price then begins to revert to the level observed on the announcement date. As shown in Panel A of Table 22.4, CAR during the post-AD window (AD + 1, CD + 7) is −0.81 %, which is not significantly different

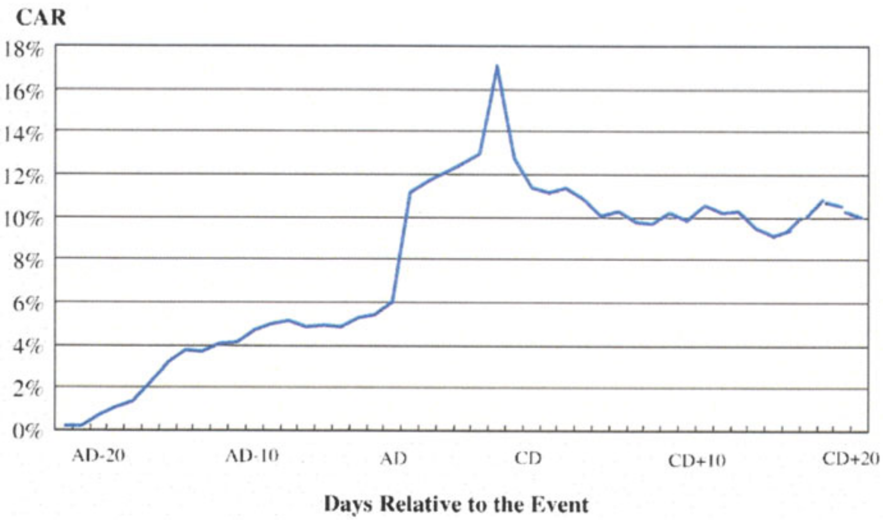


Fig. 22.1 CAR for firms newly added to the Nikkei 225 Index. Note: AD is the announcement date and CD is the effective change date. Since the number of trading days between AD and CD varies across firms, the interval between AD + 1 and CD is displayed as 5 days (actual average = 5.5 days). The CAR is displayed each daily AAR over this interval were the interval CAR divided by five. Note that the sample firms experience about a 6 % stock price increase on average during the pre-announcement window of (AD − 20, AD − 1). As shown in Table 22.3 and Panel A of Table 22.4, AAR of the sample firms on the announcement date (AD) is 5.16 %, CAR during the run-up period of (AD + 1, CD − 1) is 5.7 %, and CAR during the release window of (CD, CD + 7) is −6.51 %. Note also that CAR during the post-AD window (AD + 1, CD + 10) is −0.81 %(5.7–6.51), which means that the average stock price of the sample firms surrounding CD + 7 is almost equal to that on AD + 1

from zero. This means that the average stock price of the sample firms surrounding CD + 7 is almost equal to the level on AD + 1.[9]

This pattern of stock price behavior for the added firms is very similar to that reported by Lynch and Mendenhall (1997). Although the results are not reported in this chapter, we have confirmed that the stock price behavior of firms deleted from the Nikkei 225 Index is also similar to that of counterpart firms deleted from the

[9] As shown in Fig. 22.1, there is no significant stock price change after CD + 7. On the other hand, the added firms experience a gradual stock price increase (about 6 %) prior to the announcement date. This may be because some analysts report which firms are candidates for the Nikkei 225 Index based on the criteria for changing the index composition, and several investors take a long position based on the analysts' reports.

S&P 500. Thus, little difference is observed between the Japanese and U.S. markets in terms of the stock price behavior associated with composite changes in a popular stock index.[10]

The addition of a firm to the Nikkei 225 Index produces a significant stock price increase on the announcement date, and this rise is permanent rather than temporary. There are two hypotheses that might explain this permanent increase: the information hypothesis and the downward-sloping demand curve hypothesis. However, it is difficult to determine which one is more plausible. As mentioned by Chakrabarti et al. (2005), they are not mutually exclusive.

The addition of a stock to the Nikkei 225 Index also produces a continuous stock price increase into the run-up window. In contrast to the announcement effect, the run-up period stock price increase in neutralized by the stock price decline in the following release window and is therefore temporary. In the next section, we examine the effects of the demand stock caused by the portfolio rebalancing on the part of the index arbitrageurs.

5 Effects of Index Arbitrageurs' Demand Shock

Previous studies, such as those of Harris and Gurel (1986), Lynch and Mendenhall (1997), and Wurgler and Zhuravskaya (2002), suggest that the demand shock due to trading by index tracking investors affects the stock price behavior of firms newly added to the index. As we argue in Sect. 2, it is likely that, in the case of the Nikkei 225 Index, the rebalancing of the cash stock portfolio by index arbitrageurs affects the stock prices of the added firms. In this section, we investigate this issue in detail.

In order to test how the demand shock affects the stock price of an added firm, the cumulative abnormal return in the window of (AD, CD − 1) is regressed on the demand shock and arbitrage risk measures. As explained in Sect. 2, we use two measures of demand shock. One measure, DS, is the percentage of the added firm's shares relative to its outstanding shares to be purchased by index arbitrageurs. The other measure, DSF is the percentage of the added firm's free float to be purchased by index arbitrageurs.[11] In general, DSF is larger than DS. Under the demand shock hypothesis, we expect that DSF would have a stronger influence on the temporary stock price than DF.

If an excess demand shock hits the market, it may shift the stock price away from its fair value, since rational arbitrage activity does not function in a perfect way as predicted by the efficient market hypothesis. In an efficient market setting, rational arbitrages should have an effect that is opposite that of the unsophisticated demand shock and this should return the stock price to its fair value. However, as argued

[10]Chakrabarti et al. (2005) report essentially the same pattern of stock price behavior in Japanese markets surrounding changes in the MSCI.

[11]We alternately used 70 % in place of 80 % and observed no significant difference in the results.

by Shleifer (2000) and Wurgler and Zhuravskaya (2002), such arbitrage activity entails risk on the part of the arbitrageurs and falls short of the necessary corrective force. The larger the arbitrage risk of the added stock, the more it is overvalued by the excess demand shock. Therefore, we expect the cumulative abnormal return following the announcement date to have a positive relationship with the extent of the stock's arbitrage risk.

As a measure of arbitrage risk, which is designated as RISK, we use the variance of the residuals of the market model over a 200-day estimation period. The same arbitrage risk measure also employed by Wurgler and Zhuravskaya (2002), and Hanaeda and Serita (2003), and Chakrabarti et al. (2005).[12]

Panel A of Table 22.5 presents the regression results for the stock returns in the window of (AD, CD − 1). The results show that, during the window, the abnormal stock return of the added firms has a significant positive relationship with DF and DFS. The results of model (1) and model (2) indicate that DSF has a more significant effect than DS. This result is line with our initial prediction.

The finding that constant terms in models (1) and (2) are significantly may be attributed the information effect, which is likely to be reflected in the stock price on the announcement date. In order to remove such an effect, the cumulative abnormal return in the run-up window of (AD + 1, CD − 1) is regressed on the demand shock and arbitrage risk measures. The results of models (3) and (4) in Panel B of Table 22.5 show that only demand shock has a significant positive effect during the run-up period on the abnormal stock returns of added firms. No constant term has a significant relationship with CAR in the run-up period. Consistent with our prediction, DSF has a more significant effect than DS. It should be noted that the regression results provide plausible explanations with large R-squares. Models (3) and (4) can account for about 25 % and 34 % of the total variation in abnormal stock price increases in the run-up window, respectively. It can therefore be concluded that the index arbitrageurs' excess demand is an important factor in the temporary stock price increase that follows the announcement of additions to the Nikkei 225 Index.

Models (3) and (4) also show that abnormal stock price increases are related in a positive manner with the measure of the stock's arbitrage risk (RISK). Although the results are consistent with the qualitative prediction, they are not statistically significant. Hanaeda and Serita (2003) report that for additions to the Nikkei 225 Index, the measure of the arbitrage risk based on the standard deviation of the error term of the market model has no significant impact on the announcement effect. Such a measure of arbitrage risk may not be effective for explaining the temporary stock price changes that occur in response to changes in the Nikkei Index.[13]

[12]Wurgler and Zhuravskaya (2002) use another measure of arbitrage risk based on firms that match the characteristics of the added firm. They find that the two measures of arbitrage risk are very similar in magnitude and are highly correlated.

[13]In their sample of all firms newly added to the MSCI Index, Chakrabarti et al. (2005) find no significant relationship between stock behavior and the measure of arbitrage risk based on the residuals of the market model.

Table 22.5 Cross-sectional regressions of CARs

Panel A: Regression results of CAR in the window of (AD, CD − 1)		
	(1)	(2)
Constant	0.059(3.21)***	0.041(2.28)**
DS	2.50(3.86)***	
DSF		1.20(5.17)***
RISK	0.66(0.23)	0.77(0.28)
R^2	0.19	0.29
Sample size	69	69
Panel B: Regression results of CAR in the run-up window of (AD + 1, CD − 1)		
	(3)	(4)
Constant	−0.004(−0.12)	−0.025(−0.71)
DS	2.45(4.69)***	
DSF		1.10(5.81)***
RISK	0.50(0.42)	0.54(0.53)
R^2	0.25	0.34
Sample size	69	69
Panel C: Regression results of CAR in the release window of (CD, CD + 7)		
	(5)	(6)
Constant	−0.005(−1.20)	−0.05(−1.39)
DS	0.57(1.00)	
DSF		0.32(1.39)
RISK	−0.97(−0.87)	−0.87(−0.78)
RUN-UP	−0.29(−2.55)***	−0.33(−2.75)***
R^2	0.19	0.19
Sample size	69	69

Note: The dependent variables are CAR in the window of (AD, CD − 1) in Panel A, CAR in the run-up window of (AD + 1, CD − 1) in Panel B, and CAR in the release window of (CD, CD + 7) in Panel C. DS is the measure of demand shock based on the total outstanding shares, DSF is the measure of demand shock based on the free float, and RISK is the measure of arbitrage risk. RUN-UP represents CAR in the run-up period. The *t*-value is given in parentheses
**Significant at the 5 % level, ***Significant at the 1 % level

Next, we focus on the stock price decrease in the release window of (CD, CD + 7). Panel C of Table 22.5 presents the regressions results for the stock returns in the release window. Since the index arbitrageurs retain most of the stocks that they purchased before the change date as part of their cash portfolios, the index arbitrageurs' huge demand for the added firms' shares will likely disappear after the change date. Then, we expect that neither DS nor DSF will have a significant effect on the release-period CAR. The results are consistent with this prediction, because the coefficient of the demand shock in models (5) and (6) is not significant.[14]

[14]We find a significant positive relationship between the release-period CAR and the estimated percentage of the added firms shares to be sold or purchase by the index arbitrageurs during the

The arbitrage risk has a negative but insignificant relationship with the stock price behavior after the change date. The coefficients of RISK in models (5) and (6) show that the higher the arbitrage risk of the added stock, the more it is overvalued in the run-up period. Then, in the release period, the overvaluation due to arbitrage risk may disappear gradually.

In the regression model, we include the run-up period cumulative abnormal returns, denoted by RUN-UP. The results in models (5) and (6) show that the stock price behavior in the release period has a significant relationship with that in the run-up period. The more the stock price increases in the run-up period, the more it goes down on and immediately after the change date. We also perform the same regression analysis for cumulative abnormal returns in the 2-day period of (CD, CD + 1), and we obtain essentially the same results as those obtained for the release window.

6 Conclusion

This chapter investigates the stock price behavior of firms around the time they are added to the Nikkei 225 Index, the most popular stock index in the Japanese stock markets. The stock prices of firms to be added rise on the announcement date. They continue to rise during the run-up window (until the day before the effective change date) and subsequently decline beginning on the change date. Such stock price behavior in Japanese markets is very similar to that in the U.S., as reported by Lynch and Mendenhall (1997). With regard to the addition to the major stock index, which is not under the firm's discretion, there is little difference in the market responses of the Japanese and U.S. markets.

This chapter also investigates how the demand shock index arbitrageurs affect the stock price behavior surrounding the addition of a firm to the index. The cumulative abnormal return in the run-up window is positively related to the demand shock. This finding is consistent with the hypothesis that the excess demand of index arbitrageurs for newly added firms pushes up stock prices.

In sum, the stock prices of firms newly added to the Nikkei 225 Index rise significantly when the impending addition is first announced. This announcement effect is likely to be permanent. However, the post-announcement stock price rise in the run-up window turns out to be temporary, because most of the cumulative gains in the window are offset by the stock price declines that follow on and after the effective change day. The demand shock due to the index arbitrageurs' portfolio rebalancing significantly influences this temporary stock price increase.

release period. The decrease in the added firm's stock price following the effective change date is affected by index arbitrageurs' rebalancing of their cash positions. We are indebted to the reviewer for suggesting this point.

The magnitude of stock price reversion on and after the change day is significantly related to the extent of abnormal return during the run-up period.

Addendum: Trading Simulation[15]

Summary of Other Related Literature

The long-held assumption that stocks have perfect substitutes, as well as the perfect elasticity of demand that follows from it, is central to modern financial theory. If securities have perfect elasticity of demand, supply and demand shocks that are devoid of information should have no effect on prevailing prices. If stocks have a short-term downward demand curve, prices should be instantly affected by a demand shock due to indexing, but that effect should dissipate once the excess demand is satisfied. Early work by Harris and Gurel (1986) is consistent with the price pressure hypothesis. If, on the other hand, stocks have a long-term downward sloping demand curve, the excess returns should be permanent. Shleifer (1986), Beneish and Whaley (1996), Lynch and Mendenhall (1997), Kaul et al. (2000), Wurgler and Zhuravskaya (2002), Blume and Edelen (2004) present evidence supporting this view. The observed excess returns, however, may be the result of fundamental change that the index adoption/deletion brings about, namely, an increase in expected future cash flow and a decrease in required return.

An increase in expected future cash flow can occur for three reasons. First, inclusion in the index conveys positive information about that firm's prospects (Jain 1987; Dhillon and Johnson 1991). Second, enhanced investor awareness can lead to the expectation of cash flows, in that index-adopted firms may be forced to perform more efficiently and to make more value-enhancing decisions (Denis et al. 2003). Membership in the Nikkei 225 may also enable firms to attract new capital if financial institutions are more willing to lend to firms in the index.

A decrease in required return can also occur for three reasons. First, there may be an improvement in liquidity (Chordia 2001; Hegde and McDermott 2003). Second, the greater interest in Nikkei 225 index firms generates greater production of information by analysts and news media, which leads to reduced information asymmetry. A third explanation for a lower required return derives from heightened awareness in Merton's (1987) model. If some investors know of only a subset of all stocks and hold only the stocks they are aware of, these investors will be inadequately diversified and demand a premium for thenonsystematic risk they bear.

[15]This addendum has been newly written for the book chapter.

Merton calls this a shadow cost. If a firm is adopted as one of the composite stocks in the Nikkei 225 index, it will appeal to more investors and its shadow cost to diminish.

Market Efficiency Test

This chapter investigates how demand shock driven by index arbitrageurs affects stock price behavior. The cumulative abnormal return in the run-up window is positively related to the demand shock. This finding is consistent with the price pressure hypothesis. The stock prices of firms newly added to the Nikkei 225 Index rise significantly when the impending addition is first announced. Since this announcement effect is likely to be permanent, it also seems consistent with the downward-sloping demand curve hypothesis. However, the post-announcement stock price rise in the run-up window turns out to be temporary, because most of the cumulative gains in the window are offset by stock price declines that follow on and after the effective change day. These results are consistent with the previously documented hypotheses but inconsistent with market efficiency. Thus, it seems unlikely that reported returns around Nikkei 225 index composition changes can be attributed solely to liquidity and the information hypothesis.

To test market efficiency, weconduct a few trading strategy simulations, taking advantage of index composite changes. The trading strategies tested are the following: Strategy 1, going long \$1 in the added stock upon its announcement and holding this position until the day before the change date (CD – 1), and Strategy 2, going short \$1 in the added stock at the close of the change date (CD). Strategy 1 tests whether there is a window of opportunity to act early and take a position in the stock for a few days to sell it back to latecomers. Strategy 2 tests whether the stock price is overvalued on the change date (CD). Strategy 2 has three versions: 2-a, buying back the short position created on CD on CD + 1; 2-b, buying back on CD + 2; and 2-c, buying back on CD + 3. Addendum Table 22.6 describes the results of each strategy.

Table 22.6 Trading strategies surrounding index composite change

	Strategy 1	Strategy 2-a	Strategy 2-b	Strategy 2-c
Sample stocks (n)	69	69	69	69
Average return per stock	12.29 %	3.34 %	4.22 %	4.52 %
Max	46.97 %	12.69 %	15.99 %	18.78 %
Min	−10.87 %	−6.09 %	−3.70 %	−11.19 %
Winning ratio	91.30 %	86.96 %	82.61 %	84.06 %
Sample events (N)	24	24	24	24
Average return per event	6.17 %	2.59 %	3.50 %	3.56 %
Max	29.04 %	8.05 %	10.90 %	12.38 %
Min	−10.87 %	−3.45 %	−0.80 %	−11.19 %
Winning ratio	87.50 %	82.61 %	86.96 %	91.30 %

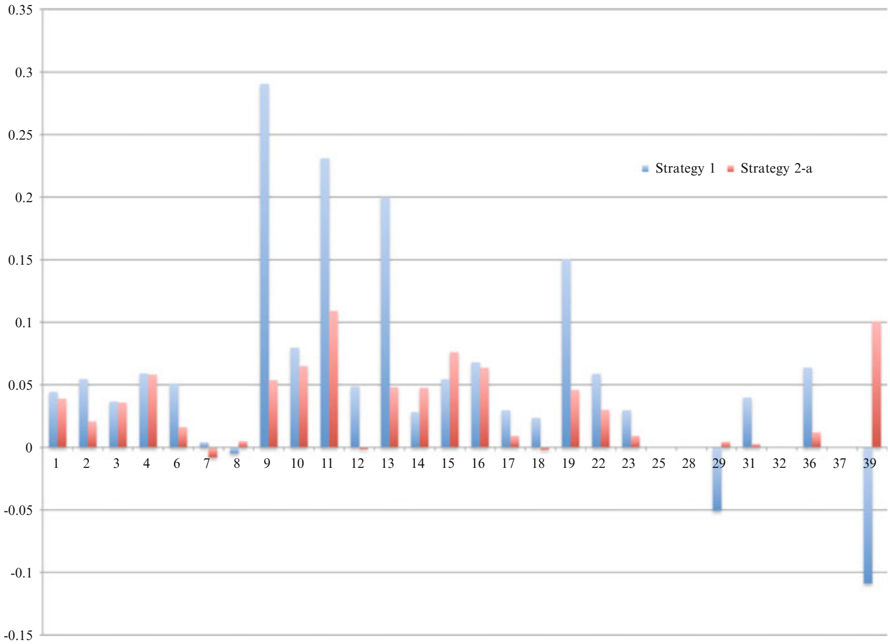


Fig. 22.2 Historical plot of average trading and loss per stock (*Note*: The numbers are calculated based on the return on $1 generated per event following Strategies 1 and 2-a, as described in the text)

All of the strategies produce significant positive returns. In the run-up window (from AD + 1 to CD − 1), investors would make 12.29 % on average, for $1 invested in a stock to be included in the Nikkei 225 Index. This is significantly more than reported by Lynch and Mendenhall (1997). These authors conduct a similar trading strategy with stocks to be included in the S&P 500 index and find such trades generate 3.968 % per stock, on average. The percentage of positive return trades over all trades (winning ratio) is a staggering 91.30 % for the Nikkei 225 Index, which is also substantially higher than the 79 % for the S&P 500 Index. Addendum Fig. 22.2 is a graphical representation of the history of trading profit and loss per stock, on average, for our 39 sample events. Strategy 1 is profitable most of the time but trading on the latest event in our sample generates a large loss. This may be because market participants became aware of the repeated pattern of the profit opportunity and had taken speculative long positions in the stocks before the announcement.[16] On the other hand, Strategy 2-a seems to remain profitable throughout, suggesting the possibility that market participants tend to overvalue newly added stocks during the run-up period.

[16]Brokerage houses issue reports on the new composite candidate stocks.

References

Beneish MD, Whaley RE (1996) An anatomy of the "S&P 500 Game": the effects of changing the rules. J Financ 51:1909–1930

Blume M, Edelen R (2004) S&P500 indexers; tracking error, and liquidity. J Portf Manag 30(3): 37–46

Chakrabarti R, Huang W, Jayaramen N, Lee J (2005) Price and volume effects of changes in MSCI indices- nature and causes. J Bank Financ 29:1237–1264

Chen H, Noronha G, Singal V (2004) The price response to S&P 500 index additions and deletions: evidence of asymmetry and a new explanation. J Financ 59:1901–1928

Chordia T (2001) Liquidity and returns: the impact of inclusion into the S&P 500 index. Working paper, Emory University

Denis D, McCollell J, Ovtchinikov A, Yu Y (2003) S&P 500 index additions and earnings expectations. J Financ 58:1821–1841

Dhillon U, Johnson H (1991) Changes in the standard and poor's 500 list. J Bus 64:75–85

Hanaeda H, Serita T (2003) Price and volume effects associated with a change in the Nikkei 225 list: new evidence from the big change on April 2000. Int Financ Rev 4:199–225

Harris L, Gurel E (1986) Price and volume effects associated with a change in the S&P500 list: new evidence for the existence of price pressure. J Financ 41:815–829

Hegde S, MeDermott J (2003) The liquidity effects of revisions to the S&P500 index: an empirical analysis. J Financ Mark 6:413–459

Henderson G (1990) Problems and solutions in conducting event studies. J Risk Insur 57:282–306

Jain P (1987) The effect on stock price of inclusion in or exclusion from the S&P500. Financ Anal J 43:58–65

Japan Securities Research Institute (2002) Securities market in Japan. Japan Securities Research Institute, Tokyo. 2000

Kaul A, Mehrotra V, Morck R (2000) Demand curves for stocks do slope down: new evidence from an index weights adjustment. J Financ 55:893–912

Liu S (2000) Changes in the Nikkei 500: new evidence for downward sloping demand curves for stocks. Int Rev Financ 1:245–267

Lynch A, Mendenhall R (1997) New evidence on stock price effects associated with changes in the S&P 500 index. J Bus 70:351–383

Merton R (1987) Presidential address: a simple model of capital market equilibrium with incomplete information. J Financ 42:483–510

Okada K, Isagawa N, Fujiwara K (2006) Addition to the Nikkei 225 Index and Japanese market response: temporary demand effect of index arbitrageurs. Pac Basin Financ J 14:395–409

Peterson P (1989) Event studies: a review of issues and methodology. Q J Bus Econ 28:36–66

Pontiff J (1996) Costly arbitrage: evidence from closed-end funds. Q J Econ 111:1135–1151

Shleifer A (1986) Do demand curves for stocks slope down? J Financ 41:579–590

Shleifer A (2000) Inefficient markets. Oxford University Press, Oxford

Wurgler J, Zhuravskaya E (2002) Does arbitrage flatten demand curves for stocks? J Bus 75: 583–608

Chapter 23
The Calendar Structure of the Japanese Stock Market: The 'Sell in May Effect' Versus the 'Dekansho-Bushi Effect'

Shigeki Sakakibara, Takashi Yamasaki, and Katsuhiko Okada

Abstract We report on a seasonal pattern that has persisted in the Japanese stock market for more than half a century: mean stock returns are significantly positive for months during the first half of the calendar year and significantly negative for months during the second half. Dubbed the "Dekansho-bushi effect," this seasonality is independent of other known calendar anomalies, such as the so-called January effect. The Dekansho-bushi effect should be distinguished from the "sell in May effect," since Japanese stocks perform well in June and poorly in November and December. The Dekansho-bushi effect varies in magnitude among firms and is particularly significant among small firms with low book-to-market ratios. Nonetheless, the effect exists, regardless of a company's size or book-to-market ratio.

Keywords Anomaly • Calendar anomaly • Seasonality

JEL Classification Codes G14

The original article first appeared in the International Review of Finance 13 (2): 161–185, 2013. A newly written addendum has been added to this book chapter.

S. Sakakibara
School of Commerce, Kwansei Gakuin University, Nishinomiya, Japan

T. Yamasaki
Graduate School of Business Administration, Kobe University, Kobe, Japan

K. Okada (✉)
Institute of Business and Accounting, Kwansei Gakuin University, 1-155 Uegahara Ichibancho, Nishinomiya, Hyogo 662-8501, Japan
e-mail: katsuokada@kwansei.ac.jp

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8_23

1 Introduction

Japanese monthly stock returns are significantly higher during the January to June versus the July to December periods. We call this seasonality the *Dekansho-bushi* effect, after a traditional Japanese ballad;[1] which advocates that people work only the first half of the year and spend the second half in leisure. Out of our 59-year time span study of the popular index of Nikkei 225, the index cumulatively advanced during the first half of the year while retreated during the second half in 39 years. The impact of this effect on stock returns was considerable. From 1950 to 2008, the price-weighted Nikkei 225 showed a cumulative gain of 3887.4 % for a buy-and-hold strategy during just the first half of each year, versus a cumulative gain of 102.2 % for just the second half of each year. Using the value-weighted Tokyo Stock Price Index (TOPIX), the disparity between these two strategies is even more dramatic: a gain of 3900.6 % versus a gain of just 69.7 %.[2]

Studies of financial markets in several countries have documented empirical regularities that appear to be inconsistent with the efficient market hypothesis. In the United States, Reinganum (1983) finds that small stocks outperform large stocks in January, while Tinic and West (1984) find that high-beta stocks outperform low-beta stocks in January. French (1980) finds that returns on Mondays are lower and are higher on Fridays, in what is termed the weekend effect. Ariel (1990) finds that returns on the days before holidays are higher, the so-called holiday effect. Ariel (1987) also reports a monthly effect on stock returns: Stocks are higher in the first half of the month and flat during the second half.

As for empirical evidence regarding the Japanese stock market, Kato and Schallheim (1985) report that the January effect is indeed at work there and Sakakibara (1994) confirms the presence of the weekend effect in the index call options market. Bouman and Jacobsen (2002) report a ‘sell in May’ effect in 36 of the 37 countries including Japan. Maberly and Pierce (2005) examine the Japanese popular price-weighted average index (Nikkei 225) and report that the sell in May effect disappeared since the introduction of Nikkei 225 index futures in September 1986.

The findings reported in this chapter add to this list of regularities that are independent of previously reported seasonal patterns. In particular, our findings are not related to the well-known January effect, because the current study’s analysis results remain robust even when January is excluded from the sample months. Although the *Dekansho-bushi* pattern is similar to that of calendar anomalies reported by Bouman and Jacobsen (2002), the Japanese stock market performs in a

[1] *Dekansho-bushi* is a well-known folk song traditionally sung by farmers since the Edo era (1603–1868) in the Sasayama district, located in western Japan. It celebrates a lifestyle of laboring only during the first half of the year and spending the rest of the year in leisure.

[2] The return over each half-year is defined as the sum of one plus the monthly return over that period.

manner different from what is described as "Sell in May and go away."[3] On the basis of our observation of 25 reference portfolios of similar size and book-to-market ratios, most Japanese stocks perform well until June but lag from July to the end of the year. This reflects the "Twain effect," which suggests that October is a dangerous month in terms of stock prices, but by no means the only one. *Dekansho-bushi* pattern in various indices and reference portfolios does not disappear even after the internationalization of the Japanese market. In fact, for comparison with Maberly and Pierce (2005), *Dekansho-bushi* effect becomes slightly more conspicuous after index future was introduced.[4]

The study of such patterns often follows a path in which the popular press mentions a supposedly profitable trading rule, which in turn prompts a scholarly inquiry. The regularity reported in this chapter, however, followed a different path: the subject seasonality was first documented in our working paper in Japanese and then reported in the popular Japanese press, suggesting that the pattern in question has not been well known among investors in the Japanese stock market.[5]

This chapter is organized as follows. Section 2 reports several tests whose results show the existence of a half-year pattern in Japanese stock market returns. Section 3 discusses the results and considers possible biases that could be responsible for the observed effect. Section 4 concludes the chapter.

2 The Half-Year Pattern in Japanese Stock Index Returns

2.1 Monthly Returns of Various Indexes

To represent the returns accruing to stocks, the following tests employ the Nikkei NEEDS Financial Quest to obtain the returns of the value-weighted TOPIX and the price-weighted Nikkei 225 index, the two most commonly quoted Japanese stock indexes. In addition, we obtain the Tokyo Stock Exchange 1st section Arithmetic Stock Price Average and the Nikkei All Stock Average. The data span the years 1950–2008 (708 months) for both TOPIX and Nikkei 225.

The data portrayed in Table 23.1 and Fig. 23.1 indicate the superiority of the trading environment in Japan in the first half of the year, compared to the second half.

[3]According to the saying, the month signals the start of a bear market, so investors are better off selling their stocks in May and holding cash. The adage ends thus: "… but buy back on St. Leger Day." St. Leger Day refers to the day when a horse race is run at Doncaster in England every September.

[4]Nikkei 225 index performs 1.0 % per month better on average in the first half year than the latter for the period 1970–1986. During 1987–2008 period, the average monthly difference between the first and the second half year is 1.1 %.

[5]The half-year seasonality in the Japanese stock market was first reported by our working paper on September 29, 2003; subsequently, the first article mentioning Japanese stock market seasonality appeared in the popular *Nihon Keizai Shimbun* on January 15, 2009.

Table 23.1 Mean monthly returns for five indexes

		First-half year (Jan–Jun)		Last-half year (Jul–Dec)				
	Period	Monthly return	Std. dev.	Monthly return	Std. dev.	Diff.	t-statistic	p-value
	Entire period							
Nikkei 225	1950/1–2008/12	0.012	0.056	0.004	0.059	0.008	1.910	0.056
TOPIX	1950/1–2008/12	0.012	0.051	0.003	0.056	0.009	2.168	0.031
TSE 1st Arithmetic Stock Price Average	1978/1–2008/12	0.010	0.048	−0.008	0.003	0.018	3.357	0.001
Nikkei All Stock Index	1981/1–2008/12	0.010	0.052	−0.002	0.056	0.013	2.144	0.033
Russell/Nomura Japan total market index	1980/1–2008/12	0.010	0.050	−0.002	0.055	0.012	2.136	0.033
	Period before the crash of 1990							
Nikkei 225	1950/1–1989/12	0.018	0.053	0.010	0.053	0.008	1.663	0.097
TOPIX	1950/1–1989/12	0.016	0.049	0.009	0.052	0.007	1.596	0.111
TSE 1st Arithmetic Stock Price Average	1978/1–1989/12	0.018	0.035	0.009	0.034	0.009	1.495	0.137
Nikkei All Stock Index	1981/1–1989/12	0.025	0.045	0.012	0.039	0.013	1.549	0.124
Russell/Nomura Japan total market index	1980/1–1989/12	0.021	0.041	0.012	0.038	0.010	1.348	0.180
	Period after the crash of 1990							
Nikkei 225	1990/1–2008/12	0.000	0.059	−0.009	0.070	0.009	1.017	0.310
TOPIX	1990/1–2008/12	0.002	0.054	−0.009	0.062	0.012	1.506	0.134
TSE 1st Arithmetic Stock Price Average	1990/1–2008/12	0.005	0.054	−0.019	0.062	0.024	3.083	0.002
Nikkei All Stock Index	1990/1–2008/12	0.004	0.054	−0.009	0.062	0.013	1.656	0.099
Russell/Nomura Japan total market index	1990/1–2008/12	0.004	0.053	−0.009	0.061	0.013	1.745	0.082

Note: Choice of periods depends on data availability for each index from Nikkei NEEDS Financial Quest

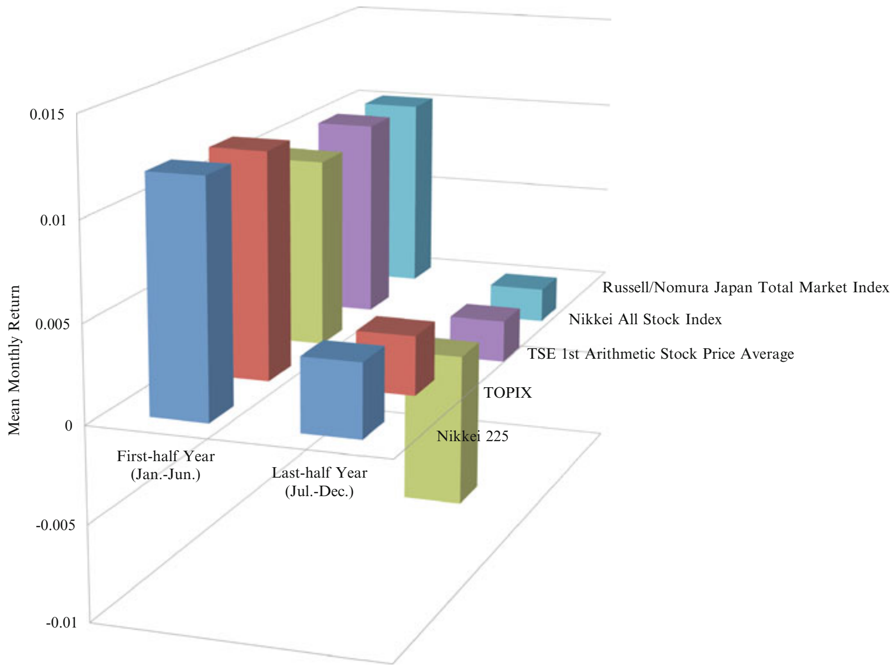


Fig. 23.1 Mean monthly returns for five indexes (*Note*: The mean monthly returns were calculated during sample periods that differed from index to index. The choice of periods depends on the data availability for each index from the Nikkei Economic Electronic Database Systems Financial Quest. TOPIX, Tokyo Stock Price Index; TSE, Tokyo Stock Exchange)

When each trading year is divided evenly into two halves, the mean monthly return for the first half is significantly more than the mean monthly return for the second half. Indeed, the first-half monthly mean is found to be positive, while the second-half monthly means is negative in some indexes. The *t*-statistics for the difference in the mean monthly returns for the two populations are 1.910 for the price-weighted index and 2.168 for the value-weighted index. Figure 23.1 is a graphic representation of Table 23.1, which shows the statistics for the entire period under study. For both TOPIX and Nikkei 225, we calculate the mean monthly return since January 1950; this period includes Japan's post-war high-growth period. Other indexes cover the maximum period, as long as the data are available. The Tokyo Stock Exchange 1st section Arithmetic Average is the simple average index of listed stocks in the Tokyo Stock Exchange, 1st section. The Nikkei All Stock Index, meanwhile, is the capitalization weighted index of all listed stocks in Japan, not limited to the Tokyo Stock Exchange but excluding JASDAQ. The Russell/Nomura Japan index is the capitalization weighted index of all listed stocks; of all the indexes, it covers the largest number of stocks (i.e., 98 % of listed stocks). Nomura/Russell Japan calculates its index based on float-adjusted market capitalization.[6]

Basically, the Japanese economy has undergone a high-growth trajectory since the end of World War II: the Nikkei 225 rose from 109.91 to 38,915.87, and the TOPIX increased from 12.66 to 2,881.37, at the end of 1989. The stock market started declining in the beginning of 1990 and has continued to remain sluggish to this day. In the period comprising January 1990 to December 2008—when the Japanese economy suffered a dramatic decline in the stock and property markets and subsequent prolonged deflationary pressure—the Nikkei 225 plummeted from 37,188.95 to 8,859.56, and the TOPIX similarly dropped from 2,737.57 to 859.24.[7] In retrospect, 1990 marks the turning point of the Japanese economy.

As seen in Table 23.1, for all indexes during the entire period, the *t*-statistic was statistically significant, thereby showing that the mean monthly return for the first half of the trading years significantly exceeded the mean monthly return for the second half. Most notably, for all indexes during the post-1990 period, the mean monthly return for the first half of the trading years was positive, while that for the second half was negative; the differences between the means in this sub-sample period, however, were not always statistically significant. It is worth noting that during a dramatic bear-market period—that is, when every single one of the indexes plummeted to less than one-third of their respective peaks—investing only during the first half of each year produced positive mean monthly returns.

2.2 χ^2 Test and the Impact of Cumulative Return

The first half of the trading years offered a better trading environment and differed significantly from the second half over our sample period. We test the sensitivity of these conclusions as follows: we divide each trading year so that equal numbers of trading months appear in each half-year. The cumulative return over each half-year is defined as the sum of one plus the monthly return over that period. We call this cumulative return a "buy and hold return" (BHR). If the returns for all months of the trading year are drawn from a single distribution, then the probability that the BHR for the first half of a trading year will exceed that for the second half of that same trading year should be 0.5. Therefore, the null hypothesis stated that the expected frequency of higher first-half-year returns would be equal to half the number of years in the test period.

Table 23.2 reports the test statistics resulting from a comparison of this expectation with the observed results. For the price-weighted index (Nikkei 225), first-half

[6]The market capitalization of each stock is calculated as its price × number of shares outstanding × (1 − stable share holdings ratio). This adjustment underweights the performance of less-tradable stocks while overweighting highly liquid stocks. For calculation details, please see www.russell.com/indexes/data/russell_nomura/russell_nomura_indexes.asp

[7]During this period, the commercial property market index issued by the Ministry of Land, Infrastructure, Transport, and Tourism declined from a peak of 271.6 to 76.8.

Table 23.2 χ^2 test

		First-half year (Jan–Jun)		Last-half Year (Jul–Dec)						
	Period	Six-month BHR	Std. dev.	Six-month BHR	Std. dev.	Diff.	t-statistic	p-value	Frequency of higher first-half year returns	χ^2
Nikkei 225	1950–2008	0.0764	0.159	0.0268	0.177	0.050	1.600	0.112	39/59**	6.119
TOPIX	1950–2008	0.0756	0.154	0.0234	0.177	0.052	1.734	0.086	38/59**	4.898
TSE 1st Arithmetic Stock Price Average	1978–2008	0.0632	0.144	−0.0419	0.169	0.105	2.638	0.011	21/31**	3.903
Nikkei All Stock Index	1981–2008	0.0666	0.152	−0.0067	0.184	0.073	1.622	0.111	19/28*	3.571
Russell/Nomura Japan total market index	1980–2008	0.0922	0.156	−0.0036	0.173	0.096	2.214	0.031	19/29*	2.793

Note: The null hypothesis states that the expected frequency of higher first-half-year returns is equal to half the number of years in the test period. The six-month BHR was defined as return over each half-year, as the sum of one plus the monthly returns over that period. * and ** indicate significance at the 10 % and 5 % levels, respectively cumulative returns exceeded those of the second half in 39 of 59 years, and the χ^2 test statistic is 6.119. For the value-weighted index (TOPIX), the first-half cumulative returns exceeded those of the second half in 38 of 59 years, and the χ^2 test statistic was 4.898. Thus, for both indexes, the null hypothesis was rejected at the 1 % significance level. Besides, the expected frequency of higher first-half-year returns among the three other indexes was not equal to half the number of years in the test period.

Table 23.2 also reports the test statistics for the difference in means. For the price-weighted index, the mean first-half BHR (i.e., 6-month BHR) is 7.64 %; this is larger than the second 6-month BHR of 2.68 %, although the difference is statistically insignificant ($p = 0.112$). Likewise, for the value-weighted index, the corresponding means of the first and second 6-month BHRs are 7.56 % and 2.34 %, respectively. Thus, the null hypothesis of equal BHR is rejected at the 10 % significance level ($p = 0.086$).

The cumulative impact of the different monthly mean returns over the 59-year time span is profound. The BHR from investing in the price-weighted index during only the first half of all trading years is 3,887.44 %, while the comparable BHR for investing in the second half is 102.15 %. Likewise, for the value-weighted index, the first and second-half BHRs were 3,900.63 % and 69.65 %, respectively.

The financial impact of following a *Dekansho-bushi* strategy—that is, holding stocks during only the first half of each trading year and hedging the position by selling the index in the second half—is enormous. During the period preceding the peak of the stock market (January 1950–December 1989), such a strategy would have yielded a cumulative return of 4,753.56 % for the price-weighted index and 3,548.38 % for the value-weighted index. Even during the period following the bubble burst (January 1990–December 2008), during which the market lost more than 70 % of its value, the *Dekansho-bushi* strategy would have lost 17.85 % for the price-weighted index and earned 9.65 % for the value-weighted index. Given the magnitude of the bear market, these numbers are surprising.

These differences between mean stock returns in the first and second halves of trading years were not due to outliers, as can be seen from the frequency histogram of those returns (Fig. 23.2). Identical numbers of trading months comprised each of the two populations, so the distributions were directly comparable. The extreme tails of the two distributions were similar; the difference in means was due to a slight shift in the overall distributions of the two populations.

To visually demonstrate the magnitude of the *Dekansho-bushi* strategy, we present a trading simulation on the Nikkei 225 futures market. We collected the closing prices of Nikkei 225 futures traded on the Osaka Stock Exchange since 1988, when the stock index futures market became tradable in Japan for the first time. We assumed that any investor who follows the *Dekansho-bushi* strategy buys at the closing price of the Nikkei 225 March futures contract on the first trading day of January 1989 and rolls the long position until the last trading day of June of the same year. We ignored futurescommissions and interest rate income on the

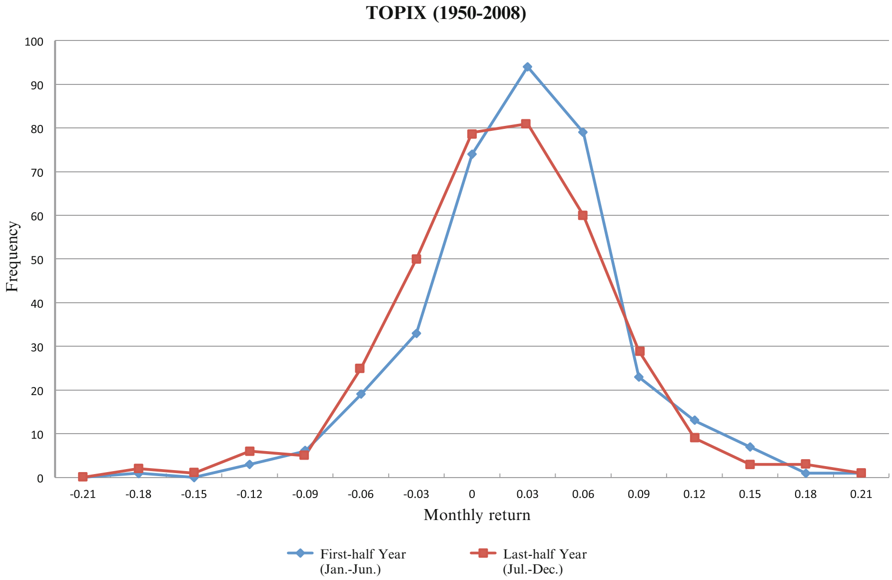


Fig. 23.2 Histograms of monthly return frequencies for the value-weighted index (TOPIX) (*Note:* The Intervals are 3 % wide; each point represents the indicated number of monthly observations with returns falling within that interval. The sub-populations are derived by splitting the year in half at the end of June, so that equal numbers of trading months fall in each half. TOPIX, Tokyo Stock Price Index)

margin balance. Since Nikkei 225 expires every 3 months—namely, in the second week of March, June, September, and December every year—the investor rolls over the position twice in the first half of the year: from the March contract to the June contract, sometime before the second week of March, and from the June contract to the September contract, before the second week of June. We assumed that the rollover had taken place when the open interest of the current futures contract exceeded that of the subsequent one.

Figure 23.3 describes the time-series margin balance of 100 as of January 1989. For comparison, we added the equivalent time-series margin balance from when an investor follows the "reverse *Dekansho-bushi*" strategy—namely, to trade long on the first trading day of July and hold the position until the last trading day of December. As 1989–2008 was a disastrous period for the equity index in Japan, all strategies will reduce the initial margin over time; however, the *Dekansho-bushi* strategy allowed investors to retain more than 90 % of their initial margin at the end of 2008, while the reverse *Dekansho-bushi* and "full investment" strategies each lost two-thirds of it.

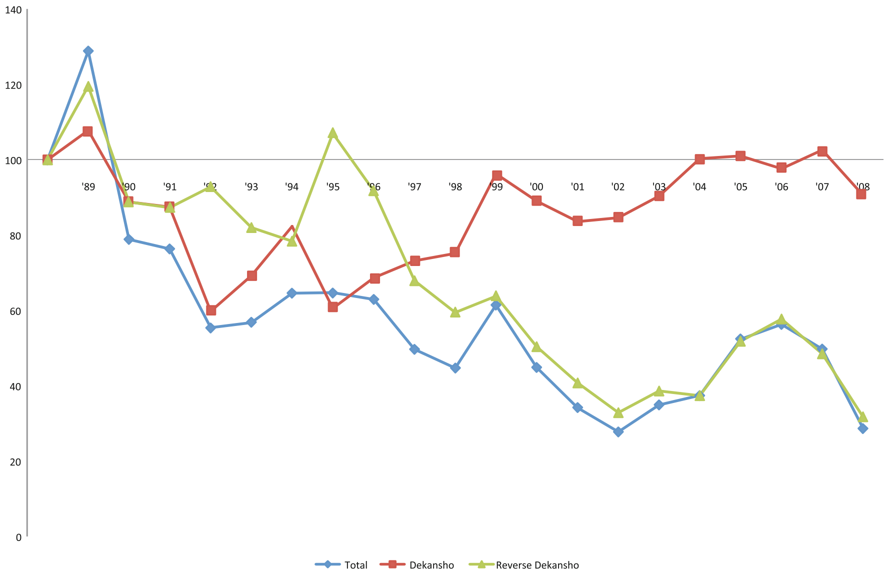


Fig. 23.3 The *Dekansho-bush*i trading strategy applied to Nikkei 225 futures (*Note:* The total indicates the balance of the margin account over time, which has a value of 100 in January 1989, assuming the investor rolls it over every 3 months. *Dekansho-bushi* is the result of a certain trading strategy: long-trade Nikkei 225 futures at the end of the last trading day of December and sell the position on the last trading day of June the following year. Reverse *Dekansho-bushi* trading strategy is to long-trade Nikkei 225 futures at the first trading day of July and sell the position on the first trading day of January the following year)

3 Possible Biases

3.1 Composite Change and New Listing Effect

There is the possibility that Nikkei 225, as a price-weighted average, may be heavily influenced by the price movement of small stocks. Further, the repetitive index composite change in Nikkei 225 may derive some impact that drives the index to move in a manner that *looks like* a seasonal pattern. The TOPIX, being value-weighted, is less susceptible to such change, but could still be potentially affected by the seasonal pattern of new listings. For example, if a large capitalization stock were frequently listed on the Tokyo Stock Exchange 1st section in the first half of the year, the index could perform well in the first half of the year, thanks to the increase in the total market capitalization.

To avoid such biases, we constructed an index of stocks that traded continuously between 1978 and 2008.[8] The equal-weighted calculation of this newly created

[8]Due to a lack of data, the sample is restricted to this period.

index generated a mean monthly return of 2.14 % in the first half of the years examined and −0.61 % in the second half of those years. The difference in the mean was significant at the 1 % CI (*t*-value 4.64). The value-weighted average of this new index also showed a significant *Dekansho-bushi* effect," but to a lesser extent (*t*-value 2.39). The mean monthly return of the first half of the years was 1.10 %, while that of the second half was −0.14 %. This difference may be attributable to the fact that the *Dekansho-bushi* effect" is more pronounced among small stocks.

3.2 Sell in May Effect

Bouman and Jacobsen (2002) report on market seasonality in the Japanese stock market. They conclude that Japanese stock market seasonality is part of a global "sell in May effect." In general, the stock market returns of 37 countries they investigate tend to be below the mean in all months from May through October, although the results tend to be mixed for July. In order for the "sell in May effect" to hold, some poor performance should be seen in May and June. The adage that starts with "sell in May and go away" ends thus: "but buy back on St. Leger Day." Because "St. Leger Day" refers to the date of a horse race run at Doncaster in England every September, the saying suggests that the market will perform poorly in May, June, July, August, and September.

We have documented that the Japanese stock market behaves in a way different from what this saying implies. For this purpose, we created 25 reference portfolios based on size and book-to-market ratios, each of which is reconstituted in August of every year. These portfolios were formed in two steps. First, in August of year t, we ranked all Tokyo Stock Exchange and JASDAQ firms in our population on the basis of market capitalization. Size quintiles were then created, based on these rankings for all Tokyo Stock Exchange 1st section firms. Second, within each size quintile, firms were sorted into quintiles on the basis of their book-to-market ratios in year $t-1$.[9] The returns on the 25 reference portfolios were calculated using equal-weighted averages over the 1978–2008 period.[10] Thus, a reference portfolio with size and book-to-market attributes of 1–1 indicated that the stocks in that category were small and growth stocks. Likewise, the 5–5 reference portfolio contained large and value stocks.

As Fig. 23.4 shows, most reference portfolios performed well in the first half of the year, but returns suddenly declined in the second half. Across all reference portfolios, there was a clear manifestation of the *Dekansho-bushi* effect.

[9]We follow Barber and Lyon's (1997) methodology for creating a reference portfolio. Due to the number of stocks in our population, we employed a quintile rather than decile classification. Further, we reconstituted in August each year, since the majority of shareholder meetings in Japan are held in May and June.

[10]Due to a lack of data, reference portfolio returns were calculated in this period.

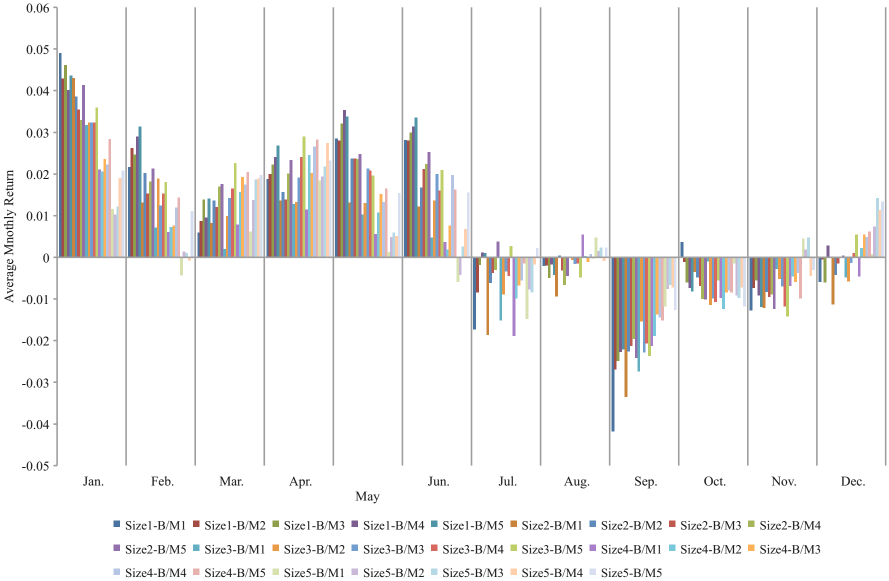


Fig. 23.4 *Dekansho-bushi* effect on 5×5 reference portfolios between 1978 and 2008 (*Note:* Japanese stock market seasonality fits the *Dekansho-bushi* pattern better than the sell in May pattern. Twenty-five reference portfolios are created based on quintile classification of size and book-to-market ratio. Each of the 25 reference portfolios return is calculated for the period between 1978 and 2008. Size1-B/M1 represents stocks that are small growth firms and Size5-B/M5 corresponds to large value firms)

3.3 *Size and Value Effects*

Nikkei NEEDS Financial Quest provides Russell/Nomura-style indexes created by the Nomura Research Institute, a think-tank based in a Japanese brokerage house. Russell/Nomura-style indexes are based on value/growth and size; using these indexes, “size effect” and “value effect” can be estimated with respect to seasonal dependencies. Due to a lack of data, however, mean monthly returns from 1980 were estimated on the basis of sampling. During the 1980–2008 period, value-firm returns exceeded growth-firm returns (value effect), and small-firm returns exceeded large-firm returns (size effect) during both the first and second halves of the trading years. Both size effect and value effect exist in the pre-1990 period (sub-period I, 1980–1989) and the post-1990 period (sub-period II, 1990–2008). The differences between the means for the first and second halves of the trading years were statistically significant in the sample of small firms, but were insignificant in the sample of large firms. Table 23.3 indicates the details of the each group’s mean returns and t-statistics in testing the null hypothesis that the mean monthly returns during the first and second halves of the trading years were equal.

During the full period, the *Dekansho-bushi* effect prevailed among the stocks of both middle- and small-size groups, regardless of book-to-market ratios. In the large-size groups, both value stocks and growth stocks showed stronger performance during the first half of trading years. Only value stocks, however, showed the *Dekansho-bushi* effect to a statistically significant degree. It is noteworthy that performance during the first half of the trading years was better than that of the second half, regardless of the sub-period or sample groups based on company size or book-to-market ratio. Figures 23.5, 23.6, and 23.7 graphically represent Table 23.3.

Table 23.3 Mean monthly returns for size-based portfolios and value-growth portfolios

Russell/Nomura Japan index		First-half year (Jan–Jun)		Last-half year (Jul–Dec)				
Size	Book-to-market	Mean monthly returns	Std. dev.	Mean monthly returns	Std. dev.	Diff.	t-statistic	p-value
Panel A: 1980/1–2008/12 (n = 348)								
Total	Total	0.0102	0.0500	−0.0019	0.0551	0.0120	2.136	0.033
	Value	0.0144	0.0525	−0.0016	0.0534	0.0160	2.816	0.005
	Growth	0.0061	0.0505	−0.0025	0.0600	0.0085	1.431	0.153
Top	Total	0.0074	0.0551	0.0004	0.0598	0.0070	1.142	0.254
	Value	0.0115	0.0573	0.0011	0.0600	0.0105	1.663	0.097
	Growth	0.0040	0.0557	−0.0002	0.0636	0.0042	0.659	0.510
Middle	Total	0.0121	0.0494	−0.0023	0.0546	0.0144	2.575	0.010
	Value	0.0152	0.0542	−0.0023	0.0539	0.0175	3.027	0.003
	Growth	0.0080	0.0490	−0.0024	0.0607	0.0104	1.753	0.080
Small	Total	0.0177	0.0541	−0.0083	0.0580	0.0260	4.315	0.000
	Value	0.0206	0.0554	−0.0074	0.0563	0.0280	4.680	0.000
	Growth	0.0129	0.0562	−0.0093	0.0645	0.0223	3.437	0.001
Panel B: 1980/1–1989/12 (n = 120)								
Total	Total	0.0215	0.0407	0.0118	0.0383	0.0097	1.348	0.180
	Value	0.0270	0.0446	0.0138	0.0403	0.0132	1.698	0.092
	Growth	0.0158	0.0399	0.0095	0.0409	0.0063	0.854	0.395
Top	Total	0.0188	0.0530	0.0131	0.0514	0.0057	0.599	0.550
	Value	0.0242	0.0561	0.0155	0.0559	0.0087	0.852	0.396
	Growth	0.0137	0.0529	0.0111	0.0527	0.0026	0.271	0.787
Middle	Total	0.0243	0.0379	0.0109	0.0348	0.0135	2.025	0.045
	Value	0.0293	0.0460	0.0128	0.0387	0.0165	2.121	0.036
	Growth	0.0185	0.0352	0.0084	0.0396	0.0101	1.474	0.143
Small	Total	0.0248	0.0338	0.0101	0.0354	0.0146	2.313	0.022
	Value	0.0292	0.0362	0.0121	0.0363	0.0171	2.585	0.011
	Growth	0.0198	0.0363	0.0081	0.0390	0.0118	1.710	0.090

(continued)

Table 23.3 (continued)

Russell/Nomura Japan index		First-half year (Jan–Jun)		Last-half year (Jul–Dec)				
Size	Book-to-market	Mean monthly returns	Std. dev.	Mean monthly returns	Std. dev.	Diff.	t-statistic	p-value
Panel C: 1990/1–2008/12 (n = 228)								
Total	Total	0.0042	0.0534	−0.0090	0.0611	0.0133	1.745	0.082
	Value	0.0078	0.0553	−0.0097	0.0577	0.0175	2.336	0.020
	Growth	0.0009	0.0547	−0.0088	0.0673	0.0097	1.191	0.235
Top	Total	0.0014	0.0554	−0.0063	0.0629	0.0077	0.986	0.325
	Value	0.0049	0.0570	−0.0065	0.0609	0.0114	1.456	0.147
	Growth	−0.0011	0.0567	−0.0062	0.0680	0.0051	0.611	0.542
Middle	Total	0.0056	0.0535	−0.0093	0.0615	0.0149	1.944	0.053
	Value	0.0078	0.0568	−0.0103	0.0590	0.0181	2.360	0.019
	Growth	0.0024	0.0542	−0.0081	0.0687	0.0105	1.283	0.201
Small	Total	0.0139	0.0620	−0.0180	0.0650	0.0319	3.794	0.000
	Value	0.0160	0.0629	−0.0177	0.0620	0.0338	4.081	0.000
	Growth	0.0093	0.0641	−0.0185	0.0730	0.0278	3.059	0.002

Note: Panel A reports the mean cumulative return comparison for size-based portfolios (Top, Middle, and Small) for the entire period. Within each size category, the portfolio was subdivided into value and growth, based on book-to-market ratios. Panel B reports the returns of each portfolio during the sub-period before 1990. Panel C reports the returns of each portfolio during the sub-period after 1990

3.4 January Effect

The *Dekansho-bushi* effect may merely be a manifestation of the "January effect." Keim (1983), Roll (1983), and Reinganum (1983) each noted a tendency for the stocks of small firms to earn significant excess returns in January, with much of the effect concentrated in the first few days of the month. Kato and Schallheim (1985), furthermore, confirm the January effect in the Japanese stock market. To determine whether the *Dekansho-bushi* effect reflects nothing more than unusually high mean returns in January, we studied mean monthly returns excluding January. Table 23.4 reports the results based on the Russell/Nomura index, which offers indexes based on size (top, middle, and small) and book-to-market ratios (value and growth). For the total Russell/Nomura index, the mean of the five monthly returns for the first half-year (excluding January) and the mean of the six monthly returns for the second half-year were 1.01 % and −0.19 %, respectively (*t*-statistic for difference of the means = 2.020; implied p = 0.044).[11] Comparable figuresfor the Russell/Nomura

[11] P-values are calculated based on a two-tailed test.

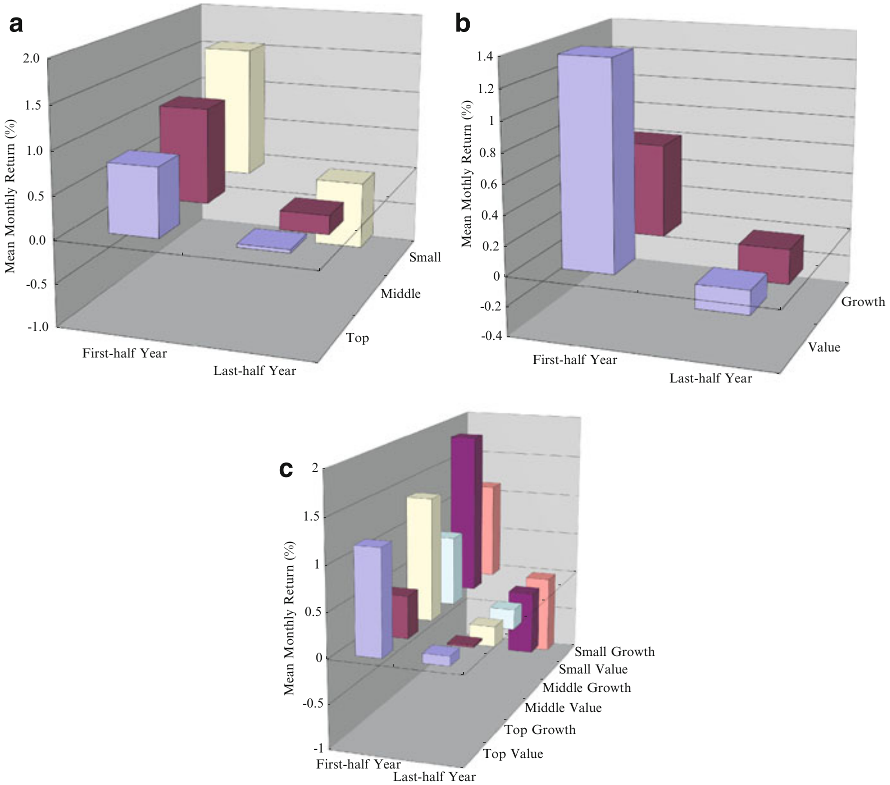


Fig. 23.5 (**a**) The *Dekansho-bushi* effect on stocks in three size categories. (**b**) The *Dekansho-bushi* effect on high book-to-market ratio stocks (growth stocks) and low book-to-market ratio firms (value stocks). (**c**) The *Dekansho-bushi* effect on six different categories

value index and growth index were 1.38 % and −0.16 % ($t = 2.596$; implied $p = 0.010$), and 0.64 % and −0.25 % ($t = 1.411$; implied $p = 0.159$), respectively.

The effect of excluding January from the monthly means was appreciable and in the direction predicted by the January effect. For all three sub-indexes based on size, the means of monthly returns during the first and second half-years were lower when January was excluded. Even with January excluded, however, the *Dekansho-bushi* effect was still found in the remaining months, as evidenced by differences between the means in the first and second half-years; these differences remain statistically significant. Hence, the observed differences in the mean returns for the first and second halves of the trading years are caused by something other than the unusually high returns that occur in early January.

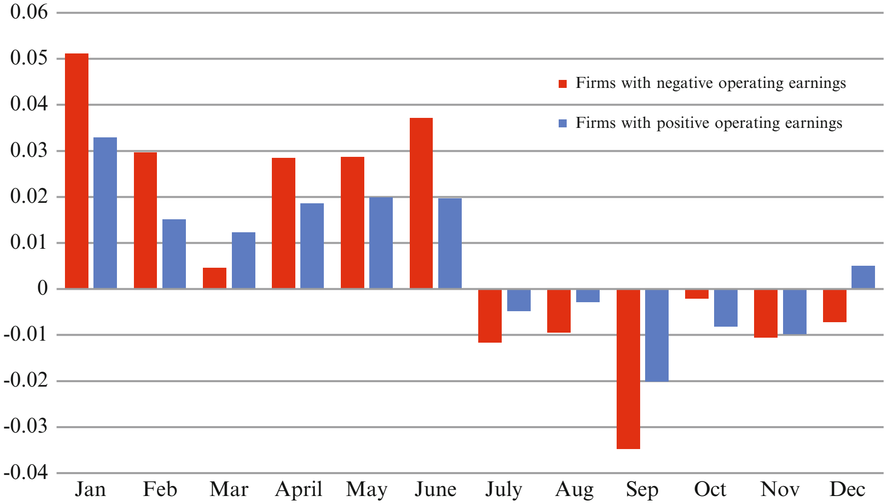


Fig. 23.6 Mean calendar month return for firms with positive and negative operating earnings between 1978 and 2008

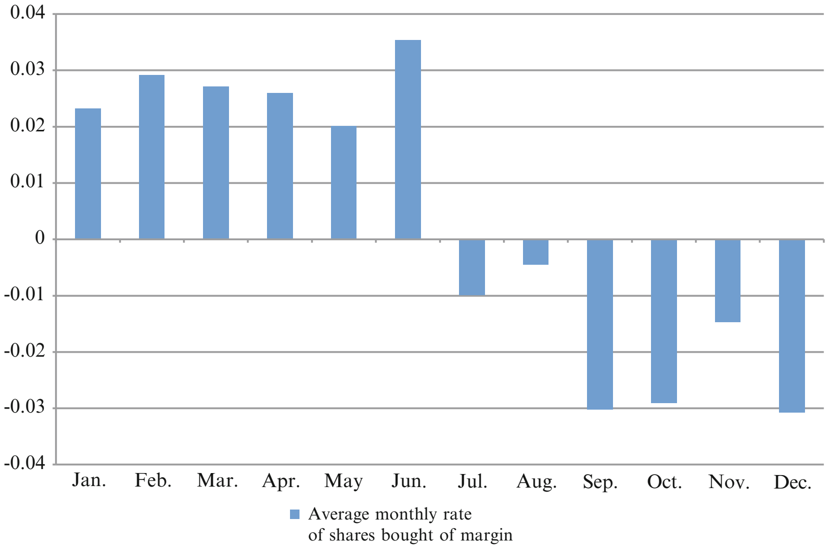


Fig. 23.7 Average monthly rate of shares bought on margin between 1995 and 2008 (*Note*: The monthly rate of shares bought on margin is calculated by dividing the cumulative number of shares bought on margin during a month by the cumulative number of shares bought on margin during the previous month)

3.5 Behavioral Explanation

We ask why the reported seasonal pattern has existed in Japan for most of the post-war era. A possible explanation is offered by the behavioral perspective. Saunders (1993) initiated a new and interesting strand of research to evolve showing that weather in New York affect the stock market return. Hirshleifer and Shumway (2003) conduct a comprehensive investigation of world stock markets excluding Japan confirming correlations between weather and the stock market return and Kato and Takahashi (2004) confirm the similar correlation in Japan. Kamstra et al.

Table 23.4 Mean monthly returns excluding January for size-based portfolios and value-growth portfolios

Russell/Nomura Japan index		First-half year (Jan–Jun)		Last-half year (Jul–Dec)				
Size	Book-to-market	Mean monthly returns	Std. dev.	Mean monthly returns	Std. dev.	Diff.	t-statistic	p-value
Panel A: 1980/2–2008/12 (n = 319)								
Total	Total	0.0101	0.0498	−0.0019	0.0551	0.0120	2.020	0.044
	Value	0.0138	0.0524	−0.0016	0.0534	0.0155	2.596	0.010
	Growth	0.0064	0.0505	−0.0025	0.0600	0.0089	1.411	0.159
Top	Total	0.0081	0.0554	0.0004	0.0598	0.0077	1.182	0.238
	Value	0.0119	0.0584	0.0011	0.0600	0.0108	1.620	0.106
	Growth	0.0048	0.0556	−0.0002	0.0636	0.0050	0.743	0.458
Middle	Total	0.0114	0.0489	−0.0023	0.0546	0.0137	2.342	0.020
	Value	0.0141	0.0532	−0.0023	0.0539	0.0164	2.719	0.007
	Growth	0.0080	0.0495	−0.0024	0.0607	0.0104	1.652	0.100
Small	Total	0.0159	0.0526	−0.0083	0.0580	0.0242	3.868	0.000
	Value	0.0187	0.0536	−0.0074	0.0563	0.0261	4.220	0.000
	Growth	0.0114	0.0559	−0.0093	0.0645	0.0207	3.033	0.003
Panel B: 1980/2–1989/12 (n = 110)								
Total	Total	0.0182	0.0410	0.0118	0.0383	0.0065	0.856	0.394
	Value	0.0243	0.0453	0.0138	0.0403	0.0105	1.284	0.202
	Growth	0.0120	0.0401	0.0095	0.0409	0.0025	0.318	0.751
Top	Total	0.0164	0.0543	0.0131	0.0514	0.0033	0.328	0.744
	Value	0.0229	0.0583	0.0155	0.0559	0.0074	0.675	0.501
	Growth	0.0104	0.0536	0.0111	0.0527	−0.0007	−0.071	0.943
Middle	Total	0.0208	0.0385	0.0109	0.0348	0.0099	1.418	0.159
	Value	0.0257	0.0459	0.0128	0.0387	0.0129	1.598	0.113
	Growth	0.0151	0.0365	0.0084	0.0396	0.0067	0.914	0.363
Small	Total	0.0199	0.0329	0.0101	0.0354	0.0098	1.486	0.140
	Value	0.0242	0.0353	0.0121	0.0363	0.0120	1.755	0.082
	Growth	0.0152	0.0357	0.0081	0.0390	0.0072	0.997	0.321

(continued)

Table 23.4 (continued)

Russell/Nomura Japan index		First-half year (Jan–Jun)		Last-half year (Jul–Dec)				
Size	Book-to-market	Mean monthly returns	Std. dev.	Mean monthly returns	Std. dev.	Diff.	t-statistic	p-value
Panel C: 1990/2–2008/12 (n = 209)								
Total	Total	0.0059	0.0536	−0.0090	0.0611	0.0149	1.853	0.065
	Value	0.0084	0.0552	−0.0097	0.0577	0.0181	2.300	0.022
	Growth	0.0035	0.0552	−0.0088	0.0673	0.0122	1.419	0.157
Top	Total	0.0037	0.0557	−0.0063	0.0629	0.0100	1.204	0.230
	Value	0.0061	0.0578	−0.0065	0.0609	0.0126	1.522	0.129
	Growth	0.0019	0.0567	−0.0062	0.0680	0.0080	0.916	0.361
Middle	Total	0.0065	0.0531	−0.0093	0.0615	0.0157	1.955	0.052
	Value	0.0079	0.0560	−0.0103	0.0590	0.0182	2.276	0.024
	Growth	0.0042	0.0549	−0.0081	0.0687	0.0123	1.412	0.160
Small	Total	0.0138	0.0606	−0.0180	0.0650	0.0318	3.633	0.000
	Value	0.0158	0.0610	−0.0177	0.0620	0.0335	3.920	0.000
	Growth	0.0093	0.0641	−0.0185	0.0730	0.0278	2.900	0.004

Note: Panel A reports the mean cumulative return comparison of size-based portfolios (Top, Middle, and Small) for the entire period. In each of the size categories, the portfolio was subdivided into value and growth, depending on book-to-market ratio. Panel B shows the return of each portfolio during the sub-period before the bubble burst. Panel C shows the return of each portfolio during the sub-period after the bubble burst

(2000) report lower stock returns after weekends with daylight saving time changes. Dichev et al. (2003) and Yuan et al. (2006) relate stock returns to lunar phases. Kamstra et al. (2003) find evidence of a relation between potential mood changes of investors due to a seasonal affective disorder. Cao and Wei (2005) link stock market returns to temperature variations.[12] Individuals in a good mood make more optimistic choices. A highly robust effect is that individuals in a good mood make more positive evaluations of many things, such as life satisfaction, past events, people, and consumer products (see, for example, Wright and Gordon 1992) and the survey of Bagozzi et al. 1999).

From this perspective, there are many seasonal events that may possibly make Japanese people more optimistic throughout the first half of the year. January begins the calendar year, and the optimism prompted by a feeling of a "fresh start" is commonly seen worldwide. This feeling may be especially strong in Japan (Nezlek et al. 2008) Like Christmas for Westerners, *Oshogatsu* (New Year) is an important yearly event for Japanese, who commonly take long vacations around this holiday. April 1 starts the fiscal year for most Japanese corporations and public institutions,

[12] Jacobsen and Marquering (2008) argue that the correlation between weather and the stock market is spurious and it is a mere representation of the Halloween effect. Kamstra et al. (2009) and Jacobsen and Marquering (2009) debate is being underway.

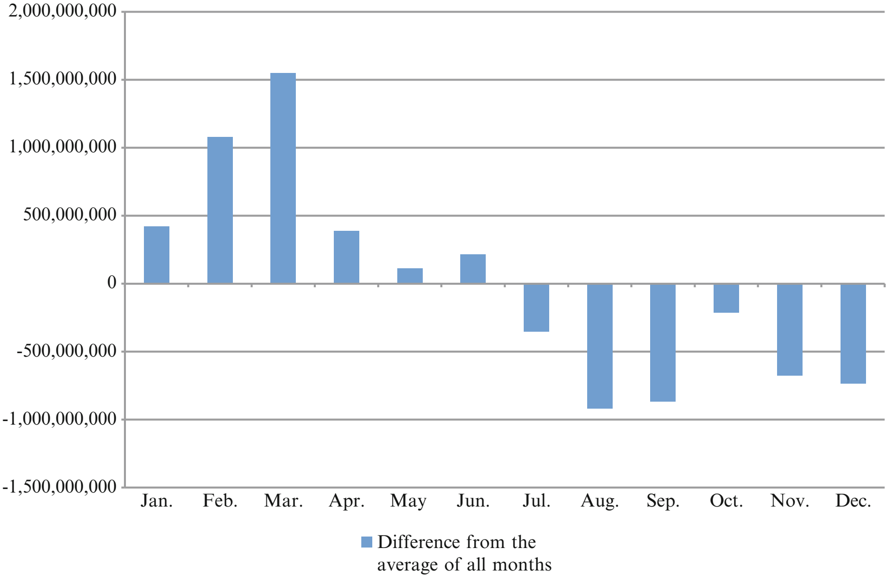


Fig. 23.8 The average number of net shares traded by individual investors between 1978 and 2008 (*Note:* The figures indicate the number of shares bought on margin minus the number of shares sold on margin by individual investors. The graph shows the deviation from the average net number of shares traded by individuals throughout the sample period)

and throughout the country, classrooms and offices fill with fresh faces. Meanwhile, the outdoors is filled with the scent and beauty of cherry blossoms, which also symbolize a fresh start. Toward the end of the month, a series of national holidays called Golden Week begins[13]—another happy time in Japan.[14]

As a naive proxy for optimism among investors, we collect margin balance data for the period 1995–2008.[15] Figure 23.8 shows the average month-over-month percentage changes of shares bought on margin during this period. The monthly rate of shares bought on margin is calculated by dividing the cumulative number of shares bought on margin during a month by the cumulative number of shares bought on margin during the previous month. As Fig. 23.8 illustrates, investors tend

[13]The current National Holiday Laws set nine official holidays, of which four are concentrated in a single week spanning from late April to early May.

[14]Obviously, the feeling of a "fresh new start" is just one example of a factor that can influence one's mood and that investors may be able to control by paying attention to the sources of their mood. On any given day, one might be able to identify myriad other possible influences, such as uncomfortable new shoes, a broken air conditioner, the triumph of a child in school, or the success of a popular local sports team.

[15]The Tokyo Stock Exchange does not disclose margin-related balance data before 1995. Accordingly, our proxy calculation for sentiment is limited to the period after disclosure restrictions were lifted.

to cumulate their margin buy positions during the first half of the year. The rate of margin purchase decreases in the July–August summer period. From September to year end, margin investors tend to unload their positions.

A substantial portion of outstanding shares on the Tokyo Stock Exchange is owned by corporations. Therefore, to substantiate the argument that psychological bias on the part of individual investors might be behind the observed seasonality in the Japanese stock market, we collect data on the margin trading volumes under "on margin transactions" of individuals disclosed by the Tokyo Stock Exchange. We then calculate the total number of shares bought on margin minus the total number of shares sold on margin by individuals each month during 1978–2008. Figure 23.9 shows the differences between each month and the average of all months (Jan–Dec) during this period.

It appearsthat investors are optimistic during the first half of the year but "grew sober" during the second half. Although the causality mechanism remains unclear, it may be that the *Dekansho-bushi* effect is the result of investor behavior triggered by psychological influences.

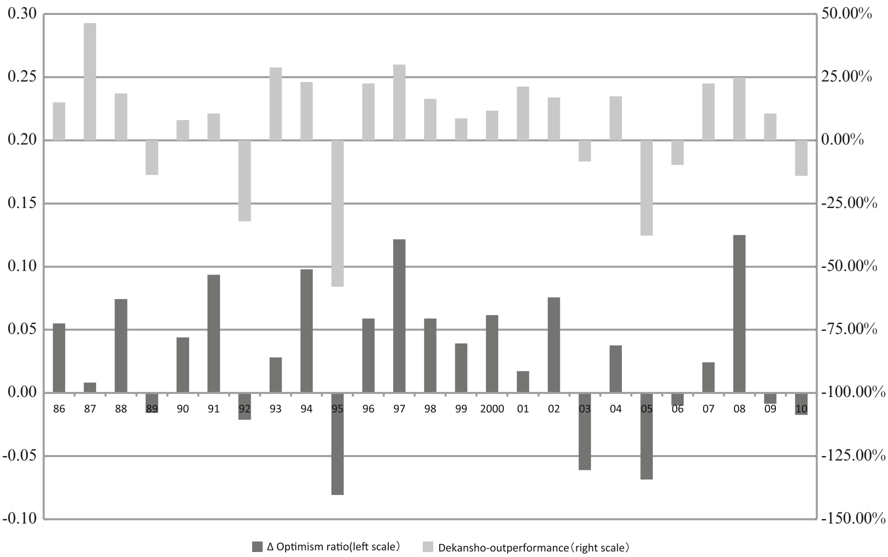


Fig. 23.9 Optimism–pessimism ratio and stock market return (1986–2010) (*Note*: Return differences between the first and second halves of the year for 1986–2010 (*upper chart, right*). The variable Δ*optimism ratio* is defined as $(n_o/(n_o+n_p)) - (n_p/(n_o+n_p))$, where n_o and n_p represent the numbers of optimistic and pessimistic articles, respectively (*lower chart, left*))

4 Conclusion

This chapter reports a longstanding, but recently discovered seasonal pattern that is unique to the Japanese stock market. This phenomenon has not been part of market practitioners' street lore and the Japanese popular press reported its existence only after we published our academic working paper. We call this half-year seasonality the *Dekansho-bushi* effect, after the famous Japanese traditional folk song that advocates a lifestyle of laboring only in the first half of the year and spending the second half in leisure.

The magnitude of this effect is significant. During the 59 years studied, every cumulative market advance occurred during the first halves of the trading years, with the second halves of those trading years contributing negatively.

Various explanations for the *Dekansho-bushi* effect have been considered, including the possibility that it is confounded by the previously reported January effect and size effect. However, after controlling these effects, observed calendar regularity still remains. Another possibility is that the indexes tested are prone to index composite changes, i.e. newly chosen stocks tend to perform better, or new exchange listings in a sense that newly listed stocks tends to underperform after the listings; when tested with our created index of currently traded stocks, however, these market events fail to explain seasonality. The *Dekansho-bushi* effect could be interpreted as part of the already documented sell in May effect on the global equity market; however, closer examination reveals that the seasonal pattern in the Japanese market is unique and does not support the implications of selling in May. Window dressing toward the fiscal year end of March could possibly contribute to the seasonal pattern of the Japanese stock market. However, the *Dekansho-bushi* effect is confirmed in the portfolio consists of firms that have only operating profits.

A number of behavioral explanations for the pattern are possible. The *Dekansho-bushi* effect may have something to do with psychological factors prompted by events in the Japanese calendar. Happy events during the first 6 months of the year lift the spirits of the Japanese people. This can lead naïve individual investors to evaluate prospects more optimistically early in the year. They then spend the second half of the year with more sober dispositions, which has the effect of tightening investment wallets and suppressing stock prices. We present evidence of individual investors' behavior consistent with this conjecture. Japanese individual investors tend to be active market participants in the first half of the year and unwind their positions in the second half. However, causal linkage in these correlations remains unclear.

Addendum: Market Psychology in the News Text[16]

Seasonal Psychology of Investors

Sakakibara et al. (2013) conclude with the behavioral conjecture that investors may be driven by positive events, which make them optimistic. These events in Japan are concentrated in the first half of the year. The calendar new year celebration (*Oshogatsu*) occurs in January, while the fiscal new year starts for companies and schools in April (*Shinnendo* means fresh, new start), and the Golden Week holidays are distributed between late April to early May. The notion that financial market participants may be impacted by psychological factors is not new. For example, the effect of indices crossing psychological barriers, such as the 9,000 level of the Dow, is discussed by Donaldson and Kim (1993). Kamstra et al. (2000) report that sleep desynchronosis caused by daylight savings time has a statistically significant impact on stock returns. The authors also present evidence that global stock market returns are affected by seasonal affective disorder. These studies focus on investor psychology and its correlation to the anomalous behavior of the stock market. If investor psychology is indeed the key driving factor behind market seasonality, the psychology of market news reporters or pundits quoted in the financial press should manifest itself with the seasonal pattern in their word usage in news texts. With this hypothesis in mind, we examine textual data from newspapers to investigate whether a more optimistic outlook is prevalent in the first half of the year than in the latter half.

News Data

We use the four business newspapers published by Nihon Keizai Shimbun Co. Ltd., whose combined circulation across Japan is over five million: *Nihon Keizai Shimbun,* whose morning edition has a daily circulation of three million while the evening edition has a daily circulation of 1.6 million; *Nikkei MJ,* with a marketing focus and a circulation of 0.25 million; *Nikkei Sangyo Shimbun,* with a manufacturing focus and a circulation of 0.18 million; and *Nikkei Veritas,* with a finance focus and a circulation of 0.1 million. From over seven million articles electronically collected for the period 1986–2010, we extract only textual data that refer to the financial market or economic outlook. We call such news predictive and end up with 102,898 market-related predictive news articles. We use these news articles to determine the bullishness/bearishness of the market participants. Given the large number of news articles, manual categorization can be inefficient.

[16]This addendum has been newly written for this book chapter.

Therefore, we used a machine learning algorithm[17] to categorize news into three classes: optimistic, pessimistic, and neutral. Based on this machine categorization, we counted the number of optimistic, pessimistic, and neutral articles published each month from January through December of each year.

Seasonality in the Nikkei News

Addendum Table 23.5 summarizes the results. October has the largest quantity of news articles that have predictive statements. This may be because October is, historically, a volatile month and the market is repeatedly reminded of past disastrous October events, such as Black Thursday in 1929 and Black Monday in 1987. The second to the rightmost column is the proportion of optimistic news. The

Table 23.5 Monthly variation of psychology in news texts, 1986–2010

Month	Predictive articles (n)	Number of optimistic articles (a)	Number of neutral articles (b)	Number of pessimistic articles (c)	Ratio (a)/(a + c)	*p*-value
Jan.	8,571	1,826	5,133	1,612	53.11 %	
Feb.	7,676	1,491	4,600	1,585	48.47 %	
Mar.	8,910	1,779	5,383	1,748	50.44 %	
Apr.	8,476	1,646	5,289	1,541	51.65 %	
May	7,915	1,543	4,881	1,491	50.86 %	
Jun.	8,682	1,777	5,270	1,635	52.08 %	
First half total	**50,230**	**10,062**	**30,556**	**9,612**	**51.14 %**	**0.00067**
Jul.	8,977	1,769	5,455	1,753	50.23 %	
Aug.	8,983	1,732	5,400	1,851	48.34 %	
Sep.	8,302	1,551	5,033	1,718	47.45 %	
Oct.	9,563	1,756	5,879	1,928	47.67 %	
Nov.	8,413	1,591	5,107	1,715	48.12 %	
Dec.	8,430	1,617	5,185	1,628	49.83 %	
Second half total	**52,668**	**10,016**	**32,059**	**10,593**	**48.60 %**	**0.00003**
Total	**102,898**	**20,078**	**62,615**	**20,205**	**49.84 %**	**0.26344**

Notes: The number of articles with optimistic, neutral, and pessimistic outlook in each month is identified by a computer algorithm. The *p*-value in the rightmost column tests the null hypothesis that the probability of having optimistic article (r) is 50 %. The null hypothesis is rejected in both for the first and last half-year samples. The null is not rejected for the entire year sample

[17]We used a machine learning algorithm called Support Vector Machines (SVMs). These are supervised learning models with associated learning algorithms that analyze data and recognize patterns for classification. For details, see Steinwart and Christmann (2008).

higher the ratio, the better the mood. This ratio is by far the highest in January. This is intuitive, in the sense that people tend to make good resolutions at the beginning of the year. The ancient Romans began each year by making promises to the god Janus, for whom the month of January is named. High optimism in January is also consistent with the well-known January effect.

As pointed out by Sakakibara et al. (2013), the Japanese stock market is unique because June exhibits strong positive returns; therefore, the adage "sell in May" is not applicable to the Japanese market. The optimistic mood in June is consistent with this phenomenon. Note that June is the second most optimistic month in our sample period. For the first half of the year, the proportion of optimistic articles ($\hat{r}$) is 51.14 %. The null hypothesis, H0, states that the ratio of optimistic articles (r) is 50 %. In alternative hypothesis H1, $r > 0.5$. Our test statistics reject H0 at the 1 % confidence interval. Optimism starts to fade in the latter half of the year and the most pessimistic month is September, closely followed by October. For the latter 6 months of the year, $\hat{r}$ is 48.60 %. This observation (H0: $r = 0.5$ and H1: $r < 0.5$) rejects H0 at the 1 % confidence interval. The second half of the year is significantly pessimistic. For the entire year, the ratio is 49.84 % and the null hypothesis is not rejected, which confirms that our sample is not skewed toward optimism or pessimism throughout the year.

Thus far, our results indicate that the *Dekansho-Bushi* seasonal pattern in the Japanese stock market is synchronous with optimism in newspaper articles. In this section, we observe a 25-year period of stock market returns and determine whether non-*Dekansho-Bushi* years are synchronous with the optimism–pessimism ratio of newspaper articles.

Addendum Fig. 23.9 shows two bar charts. The top chart represents a simple seasonal return difference for the Tokyo Stock Exchange first-section value-weighted average or the first and second halves of each year of our sample period. The lower chart indicates changes in the difference of the optimism ratio for each year, defined as

$$\Delta Optimism\ ratio = \left(n_o / \left(n_o + n_p\right)\right) - \left(n_p / \left(n_o + n_p\right)\right)$$

where n_o and n_p are the numbers of optimistic and pessimistic articles, respectively.

As shown in Addendum Fig. 23.9, Δ*Optimism ratio* is almost perfectly correlated with stock market seasonality. A year with a more optimistic outlook in the first half demonstrates higher returns in the January–June period than in the July–December period, without exception. A year with a pessimistic outlook in the first half of the year demonstrates lower returns in the January–June period than in the July–December period, with only one exception, 2009. This was the year after the financial crisis and the media outlook on the market was presumably bleak. As a result, pessimism prevailed in the first half of 2009 while the stock market rebounded sharply from its oversold condition.

References

Ariel RA (1987) A monthly effect in stock returns. J Financ Econ 18:161–174

Ariel RA (1990) High stock returns before holidays: existence and evidence on possible causes. J Financ 45:1611–1626

Bagozzi R, Gopinath M, Nyer P (1999) The role of emotions in marketing. J Acad Mark Sci 27:184–206

Barber BM, Lyon JD (1997) Detecting long-run abnormal stock returns: the empirical power and specification of test statistics. J Financ Econ 43:341–372

Bouman S, Jacobsen B (2002) The halloween indicator, "sell in may and go away": another puzzle. Am Econ Rev 92:1618–1635

Cao M, Wei J (2005) Stock market returns: a note on temperature anomaly. J Bank Financ 29:1559–1573

Dichev I, Janes D, Troy D (2003) Lunar cycle effects in stock returns. J Priv Equity 6:8–29

Donaldson G, Kim H (1993) Price barriers in the dow jones industrial average. J Financ Quant Anal 28:313–330

French KR (1980) Stock returns and the weekend effect. J Financ Econ 8:55–70

Hirshleifer D, Shumway T (2003) Good day sunshine: stock returns and the weather. J Financ 58:1009–1032

Jacobsen B, Marquering W (2008) Is it the weather? J Bank Financ 32:526–540

Jacobsen B, Marquering W (2009) Is it the weather? Response. J Bank Financ 33:583–587

Kamstra MJ, Kramer LA, Levi MD (2000) Losing sleep at the market: the daylight savings anomaly. Am Econ Rev 90:1005–1011

Kamstra MJ, Kramer LA, Levi MD (2003) Winter blues: a SAD stock market cycle. Am Econ Rev 93:324–343

Kamstra MJ, Kramer LA, Levi MD (2009) Is it the weather? Comment. J Bank Financ 33:578–582

Kato K, Schallheim JS (1985) Seasonal and size anomalies in the Japanese stock market. J Financ Quant Anal 20:243–260

Kato K, Takahashi T (2004) Tenki seirou naraba kabu takashi. Gendai Financ 3:35–50 (in Japanese)

Keim DB (1983) Size-related anomalies and stock return seasonality: further empirical evidence. J Financ Econ 12:13–32

Maberly ED, Pierce RM (2005) The halloween effect and Japanese equity prices: myth or exploitable anomaly. Asia-Pacific Financ Markets 10:319–334

Nezlek J, Sorrentino R, Yasunaga S, Otsubo Y, Allen M, Kouhara S, Shuper P (2008) Cross-cultural differences in reactions to daily events as indicators of cross-cultural difference in self-construction and affect. J Cross Cult Psychol 39:685–702

Reinganum MR (1983) The anomalous stock market behavior of small firms in January: empirical tests for tax-loss effects. J Financ Econ 12:89–104

Roll R (1983) Vas ist Das? The turn-of-the-year effect and the return premia of small firms. J Portf Manag 9:18–28

Sakakibara S (1994) Weekend effect in the Japanese stock index options market. Investment 47(1):4–20 (in Japanese)

Sakakibara S, Yamasaki T, Okada K (2013) The calendar structure of the Japanese stock market: the 'sell in May effect' versus the 'Dekansho-Bushi effect'. Int Rev Financ 13:161–185

Saunders EM (1993) Stock prices and wall street weather. Am Econ Rev 83:1337–1345

Steinwart I, Christmann A (2008) Support vector machines. Springer, New York

Tinic SM, West R (1984) Risk and return: January vs. the rest of the year. J Financ Econ 13:561–574

Wright WF, Gordon BH (1992) Mood effects on subjective probability assessment. Organ Behav Hum Decis Process 52:276–291

Yuan K, Zheng L, Zhu Q (2006) Are investors Moonstruck? – lunar phases and stock returns. J Empir Financ 13:1–23

ERRATUM

Chapter 22
Addition to the Nikkei 225 Index and Japanese Market Response: Temporary Demand Effect of Index Arbitrageurs

Katsuhiko Okada, Nobuyuki Isagawa, and Kenya Fujikawa

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8

DOI 10.1007/978-4-431-55501-8_24

The name of an author was spelled incorrectly in the Table of Contents. For Chapter 22, the third author's name should read **Kenya Fujiwara**.

Also, on the opening page of Chapter 22, in the list of authors following the title, the name of the third author should be **Kenya Fujiwara**, and at the bottom of page, in the list of authors' affiliations, the name should be shown as **K. Fujiwara**.

The online version of the original chapter can be found at
http://dx.doi.org/10.1007/978-4-431-55501-8_22

Index

S. Ikeda et al. (eds.), *Behavioral Interactions, Markets, and Economic Dynamics*,
DOI 10.1007/978-4-431-55501-8