Title	Capital Market Dynamics and Prices
Author(s)	Ikeda, Shinsuke
Citation	大阪大学, 1997, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3132581
rights	
Note	

①

Capital Market Dynamics and Prices

Shinsule IKEDA
The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567, JAPAN

July 8, 1997

1

Capital Market Dynamics and Prices

Shinsule IKEDA
The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567, JAPAN

July 8, 1997

Preface

This dissertation is prepared for partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Economics, Osaka University. It contains eight chapters on capital market dynamics. The respective chapters are written based upon the articles that were individually published from 1991 through 1997 in academic journals and discussion papers. The original publications are as follows (Chapter 1 is an introduction):

1. Chapter 2: Shinsuke Ikeda, 1997, Consumer interdependence and dynamics, a revised version of: Shinsuke Ikeda, 1995, Time preference, intertemporal substitution, and dynamics under consumer interdependence, ISER Discussion Paper No. 386.

2. Chapter 3: Shinsuke Ikeda and Ichiro Gombi, 1997, Habits, costly investment, and current account dynamics, ISER Discussion Paper No. 442.

3. Chapter 4: Shinsuke Ikeda and Akihisa Shibata, 1992, Fundamentals-dependent bubbles in stock prices, The Journal of Monetary Economics 30, 143-168.

4. Chapter 5: Shinsuke Ikeda and Akihisa Shibata, 1995, Fundamentals uncertainty, bubbles, and exchange rate dynamics, The Journal of International Economics 38, 199-222.

5. Chapter 6: Shinsuke Ikeda, 1991, The continuous-time APT with diffusion factors and rational expectations: A synthesis, The Economic Studies Quarterly 42, 124-138.

6. Chapter 7: Shinsuke Ikeda, 1991, Arbitrage asset pricing under exchange risk, The Journal of Finance 46, 447-455.

7. Chapter 8: Shinsuke Ikeda, 1996, An intertemporal capital asset pricing model with stochastic differential utility, Japanese Journal of Financial Economics 1, 55-68.

8. Chapter 9: Shinsuke Ikeda, 1997, Optimal consumption and asset pricing under market incompleteness: A simple approach, a substantially revised version of: Shinsuke Ikeda, 1995, A simple approach to arbitrage asset pricing in incomplete markets, ISER Discussion Paper 380

In preparing this dissertation I owe substantial debts, directly or indirectly, to many persons including my teachers, colleagues, and students. It is impossible to make a complete list of acknowledgments. Among them, however, I would like to give special thanks to Akihiro Amano, Yoshiyasu Ono, and Yoshiro Tsutsui for continuous encouragement, invaluable suggestions on the selection of topics, and helpful comments on the original articles. I am also grateful to Ichiro Gombi and Akihisa Shibata for giving me a willing consent to incorporate several articles coauthored by them into this dissertation. Acknowledgements for the individual articles are listed below. Last, but by no means least, thanks go to my family and parents for their encouragement and help.

Acknowledgements for the individual articles

Chapter 2: My special thanks go to H. Hayakawa, Y. Ono, and especially an anonymous referee for helpful comments to improve the paper. I am also indebted for discussions to H. Cole, F. Dei, Y. Fukuta, K. Futagami, K. Itogawa, H. Izawa, C. Moriguchi, F. Ohtake, H. Osano, S. Serizawa, A. Shibata, K. Shimomura, Y. Tsutsui, K. Ueda, R. Watson, H. Yoshikawa, and seminar participants at the Finance Forum, the Kada Conference, Kobe University, Kobe-Gakuin University, Osaka University, Ritsumeikan University, and the University of Tokyo. This research is financially supported by the Japan Securities Scholarship Foundation and Grants-in-Aid for Scientific Research B, Ministry of Education, Science and Culture (#06451106).

Chapter 3: Our special thanks go to S. Ishizawa, Y. Ono, and S. Turnovsky for invaluable comments on earlier versions of this paper. We are also indebted to M. Kurima, B. Rowlinson, and participants at the Kobe Conference on International Economics and Finance 1997 and at seminars in Tezukayama and Osaka Universities for helpful discussions.

Chapter 4: This research is supported by Grant-in-Aid for Modern Management Science (Kobe University) and the Segawa Foundation (Osaka City University). Our special thanks go to two anonymous referees for invaluable comments. We are also indebted to A. Amano, K. Asako, F. Dei, J. Eaton, H. Hayakawa, K. Hirayama, H. Hisamoto, C. Y. Horioka, Toshiaki Itoh, K. Kiyono, T. Kobayashi, M. Nishio, Y. Ono, and the seminar participants at Kansai, Kobe, Kyoto, Nagoya Gakuin, and Yokohama National Universities and the University of Tokyo for helpful discussions. Of course, all remaining errors are ours.

Chapter 5: We wish to thank C. McKenzie and two anonymous referees for their helpful comments. We are also indebted to K. Ban, T. Fukushima, H. Hori, M. Ikeda, O. Kamoike, K. Mino, M. Otsuki, S. Takagi , Y. Tsutsui, M. Yoshida, and seminar participants at Musashi, Tohoku, and Osaka Universities for their valuable comments. Y. Nakanishi kindly helped us in preparing figures. The first author acknowledges the financial support of the Economics Faculty Fortieth Anniversary Program, Osaka University.

Chapter 6: I thank Yoshio Iihara for invaluable comments on the earlier version of this paper at the meeting of the Japan Association of Financial Research, which was held at Shiga University on May 13, 1989. I am also indebted to an anonymous referee and the participants in workshops at Kansai, Kobe, Osaka, and Toyama Universities where the earlier version was presented for their helpful comments and discussions.

Chapter 7: I am grateful to Akihiro Amano, Fumio Dei, Jonathan Eaton, Yoshiyasu Ono, René Stulz (Managing Editor of the Journal of Finance), an anonymous referee, and especially Charles Yuji Horioka for helpful comments. I also thank Kenjiro Hirayama, Toshio Serita, Yasuhiko Tanigawa, Yoshiro Tsutsui, and other members of the Monetary Economics Workshop for fruitful discussions. All errors are solely my responsibility.

Chapter 8: I thank Y. Iihara, T. Inoue, T. Kobayashi, H. Morita, R. Watson, an anonymous referee, and seminar participants at Kobe, Nanzan, Rikkyo, and Yokohama National Universities and the Finance Forum for helpful comments on earlier versions of this paper. This research is financially supported by Grants-in-Aid for Scientific Research (B), Ministry of Education, Science and Culture (#06451106).

Chapter 9: I am grateful to Y. Fukuta and H. Toda for helpful discussions. This research is financially supported by the Japan Securities Scholarship Foundation.

Contents

Chapter 1

Introduction

The last two decades have witnessed rapid developments in capital asset markets. On the one hand this progress enables more efficient intertemporal resource allocations while, on the other hand, it has caused much fluctuation and sometimes "chaos" in asset markets, such as stock markets and foreign exchange markets. Stimulated by this, of course, economists have done much from the empirical and theoretical viewpoints and from practical as well as academic interests (the resultant cumulative research is merged as a central field of economics under the name of intertemporal macroeconomics, financial macroeconomics, or macro finance.) However, there are many problems unsolved, as will be demonstrated in respective chapters. My main interest in the last one and a half decades has been to make thrusts on them.

The objective of this dissertation is to make some contributions to the theory of capital market dynamics. The examination is conducted especially in wealth dynamics, asset price bubbles, and asset pricing. The following part of the dissertation is structured in three parts, I through III. Part I treats wealth dynamics. Part II develops a new theory of rational bubbles. And Part III examines asset pricing theory.

Part I contains Chapters 2-3. In these two chapters I examine the implications of interdependent consumption preferences for capital market dynamics. In Chapter 2, the effects of consumer interdependence through consumption externalities are examined from the viewpoints of the saving decisions and macrodynamics. This old [e.g., Duesenberry (1949)], but not commonplace issue is reexamined by using a modern rational consumer model. By developing new notions of preferences which captures consumers' interactions, the dynamics of the wealth distribution and of the interest rate are analyzed. It will be stressed that the new notions can address several stylized facts regarding intertemporal substitution, and that the resulting properties of equilibrium dynamics substantially differ from those which are obtained by "Ramsey's conjecture."

Chapter 3 examines the implications of consumption habits in a small open economy. The main interest is, first, to explore several stylized facts regarding the dynamics of the current account and, secondly, to derive the welfare implications. I do this by combining the model of the capital adjustment costs and that of consumption habit formation. As an important message, several empirical facts regarding the current account will be consistently rationalized by choosing the adjustment costs for savings (psychological costs owing to habits) and investment (physical costs). Concerning the welfare implications, it will be pointed out that under strong habit persistence a fiscal policy which is beneficial from the initial welfare viewpoint may have a harmful hangover effect on the future welfare.

In Part II, two chapters 4 and 5 present the theory of rational bubbles which depend on "market fundamentals." Using a continuous-time model of stock prices with dividends growing stochastically, Chapter 4 examines stock price bubbles which depend on dividends. The dynamic properties of these new bubbles are analyzed especially on the stochastic stability, the possible patterns of sample paths, and the correlation with fundamentals.. These contributions might be useful to resolve the problem that the existing models of fundamentals-*independent* rational bubbles are often rejected by empirical data [e.g., Journal of Economic Perspective 4]. Indeed I will show that the empirical methods that have been used for testing the absence of bubbles are not robust in testing for our fundamentals-*dependent* bubbles.

Chapter 5 applies the same idea to the exchange rate dynamics. Besides the similar dynamic analysis to that of the previous chapter, two main contributions are conducted. First, the possibility of cyclic bubbles are derived. Second, a closed form solution of fundamentals-dependent bubbles can be utilized to obtain solutions for the equilibrium exchange rate under any regimes of stochastic policy switchings [e.g., target zones (Krugman, 1991)].

Part III, Chapters 6 through 9, extends the existing asset pricing theory in four directions. The first two chapters treat arbitrage pricing, especially the arbitrage pricing theory (APT) developed by Ross (1977) whereas the remaining two examine asset pricing based upon the market equilibrium. In Chapter 6, the APT, which was developed in a static setting, is recast in a dynamic model as a general theory of arbitrage asset valuation. Reformulating the APT-type arbitrage-free condition in terms of an asset price function, I reduce the condition to a partial differential equation with respect to the function. The pricing formulae are derived by solving this equation to consistently demonstrate the various existing ideas of arbitrage asset evaluation.

Chapter 7 extends the APT to an international setting. Specifying a linear factor return-generating model in local currency terms, I show that the arbitrage-free expected returns do not satisfy the simple APT equation unless they are adjusted by the cost of exchange risk hedging.

In Chapter 8, intertemporal capital asset pricing and the stochastic properties of optimal portfolios and consumption are examined in a continuous-time recursive utility (stochastic differential utility) model with multiple state variables dynamically affecting investment opportunities. The main concern in this chapter is to clarify how this generalization of preferences affects the intertemporal capital asset pricing model (ICAPM) and the consumption-based CAPM.

Chapter 9 analyzes consumption/portfolio choices and asset pricing under market incompleteness. To avoid difficulties (explained in the chapter) in the incomplete market analysis, I here propose that the consumers' problem is transformed into a reduced space induced by the security markets. Introducing new concepts to represent the assets which span this reduced space, I characterize the security market equilibrium. This approach enables us to apply various theorems in complete market settings [see Duffie (1996)], such as the martingale approach, the representative-agent pricing formula, and the APT, to the case of incomplete markets.

Bibliography

[1] Duesenberry, J., 1949, Income, saving, and the theory of consumer behavior (Harvard University Press, Boston).

[2] Duffie, D., 1996, Dynamic asset pricing theory, second edition (Princeton University Press, Princeton, NJ).

[3] Krugman, P., 1991, Target zones and exchange rate dynamics, Quarterly Journal of Economics 106, 669-682.

[4] Journal of Economic Perspectives 4 (Symposium on bubbles), 1990.

[5] Ross, S., 1976, The arbitrage theory of capital asset pricing, Journal of Economic Theory 13, 341-360.

Part I

Wealth Dynamics

Chapter 2

Consumer Interdependence and Dynamics

Abstract: Implications of consumer interdependence for saving decisions and macrodynamics are examined using a rational consumer model with consumption externalities. The consumption-saving choices of an agent are governed by an externality-induced time preference and elasticity of intertemporal substitution, which hinge on other consumers' preferences and consumption. Using this property, recent empirical studies supporting nonconcave utilities and various time preference schedules can be rationalized. The Ramsey equilibrium under time-preference differentials is substantially modified. 'Social addiction' to consumption resulting from mutual bandwagon externalities causes pathological dynamics. Under asymmetric mutual externalities, the equilibrium dynamics depend crucially on the initial wealth distribution.

JEL Classification Numbers: D90, E21, D58.

Keywords: Consumer interdependence, consumption externalities, intertemporal substitution, time preference, addiction.

2.1 Introduction

The consumption choice of one agent is usually affected by that of the other agents due to sociological factors, such as envy, emulation, mob-psychological imitation, etc. As stressed by classical authorities [e.g., Morgenstern (1948) and Duesenberry (1949)], this consumer interdependence is particularly important in understanding macroeconomic phenomena since consumers' interactions often affect qualitatively their aggregate behavior. Except for only several papers,[1] however, this social aspect of consumption has not been shed light on in recent economic analyses. In particular, little is known about how consumer interdependence affects the consumption-saving choice and macroeconomic dynamics.

The goal of this paper is to fill the void: Focusing on consumption externalities which affect intertemporal substitution, I shall examine implications of

[1] For example, see Abel (1990), Robson (1992), Blomquist (1993), Galí (1994), Zou (1995), and Bakshi and Chen (1996). Examples of the microeconomic analysis incorporating consumer interdependence include Becker (1974) and Hayakawa and Venieris (1977).

consumer interdependence for saving decisions and the dynamic property of the perfect foresight equilibrium. Consumers' interactions are incorporated into an optimizing two-agent model by assuming that the instantaneous utility function of one agent contains the other agent's consumption rate as an external preference-shifting parameter.

This economy thus differs from the standard heterogeneous-agent economy [e.g., Becker (1980)] only in that there is contemporaneous interdependence of preferences through consumption externalities. This difference, however, is sufficient to generate a distinct feature of interdependent consumption-saving behavior: the optimal consumption-saving choice is governed by an effective time preference rate and an effective elasticity of intertemporal substitution which are induced by consumption externalities. The two preference parameters will be called induced time preference rates and the induced elasticity of intertemporal substitution, respectively. Both parameters hinge on the other agents' subjective preference-parameters and consumption through externalities. In particular, even under constant subjective utility-discounting rates, the induced time preference rate is endogenously determined by the two agents' consumption rates. In this sense, the model developed here could be taken as a new model of endogenous time preferences. It will be shown that various time preference schedules including the ones proposed by the existing literature [e.g., Fisher (1907), Uzawa (1968), Fukao and Hamada (1991)] can be derived by specifying consumption externalities in alternative forms.

The distinct feature has two implications. First, lots of articles which estimate Euler equations for optimum consumption have reported empirical results against the concavity of the utility function [e.g., Mankiw, Rotemberg, and Summers (1985)]. As often commented, do these results imply that the consumption data observed cannot be demonstrated by using interior optimum solutions of competitive households? The reply of the present paper is positive: It will be shown that under strong mutual bandwagon externalities the induced elasticity of intertemporal substitution is negative, implying that the utility function is seemingly or socially convex while it is concave from the subjective viewpoint. This socially negative intertemporal substitution will be referred to as social addiction, meaning that an increase in consumption enhances the marginal rate of substitution through mutual bandwagon externalities.

Second, as well-known [e.g., Becker (1980)], under independent preferences, the dynamics under time preference differentials are extremely simple: in the long-run, the most patient agent owns all the assets, with the real interest rate being equal to the lowest subjective discount rate whereas the less patient agents consume just the quantity required for subsistence. In contrast, I shall show that the equilibrium dynamics under consumption externalities display richer patterns. For example, (i) the long run interest rate, which is determined by induced time preferences, can be either lower than the lowest subjective discount rate or higher than the highest one, depending on the sign and magnitude of consumption externalities. (ii) Under mutual social addiction, the interest rate increasingly approaches some long run rate, and asset prices continue to decline from the initial highest level. (iii) When consumption externalities induce strong mutual substitutability between individual agents' consumption, the order of impatience and of the long-run wealth holdings among consumers are reversed. (iv) Under asymmetric mutual externalities, there may exist two "sinks of attraction," so that which steady state is attained depends on the initial wealth distribution.

The contribution might be limited by the underlying restrictive setting in which any strategic aspects are assumed away in this two-agent model. An implicit assumption which justifies this is that "agent" i is the representative agent of group i (more simply, of country i) which is composed of an infinite number of identical agents, so that each agent in group i considers himself as too small to affect any other agents in groups i and j.[2] In this sense, the consumption externality treated in this paper is a consumption version of the Marshallian externality.

The organization of the paper is as follows. In Section 2.2 an intertemporal consumption choice problem under consumption externalities is examined. The optimal consumption rule is then summarized as a modified Keynes-Ramsey rule. In Section 2.3 the equilibrium dynamics with constant symmetric parameters are examined. In Section 2.4 the analysis is extended to the case of asymmetric externalities. In Section 2.5 conclusions are summarized and possible future directions are suggested.

2.2 Intertemporal Consumption Choices under Consumption Externalities

2.2.1 Utility and Consumption-Saving Choices

Consider a one-good exchange economy populated with two immortal (representative) agents 1 and 2. The good is nonstorable and there is no capital accumulation. Faced with perfect bond markets, each agent maximizes his or her lifetime utility which is defined as the present value of the stream of instantaneous utility or felicity.

The original feature of the model lies in that consumer interdependence is caused by consumption externalities. These externalities are modeled by assuming that the time-t felicity of agent i $(i = 1, 2)$, $u^i(t)$, contains the other agent's current consumption rate, $c_j(t)$ $(j \neq i)$, as an external preference-changing parameter, as well as his own, $c_i(t)$: $u^i(t) = u^i[c_i(t); c_j(t)]$. This felicity function is assumed to be strictly increasing and strictly concave in c_i, and twice-differentiable with respect to the two arguments. To focus on consumption externalities which affect intertemporal substitution, I assume that the cross derivatives, $u^i_{ij}(c_i; c_j) \equiv \partial^2 u^i(c_i; c_j)/\partial c_i \partial c_j$ $(i \neq j)$, are nonzero.

Given these assumptions, the lifetime utility is specified as:

$$U_i = \int_0^\infty u^i[c_i(t); c_j(t)] \exp(-\delta_i t)\, dt, \quad i, j = 1, 2\, (i \neq j), \tag{2.1}$$

where δ_i denotes a constant subjective discount rate. The magnitudes of δ_1 and δ_2 are different from each other. Without loss of generality, agent 1 is assumed

[2] Similar assumptions are usually (and implicitly) made in various contexts. In the representative agent models of a closed economy [e.g., Blanchard and Fischer (1989)], the representative agent is assumed to behave as a competitor although she is really a singleton in markets. In the two-country model [e.g., Frenkel and Razin (1985) and Ikeda and Ono (1992)], the representative agent in each country is assumed to behave in a nonstrategic manner despite the fact that each agent really affects the utility of the representative agent in the other country through pecuniary externalities.

to discount the future felicity less than agent 2:

$$\delta_1 < \delta_2.$$

If it were not for consumer interactions, the discount rate would always equal the time preference rate [see Obstfeld (1990)] and agent 1 could be taken as more patient than agent 2. As shown below, however, agent 2 here can be essentially more patient than agent 1.

In Eq. (3.2), cross derivative u^i_{ij} determines the property of the external effect of agent j's consumption on agent i's marginal felicity. When $u^i_{ij} > 0$, c_j is an Edgeworth-complement with c_i for agent i. Borrowing Leibenstein (1950)'s terminology, a positive external effect on one's marginal felicity (i.e., $u^i_{ij} > 0$) could be referred to as a "bandwagon" effect since agent i's preference to consume a commodity is enhanced by an increase in agent j's contemporaneous consumption. This externality could be taken as reflecting the mob-psychological inclination of consumers to imitate or follow the other agents' behavior. In contrast, when $u^i_{ij} < 0$, the externality induces substitutability of c_j for c_j in the Edgeworth sense.

These consumption externalities can take place for various reasons, as exemplified as follows:

Example 1 (the "keeping-up-with-Joneses" externality): Suppose that agent i is interested in his consumption relative to the social average of consumption $C = c_1^{\alpha} c_2^{1-\alpha}$: $u^i = \mu^i (c_i/C)$, which captures the "keeping-up-with-Joneses" externality. In this case, u^i_{ij} is positive.[3]

Example 2 (apparent status): As a modified case of Example 1, suppose that agent i derives felicities from pure consumption c_i and "apparent status" c_i/C : $u^i = \varsigma^i (c_i, c_i/C)$. In this case, u^i_{ij} can be either positive or negative depending on complementarity between pure consumption and "conspicuous consumption" used to demonstrate a high standard of living.[4]

Example 3 (hostility and benevolence): Externalities caused by "hostility" or "benevolence" could be considered by specifying felicity functions in recursive manners as $u^i = (c_i)^{\alpha_i} (u^j)^{\beta_i}$ $(\alpha_i > 0)$. If $| \beta_i \beta_j |< 1$, this can be solved as $u^i = (c_i)^{\sigma_i} (c_j)^{\nu_i}$, where $\sigma_i = \frac{\alpha_i}{1-\beta_i\beta_j}$ and $\nu_i = \frac{\alpha_j \beta_i}{1-\beta_i\beta_j}$. Thus, in the presence of hostility $(\beta_i < 0)$, the other agent's consumption behaves as a substitute whereas it is a complement under benevolence $(\beta_i > 0)$.[5]

---

[3] This felicity function is a version of a utility function developed by Duesenberry (1949, Chap.3). As he considered, it would be more plausible to assume that the weights $(\alpha, 1-\alpha)$ differ among agents depending on their reference groups. See also Abel (1990) and Galí (1994) in which versions of this preference are applied to asset pricing.

[4] The idea that the relative consumption (c_i/C here) is a measure of social status can be found in Duesenberry's work (Chap. 3.5). Robson (1992) considers a von Neumann-Morgenstern utility function of wealth and status defined by relative wealth holdings. It can be easily conjectured that in Example 2 the *indirect* utility is a function of consumers' own wealth holdings and the wealth distribution among consumers, as in Robson's *direct* utility function. See also Zou (1995) and Bakshi and Chen (1996) who specify felicity as functions of consumption and wealth as a status index.

[5] This is a version of interdependent utility functions developed by Becker (1974, pp.1080-1081).

Example 4 (public goods): Even without sociological interactions considered above, consumption externalities occur in the presence of public goods if the provision of public goods depends on the other agents' consumption. As the simplest example, consider a felicity function $u^i = v^i(c_i, g)$, where g is a public good, and assume that g is financed by imposing a constant rate of consumption tax τ. Then, the felicity function is given by $u^i = v^i(c_i, \tau(c_1 + c_2))$. Again, u^i_{ij} can be either positive or negative depending on substitutability between c_i and g.

Although these examples may have different economic implications, the purpose here is not to describe a certain economic phenomenon caused by a specific consumption externality, but to derive general properties resulting from consumer interdependence. Throughout this paper, therefore, instead of specifying an explicit form of consumption externality, I characterize consumption externalities in terms of complementarity-inducing externalities ($u^i_{ij} > 0$) and substitutability-inducing externalities ($u^i_{ij} < 0$).

At time zero, agent i chooses his time schedule of consumption $\{c_i(t)\}_{t=0}^{\infty}$ and nonhuman wealth holding $\{a_i(t)\}_{t=0}^{\infty}$ so as to maximize Eq. (3.2), taking as given the time paths of the endowment flow $\{w_i(t)\}_{t=0}^{\infty}$, the real interest rate $\{r(t)\}_{t=0}^{\infty}$, and the other agent's consumption schedule $\{c_j(t)\}_{t=0}^{\infty}$ $(j \neq i)$.[6] Maximization is carried out subject to the initial condition, $a_i(0) = a_{i0}$ (a constant), and the flow and intertemporal constraints,

$$\dot{a}_i(t) = r(t)a_i(t) + w_i(t) - c_i(t), \quad i = 1, 2, \tag{2.2}$$

$$\lim_{T\to\infty} a_i(T)\exp\left\{-\int_0^T r(t)\,dt\right\} = 0, \quad i = 1, 2, \tag{2.3}$$

where dots denote the time derivatives. Eq. (2.2) is self-explanatory. Eq. (2.3) represents the no-Ponzi-game condition, which prohibits the agent from rolling over his debts forever.

Define the current value Hamiltonian function as

$$H_i(t) = u^i[c_i(t); c_j(t)] + \lambda_i\{r(t)a_i(t) + w_i(t) - c_i(t)\}, \quad i = 1, 2, \tag{2.4}$$

where λ_i is the shadow price of saving. The necessary conditions are given by

$$u^i_i[c_i(t); c_j(t)] = \lambda_i(t), \quad i = 1, 2, \tag{2.5}$$

$$\dot{\lambda}_i(t)/\lambda_i(t) = \delta_i - r(t), \quad i = 1, 2, \tag{2.6}$$

where $u^i_i = \partial u^i/\partial c_i$. Conditions (2.5) and (2.6) together with (2.2) and (2.3) are also sufficient for the solution to be optimum because the Hamiltonian function H_i is concave in c_i and a_i.

Note from Eq. (2.5) that the shadow price of saving λ_i also depends on the other's consumption c_j through externalities, implying that the consumption plans for agents 1 and 2 are interdependent. This can be clarified by obtaining

[6] By the definition of externalities, agent i takes time path $\{c_j(t)\}_{t=0}^{\infty}$ as given. This amounts to assuming away any charitable behavior (e.g., donations) and predatory behavior (e.g., theft). In the present model, agents can feel benevolent and/or hostile toward their neighbors but they never control other agents' consumption by any means. Concerning this issue, see Becker (1974).

from Eqs. (2.5) and (2.6) the Keynes-Ramsey rule under consumption externalities:

$$\eta_{ii}\left[c_i(t);c_j(t)\right]\left\{\dot{c}_i(t)/c_i(t)\right\}-\eta_{ij}\left[c_i(t);c_j(t)\right]\left\{\dot{c}_j(t)/c_j(t)\right\}+\delta_i=r(t),\quad i,j=1,2\,(i\neq j),\tag{2.7}$$

where $\eta_{ii}(c_i;c_j)$ denotes the elasticity of agent i's marginal felicity with respect to his own consumption; and $\eta_{ij}(c_i;c_j)$ is the elasticity of agent i's marginal felicity with respect to agent j's consumption:

$$\eta_{ii}(c_i;c_j)=-c_iu^i_{ii}(c_i;c_j)/u^i_i(c_i;c_j)>0,\quad \eta_{ij}(c_i;c_j)=c_ju^i_{ij}(c_i;c_j)/u^i_i(c_i;c_j),\quad i,j=1,2\,(i\neq j).$$

From Eq. (2.7), the instantaneous pure marginal rate of substitution of one agent (the left hand side),[7] which depends on the other agent's consumption as well as his own, must equal the market interest rate. I focus on the Nash-equilibrium consumption plans of two agents, in which the consumption growth rates of the two agents simultaneously satisfy the two Euler equations, (2.7). The determination of the Nash-equilibrium consumption plans can be summarized by the following rule:

Proposition *(a modified Keynes-Ramsey rule): The Nash-equilibrium consumption schedules satisfy:*

$$\dot{c}_i(t)/c_i(t)=\pi_i(t)\left\{r(t)-\rho_i(t)\right\},\quad i=1,2,\tag{2.8}$$

where π_i and ρ_i represent, respectively,

$$\pi_i=\frac{\eta_{jj}(c_j;c_i)+\eta_{ij}(c_i;c_j)}{\eta_{11}(c_1;c_2)\,\eta_{22}(c_2;c_1)-\eta_{12}(c_1;c_2)\,\eta_{21}(c_2;c_1)},\quad i,j=1,2\,(i\neq j),\tag{2.9}$$

$$\rho_i=\frac{\eta_{jj}(c_j;c_i)}{\eta_{jj}(c_j;c_i)+\eta_{ij}(c_i;c_j)}\delta_i+\frac{\eta_{ij}(c_i;c_j)}{\eta_{jj}(c_j;c_i)+\eta_{ij}(c_i;c_j)}\delta_j,\quad i,j=1,2\,(i\neq j).\tag{2.10}$$

Proof: Solving the two equations in (2.7) for $\dot{c}_1(t)/c_1(t)$ and $\dot{c}_2(t)/c_2(t)$ yields Eqs. (2.8).□

In accordance with Eq. (2.8), the consumption growth rate of agent i is governed by two parameters π_i and ρ_i, both of which also depend on the other agent's subjective preferences and consumption. Parameter π_i captures the sensitivity of agent i's consumption growth rate to a change in the interest rate. I thus call parameter π_i the elasticity of intertemporal substitution induced by consumption externalities, or more simply the *induced (or socially-induced) elasticity of intertemporal substitution.* Given the sign of π_i, on the other hand,

[7] To be precise, denoting agent i's marginal rate of substitution between time t and $\tau\,(>t)$ by $M_i(t,\tau)$, the instantaneous pure marginal rate of substitution, $\mathrm{IMRS}_i(t)$, is given as: $\mathrm{IMRS}_i(t)=\lim_{\tau\to t}\partial M_i(t,\tau)/\partial\tau$. Since in this setting $\mathrm{IMRS}_i(t,\tau)=u^i_i(t)\exp\{\delta_i(\tau-t)\}/u^i_i(\tau)$ for agent i, his instantaneous pure marginal rate of substitution can be represented by the left hand side of Eq. (2.7).

the relative magnitudes of the interest rate and parameter ρ_i determine whether agent i's consumption is increasing or decreasing in t. The ρ_i is thus referred to as the rate of time preference induced by consumption externalities, or more simply, the *induced (or socially-induced) time preference rate.* For comparisons with the usual corresponding notions, I give the formal definitions of these parameters in the following.

Definition: The induced elasticities of intertemporal substitution, π_i $(i = 1, 2)$, are defined as increases of respective agents' Nash-equilibrium consumption growth rates required for the instantaneous pure marginal rates of substitution for each agent to rise uniformly by one percent. That is, (π_1, π_2) is a solution, $(\triangle(\dot{c}_1/c_1), \triangle(\dot{c}_2/c_2))$, to equation

$$\begin{pmatrix} \eta_{11} & -\eta_{12} \\ -\eta_{21} & \eta_{22} \end{pmatrix} \begin{pmatrix} \triangle(\dot{c}_1/c_1) \\ \triangle(\dot{c}_2/c_2) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Definition: The induced time preference rate, ρ_i, is defined as the instantaneous pure marginal rate of substitution in agent i's Nash-equilibrium consumption schedule (IMRS_i) evaluated at $\dot{c}_i = 0$ and $\mathrm{IMRS}_j = \mathrm{IMRS}_i$ $(j \neq i)$:[8,9]

$$\rho_i(t) = \mathrm{IMRS}_i(t)\,|_{\dot{c}_i=0;\ \mathrm{IMRS}_j=\mathrm{IMRS}_i}\,. \tag{2.11}$$

2.2.2 The Induced Elasticity of Intertemporal Substitution

In the absence of externalities, the elasticity of intertemporal substitution is given by the reciprocal of the elasticity of the marginal felicity, e.g., $1/\eta_{ii}$ for agent i. On the other hand, from Eq. (2.9), the induced elasticity of intertemporal substitution can be rewritten as

$$\pi_i = \frac{1}{\eta_{ii}}\left(1 + \frac{\eta_{ij}}{\eta_{jj}}\right)\frac{1}{1 - (\eta_{ji}/\eta_{ii})(\eta_{ij}/\eta_{jj})}, \quad i = 1, 2\,(i \neq j). \tag{2.12}$$

This implies that the value of π_i can be larger or smaller than $1/\eta_{ii}$, the no-externality case, depending on (i) whether agent i is subjected to complementarity-inducing $(\eta_{ij} > 0)$ or substitutability-inducing externalities $(\eta_{ij} < 0)$, and (ii) whether externalities are symmetric $(\eta_{ij}\eta_{ji} > 0)$ or not $(\eta_{ij}\eta_{ji} \leq 0)$. To focus

[8] The time preference rate is usually defined as the instantaneous pure marginal rate of substitution evaluated at $\dot{c}_i = 0$ [e.g., Uzawa (1969)]. Here $\mathrm{IMRS}_i(t)$ (the left hand side of Eq. (2.7)) also depends on $\dot{c}_j(t)/c_j(t)$ $(j \neq i)$. Eliminating it from $\mathrm{IMRS}_i(t)$ by using the definition of $\mathrm{IMRS}_j(t)$, $\mathrm{IMRS}_i(t)$ can be rewritten as dependent upon $\mathrm{IMRS}_j(t)$ as well as $\dot{c}_i(t)/c_i(t)$. In definition (2.11), $\mathrm{IMRS}_j(t)$ is set equal to $\mathrm{IMRS}_i(t)$ by taking into account the fact that they are equalized in equilibrium.

[9] As in Obstfeld (1990), definition (2.11) can be rewritten in terms of the Volterra derivative, $D_V[U_i(0), c_i(t)] = u_i^i(t)\exp(-\delta_i t)$, as:

$$\rho_i = -\frac{d\ln D_V[U_i(0), c_i(t)]}{dt}\,\Big|_{\dot{c}_i=0;\ \frac{d\ln D_V[U_j(0),c_j(t)]}{dt} = \frac{d\ln D_V[U_i(0),c_i(t)]}{dt}}\,.$$

on (i), suppose that agent j is not subjected to any externalities, $\eta_{ji} = 0$. Then, Eq. (2.12) reduces to

$$\pi_i = \frac{1}{\eta_{ii}}\left(1 + \frac{\eta_{ij}}{\eta_{jj}}\right) \gtrless \frac{1}{\eta_{ii}} \text{ as } \eta_{ij} \gtrless 0.$$

That is, complementarity-inducing (resp. substitutability-inducing) externalities to agent i raise (resp. lower) his elasticity of intertemporal substitution. This is because a one percent increase in r raises the optimal consumption growth rate for agent j by $1/\eta_{jj}$, which *ceteris paribus* lowers the instantaneous pure marginal rate of substitution of agent i by η_{ij}/η_{jj}, so that the consumption growth rate for agent i must rise by $(1/\eta_{ii})\{1 + (\eta_{ij}/\eta_{jj})\}$ to satisfy the Euler equation, (2.7). As for the second factor (ii), asymmetric externalities ($\eta_{ij}\eta_{ji} < 0$) weaken intertemporal substitution, whereas symmetric externalities ($\eta_{ij}\eta_{ji} > 0$) promote it insofar as $\pi_i > 0$.

It is noteworthy that, from Eq. (2.9) or (2.12), the induced elasticity of intertemporal substitution can be negative, implying that the felicity function can be socially convex. In this pathological case, the agent is, say, *socially addicted* to the present consumption (or the present living standard) in the sense that an increase in the present consumption raises the pure marginal rate of substitution of the present for the next instant consumption due to social interactions. From Eq. (2.9) or (2.12), the social addiction occurs, for example, under strong mutual Edgeworth-complementarities (i.e., strong mutual bandwagon externalities). In that case, an increase in the present consumption of agent i, *ceteris paribus*, increases the other agent's present consumption. This in turn raises the pure marginal rate of substitution of the present for the next instant consumption for agent i, despite the increase of his present consumption.

Two comments are in order regarding this increasing-return property. First, social addiction is similar in phenomenon to, but substantially different in mechanism from the addiction discussed by Becker and Murphy (1988). In their model, addiction occurs through an intertemporal habit formation process, whereas it is generated here by consumer interdependence. Becker and Murphy show that addiction occurs under "adjacent complementarity" between consumption and habit. In the present setting, the Edgeworth-complementarity between two agents' consumption rates plays a similar role.

Second, the possibility of a negative π_i might be consistent with many empirical studies which fail to obtain significantly positive estimates of the elasticity of intertemporal substitution [Giovannini (1985), Mankiw, Rotemberg, and Summers (1985), Hall (1988), Browning (1989)]. For example, Hall (1988) obtains a negative value of the elasticity from US. annual data, commenting that his finding cannot be taken literally since it implies nonconcave utility [Hall (1988), p.353]. Mankiw, Rotemberg, and Summers (1985) provide "disappointing" estimates of the Euler equation which support nonconcave utility functions.[10] As a theoretical possibility, however, a felicity function can be socially or seemingly convex under consumption externalities even though it is concave from

[10] Although Mankiw, Rotemberg, and Summers (1985) [and Kugler (1988) cited below] incorporate leisure, it is possible to apply the present idea to the felicity function containing leisure so as to derive a socially convex felicity function of consumption and leisure.

the subjective viewpoint.[11,12] This interpretation is also consistent with Kugler (1988)'s empirical finding that a cointegration approach which is robust with respect to stationary taste shocks results in a significantly positive estimate of the elasticity of intertemporal substitution.

A similar implication of a negative π_i can be derived by examining the interest rate elasticity of the saving propensity out of the permanent income. For the ease of exposition, assume that all the η_{ij} and r are constant. Then, from Eqs. (2.3), (2.2), and (2.8), I can derive the saving propensity from the permanent income, $y_i^P(t) = r\left\{a_i(t) + \int_0^\infty w_i(t+\tau)\exp(-r\tau)d\tau\right\}$, as $s_i(r) = \pi_i \cdot (1-\rho_i/r)$.[13] It follows that:

$$s_i'(r) \gtreqless 0 \Longleftrightarrow \pi_i \gtreqless 0, \quad i = 1, 2. \tag{2.13}$$

Thus, in contrast to the usual case [e.g., Takayama (1985)], the marginal propensity to save can be negative. Indeed, under social addiction ($\pi_i < 0$), an increase in the real interest rate lowers his saving propensity. This possibility could explain insignificant and/or unstable estimates of the empirical relationship between the saving propensity and the interest rate which are reported in recent studies [e.g., Giovannini (1985) and Gylfason (1993)].

2.2.3 The Induced Rate of Time Preference

From Eq. (2.10), the induced rate of time preference of one agent, ρ_i, is a weighted average of his own subjective discount rate δ_i and the other agent's δ_j, where the weight applied to δ_j is given by the relative magnitude of the externality effect of agent j's consumption. Roughly speaking, a complementarity-inducing externality draws the induced time preference rate of one agent toward the other agent's subjective discount rate whereas a substitutability-inducing one magnifies the discrepancy between the two.

A distinct property embedded in Eq. (2.10) is that, when the elasticities of marginal felicities are variable, induced time preference rates are endogenously determined by the two agents' consumption rates. In particular, the induced time preference rate can be either positively or negatively related to each agent's consumption depending on how the elasticities of marginal felicities η_{ij} rely on the two agents' consumption rates.

Using this property, the present model can mimic, or give a micro foundation to, various forms of time preference schedules proposed by the existing literature. To give examples below, suppose that η_{22} is constant.

---

[11] This explanation for apparent non-concavity might be somewhat limited because convexity in a "composite consumption" of c_1 and c_2 is directly incorporated to explain social addiction. [e.g., consider the generalized Cobb-Douglas felicity function, $u^i(c_1, c_2) = \left(c_i^\alpha c_j^{1-\alpha}\right)^\gamma$, in which social addiction occurs if and only if $\gamma > 1$, i.e., u^i is convex in $c_i^\alpha c_j^{1-\alpha}$.] There are other explanations which propose, for example, intertemporal non-separability [e.g., Eichenbaum, Hansen, Singleton (1988)].

[12] This structure is similar to that of endogenous growth models which ensure well-defined competitive equilibria with increasing returns by assuming production externalities [e.g., Romer (1986)]. The same logic can be found in Robson (1993) who derives the "concave-convex-concave" utility of the Friedman-Savage type.

[13] Here $r - \pi_i \cdot (r - \rho_i)$ is assumed strictly positive to ensure that the present value of the lifetime consumption stream is bounded.

Example 5 (increasing marginal impatience): If agent 1 is subjected to an increasing bandwagon externality: $\eta_{12}(c_1; c_2) > 0$ and $\partial\eta_{12}(c_1; c_2)/\partial c_1 > 0$, then, from Eq. (2.10), ρ_1 displays an "increasing marginal impatience," $\partial\rho_1/\partial c_1 > 0$, as is assumed by Uzawa (1968) and Epstein and Hynes (1983).

Example 6 (decreasing marginal impatience): If agent 1 is subjected to a decreasing bandwagon externality: $\eta_{12}(c_1; c_2) > 0$ and $\partial\eta_{12}(c_1; c_2)/\partial c_1 < 0$, then, ρ_1 displays a "decreasing marginal impatience," $\partial\rho_1/\partial c_1 < 0$, as proposed by Fisher (1907).

Example 7 (nonmonotonic impatience): Assume that agent 1 is subjected to a bandwagon externality: $\eta_{12}(c_1; c_2) > 0$, and that there exists some positive constant k, such that $\partial\eta_{12}(c_1; c_2)/\partial c \gtrless 0$ as $c_1 \gtrless k$. Then, ρ_1 displays a U-shaped curve as proposed by Fukao and Hamada (1991).

The present model differs from the usual endogenous time preference models in two points. First, the induced time preference rate can depend on the other agent's consumption. This is because the extent to which an agent is subjected to consumption externalities depends on how much the other agent consumes. Second, an attractive property of endogenous time preference models is that it allows for a steady state in which both c_1 and c_2 are strictly positive, regardless of differences in the degree of impatience. In contrast, the present model still retains the property that there is no such steady state, which is an inherent property of constant subjective time preference models [indeed, Eq. (2.7) implies that, insofar as $c_i > 0$, any levels of the interest rate cannot satisfy the steady state condition, $\dot{c}_1 = \dot{c}_2 = 0$].

2.3 Equilibrium Dynamics: A Simple Case

Let us examine the equilibrium dynamics by incorporating the market equilibrium condition with fixed commodity supply, $w_i(t) = y_i$ (constant):

$$c_1(t) + c_2(t) = y \ \ (= y_1 + y_2). \tag{2.14}$$

Here I focus on a simple case with constant and symmetric elasticities: $\eta_{11} = \eta_{22}$ and $\eta_{12} = \eta_{21}$ where η_{ij} are all constant. It will turn out that this simple model is useful to shed light on several typical properties arising from consumer interdependence.

The induced substitution elasticities and the induced time preference rates for two agents are now given by

$$\pi_1 = \pi_2 = \eta_{11} - \eta_{12}, \tag{2.15}$$

$$\rho_1 = \frac{\eta_{11}}{\eta_{11}+\eta_{12}}\delta_1 + \frac{\eta_{12}}{\eta_{11}+\eta_{12}}\delta_2, \text{ and } \rho_2 = \frac{\eta_{11}}{\eta_{11}+\eta_{12}}\delta_2 + \frac{\eta_{12}}{\eta_{11}+\eta_{12}}\delta_1, \tag{2.16}$$

respectively. Thus, social addiction occurs if and only if $\eta_{11} < \eta_{12}$. The relation between induced time preference rates ρ_i and externality elasticity η_{12} is depicted by Fig. 1. As is shown in the figure, for the induced time preference rates to be positive, parameter η_{12} must satisfy

$$\eta_{12} > -\eta_{11}(\delta_1/\delta_2) \text{ or } \eta_{12} < -\eta_{11}(\delta_2/\delta_1). \tag{2.17}$$

In this section, this condition is assumed to be valid.[14]

The equilibrium interest rate can be derived by logarithmically differentiating Eq. (2.14) by t and substituting Eq. (2.8) into the result as

$$r(t) = \{c_1(t)/y\}\rho_1 + \{c_2(t)/y\}\rho_2. \tag{2.18}$$

That is, $r(t)$ is determined as a weighted average of ρ_1 and ρ_2 where the weights are given by the consumption shares of respective agents. To analyze the explicit dynamics of $r(t)$, substitute Eq. (2.14) into (2.18), differentiate logarithmically the result by t, and finally substitute Eqs. (2.8) into the equation. Then, the following autonomous equation is obtained:

$$\dot{r}(t) = \{r(t) - \rho_1\}\{r(t) - \rho_2\}/(\eta_{11} - \eta_{12}). \tag{2.19}$$

As for the equilibrium consumption, differentiating Eq. (2.14) by t and substituting successively Eqs. (2.8) and (2.14) into the result yields, if $\rho_1 \neq \rho_2$,

$$c_1(t) = \{\rho_2 - r(t)\}y/(\rho_2 - \rho_1), \tag{2.20}$$

$$c_2(t) = \{r(t) - \rho_1\}y/(\rho_2 - \rho_1). \tag{2.21}$$

Given the initial interest rate, $r(0)$, the entire time path of (r, c_1, c_2) is determined by Eqs. (2.20), (2.21), and (2.19). Given this, in turn, the time path of a_i is decided by Eq. (2.2) under given initial value a_{i0}. From this recursive structure, the perfect foresight equilibrium dynamics are determined such that $r(0)$ generates sequence $\{a_i(t)\}_{t=0}^{\infty}$ which satisfies the no-Ponzi-game condition, (2.3). As demonstrated below, if this perfect foresight dynamic equilibrium exists, it is unique.

As conjectured from Eq. (2.8), the equilibrium dynamics depend on (i) $\eta_{12} \gtrless 0$, (ii) $\pi_1 (= \pi_2) \gtrless 0$, and (iii) $\rho_1 \gtrless \rho_2$. Given this observation, in order to describe fully the equilibrium dynamics, I must consider the four following cases:

- case(a): $\eta_{12} < -\eta_{11}(\delta_2/\delta_1)$,
- case(b): $-\eta_{11}(\delta_1/\delta_2) < \eta_{12} < 0$,
- case(c): $0 < \eta_{12} < \eta_{11}$,
- case(d): $\eta_{12} > \eta_{11}$,

where each interval for η_{12} is depicted in Fig. 1. As the case changes from (a) to (d), externality elasticity η_{12} becomes larger: e.g., case (a) represents the most substitutability-inducing externality whereas (d) the most complementarity-inducing. The property of each case is summarized by Table 1 from the viewpoints of (i) through (iii) above.

	(i) η_{12}	(ii) $\pi_1 (= \pi_2)$	(iii) ρ_i and δ_i
Case (a)	$-$	$+$	$\rho_2 < \delta_1 < \delta_2 < \rho_1$
Case (b)	$-$	$+$	$\rho_1 < \delta_1 < \delta_2 < \rho_2$
Case (c)	$+$	$+$	$\delta_1 < \rho_1 < \rho_2 < \delta_2$
Case (d)	$+$	$-$	$\delta_1 < \rho_2 < \rho_1 < \delta_2$

Table 1: Simple cases

[14] Kocherlakota (1990) proves that in growing economies negative utility-discounting can be consistent with well-defined competitive equilibria. This proposition does not apply to the present model since growth is assumed away here.

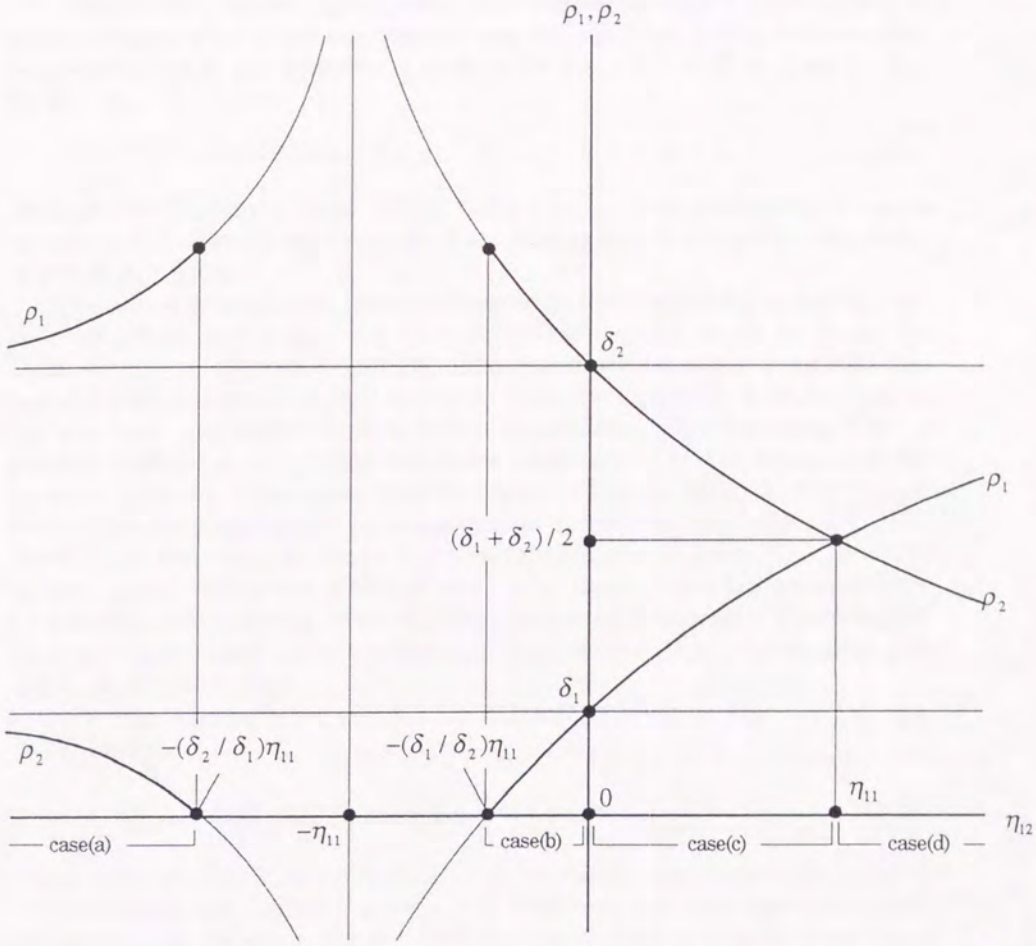


Figure 1. The induced time preference rates: $\eta_{11} = \eta_{22}, \eta_{12} = \eta_{21}$.

The equilibrium dynamics in cases (a) - (d) are depicted in Figs. 2(a) - 2(d), respectively. In each figure, the upper diagram illustrates the interest rate dynamics generated by the quadratic differential equation, (2.19); and the lower ones depict the (r, a_1) dynamics generated by Eq. (2.2) with c_1 given by Eq. (2.20), i.e.,

$$\dot{a}_1(t) = r(t)\, a_1(t) + y_1 - \{\rho_2 - r(t)\}\, y / (\rho_2 - \rho_1). \qquad (2.22)$$

As shown by the figures, when $-y_1/\rho_2 < a_{10} < y_2/\rho_1$, the equilibrium time-path is determined uniquely by the initial asset holding a_{10} on the saddle trajectory in the (r, a_1) space.

The effects of consumer interdependence on the equilibrium dynamics can be summarized as follows: First, the equilibrium interest rate is no longer determined by the subjective discount rates δ_i and can be either lower than the lowest subjective discount rate or higher than the highest one, depending on the sign and magnitude of consumption externalities. For example, when a substitutability-inducing externality takes place ($\eta_{12} < 0$), the equilibrium interest rate can be either lower than the lowest subjective discount rate (δ_1), or higher than the highest one (δ_2) depending on a_1 [see Figs. 2(a) and 2(b)]. Particularly, in this case, the long-run interest rate is always lower than the most patient agent's subjective discount rate. In contrast, when the externality is complementarity-inducing ($\eta_{12} > 0$), the long-run interest rate is always higher than the most patient agent's subjective discount rate, the no-externality case [see Figs. 2(c) and 2(d)].

Secondly, since $\min(\rho_1, \rho_2) < r(t) < \max(\rho_1, \rho_2)$ from Eqs. (2.16), Eq. (2.18) implies

$$\dot{r}(t) \underset{<}{\overset{>}{=}} 0 \Longleftrightarrow \eta_{11} \underset{>}{\overset{<}{=}} \eta_{12} \left(\Longleftrightarrow \pi_1 (= \pi_2) \underset{>}{\overset{<}{=}} 0 \right). \qquad (2.23)$$

Thus, r monotonically rises over time when the two agents are socially addicted to the present consumption [i.e., case (d)]. This contrasts with the independent preference case, in which the real interest rate declines over time toward the lowest discount rate (δ_1, here).[15] Under social addiction, as agent 1 gradually increases consumption, his preference for the present consumption escalates over time, leading the interest rate to monotonically rise over time. This intuition can be put otherwise by considering a security which yields one unit of the commodity at each instant. Letting q denote the price of the security, the no-arbitrage condition is given by: $r = (\dot{q} + 1)/q$, or equivalently $\dot{q} = rq - 1$. As can easily be shown, the equilibrium relationship between r and q can be represented by a negatively-sloped saddle trajectory defined in the (r, q) space. Under social addiction, as r rises over time, q declines monotonically from the initial highest level, meaning that the present value of the future consumption stream in terms of the present consumption falls down over time.

Thirdly, strong mutual substitutability-inducing externalities reverse the order of impatience and of the long-run wealth holdings among consumers. As seen from Table 1 and Fig. 2(a), when case (a) is true, agent 2 (i.e., the less patient agent from the viewpoint of the subjective discount rate) is effectively more patient than agent 1 due to strong substitutability-inducing externalities. Agent 2 thus eventually occupies all the available wealth and consumes all the

[15] See Frenkel and Razin (1985).

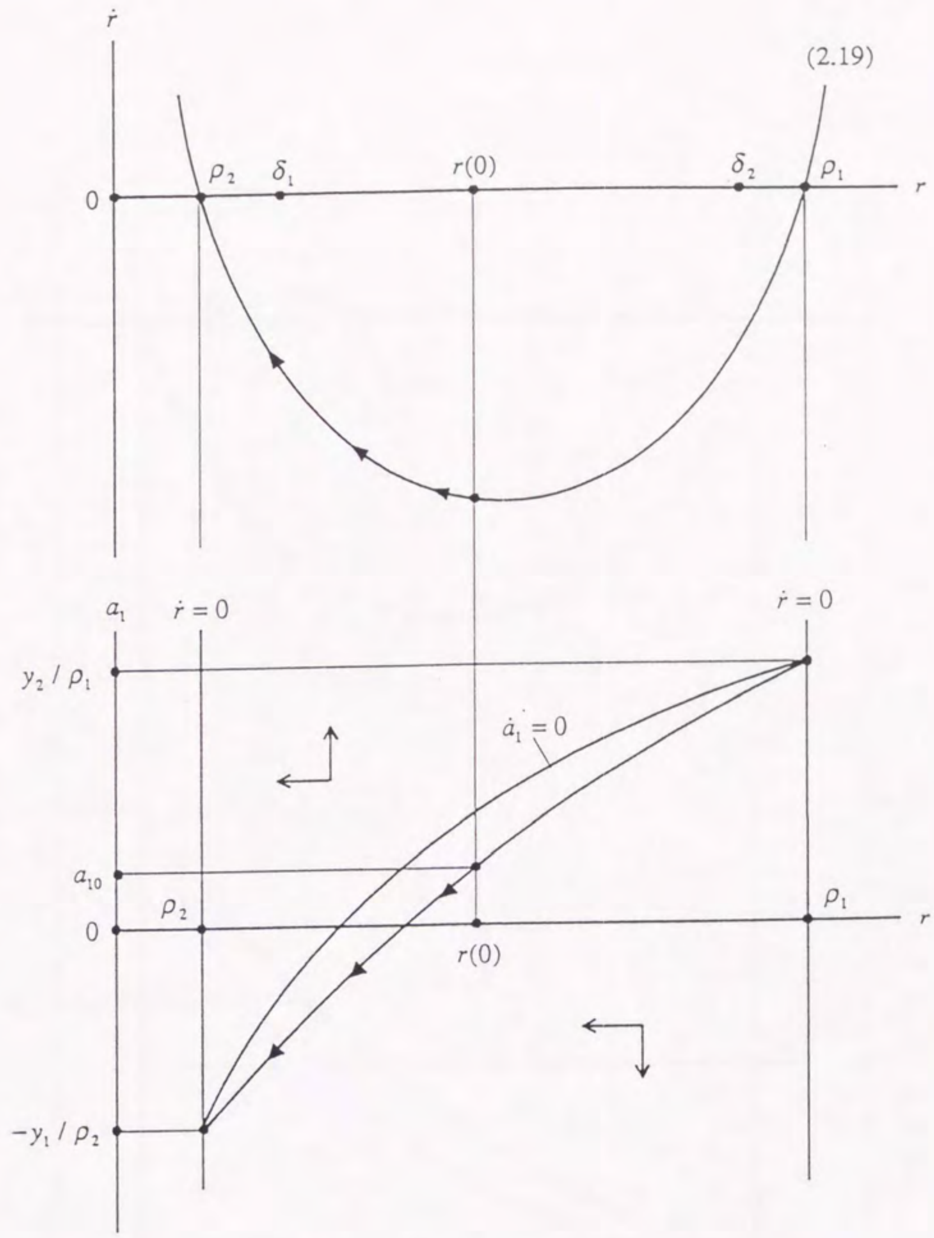


Figure 2(a). Symmetric elasticities: case (a), $\eta_{12} < -(\delta_2 / \delta_1)\eta_{11}$.

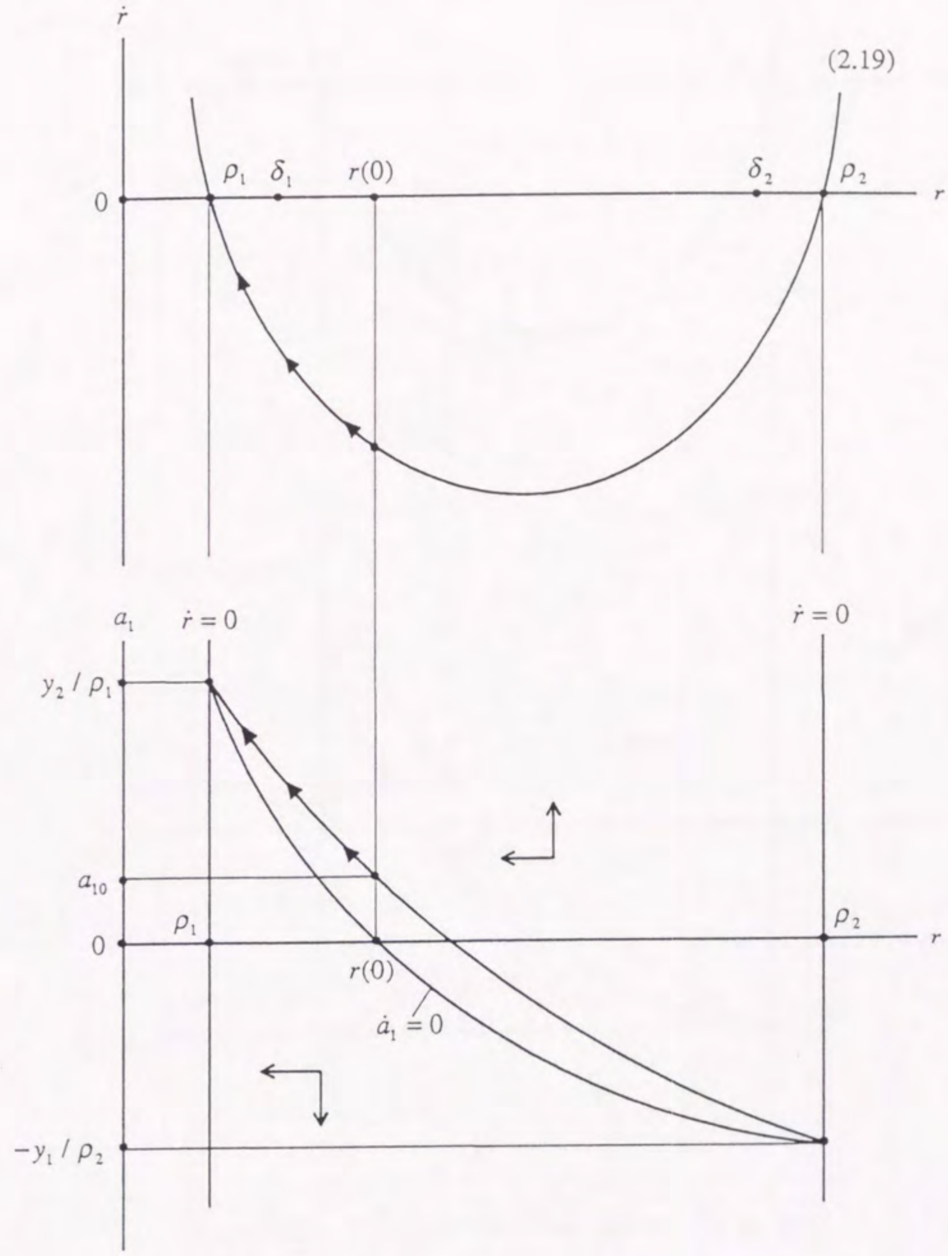


Figure 2(b). Symmetric elasticities: case (b), $-(\delta_1 / \delta_2)\eta_{11} < \eta_{12} < 0$.

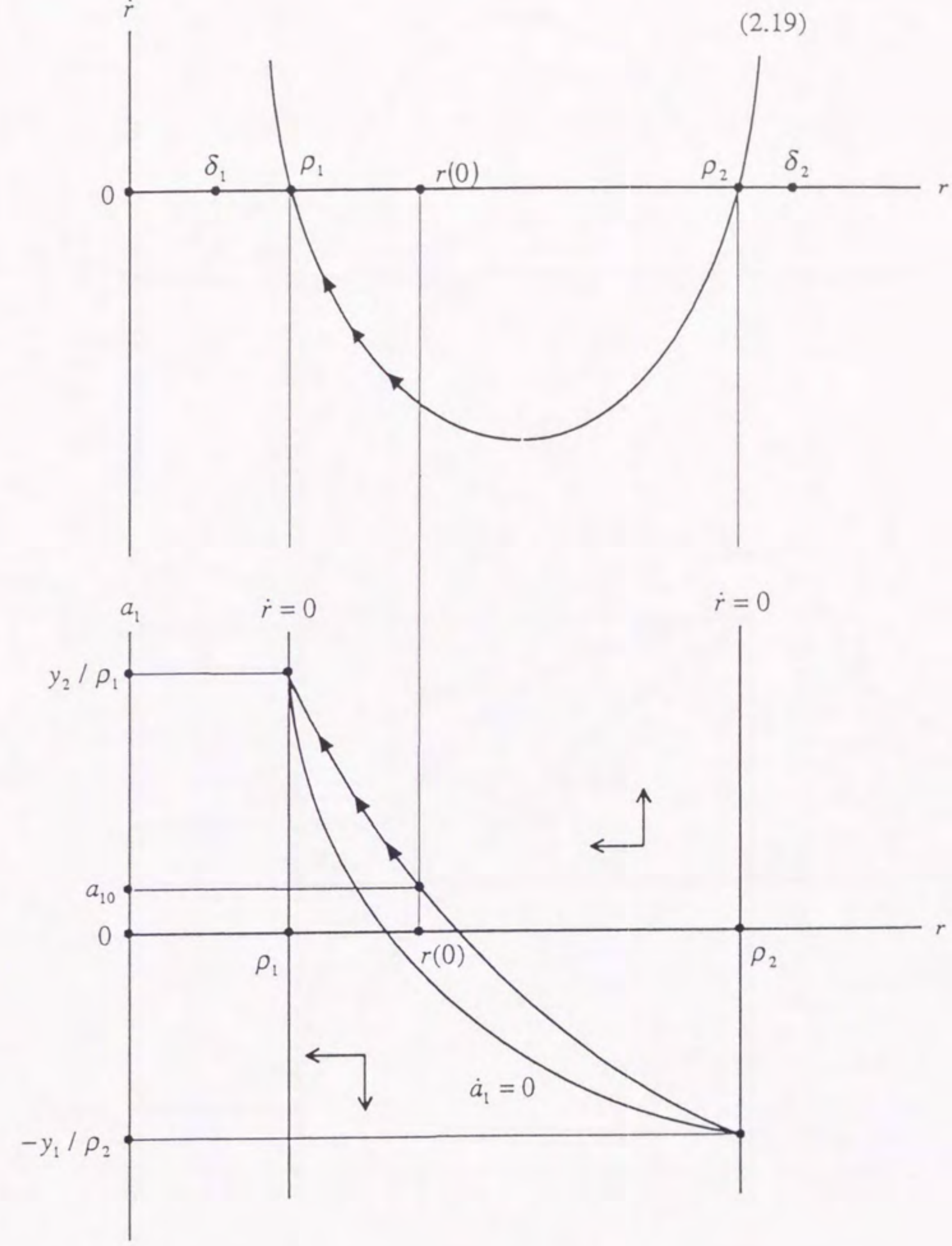


Figure 2(c). Symmetric elasticities: case (c), $0 < \eta_{12} < \eta_{11}$.

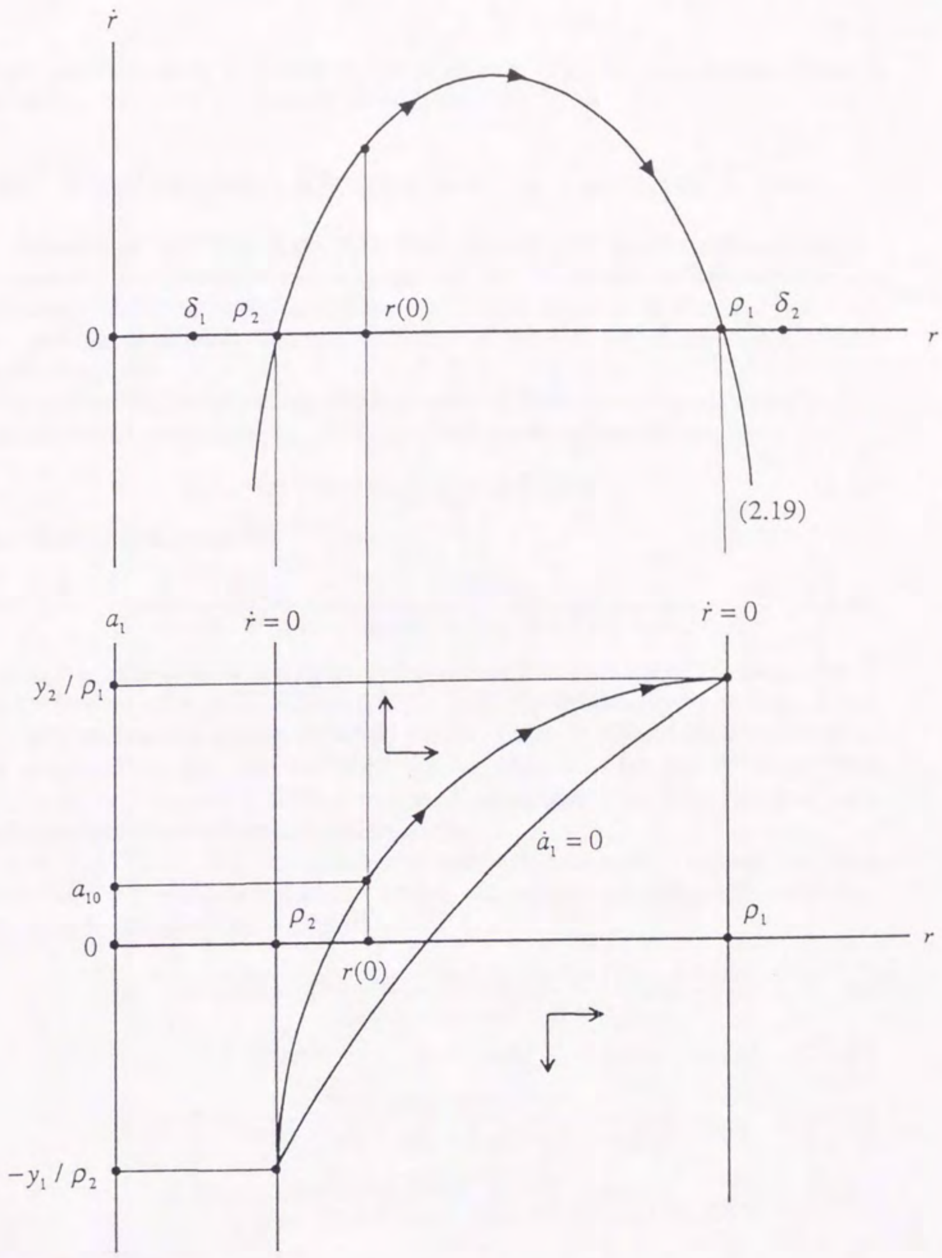


Figure 2(d). Symmetric elasticities: case (d), $\eta_{12} > \eta_{11}$.

output whereas agent 1 shrinks in the long-run. The long-run interest rate is determined by agent 2's induced time preference.[16]

2.4 Equilibrium Dynamics: A General Case

Let us finally extend the model to a more general case in which elasticities η_{ij} are constant but possibly asymmetric. As will be shown below, asymmetric externalities bring several possibilities which cannot occur in the previous case, e.g., multiple equilibria, the non-existence of equilibrium under equal wealth distributions, etc.

To derive the equilibrium interest rate, differentiate logarithmically Eq. (2.14) by t and substitute Eq. (2.8) into the result. I then obtain

$$r(t) = \omega(t)\rho_1 + \{1 - \omega(t)\}\rho_2, \tag{2.24}$$

where weight ω is given by

$$\omega = \frac{c_1\pi_1}{c_1\pi_1 + c_2\pi_2} = \frac{c_1(\eta_{22} + \eta_{12})}{c_1(\eta_{22} + \eta_{12}) + c_2(\eta_{11} + \eta_{21})}. \tag{2.25}$$

That is, the interest rate is determined as a weighted average of the two induced time preference rates as in Section 2.3, but here the weights, given by Eq. (2.25), are the consumption shares adjusted by the induced substitution elasticities. The interest rate can thus be either higher than $\max(\rho_1, \rho_2)$ or lower than $\min(\rho_1, \rho_2)$ when social addiction occurs. In that case, therefore, the positivity condition for r must be considered carefully.

From Eq. (2.24), the autonomous dynamic equation with respect to r and the equilibrium relations between r and c_i can be derived using the same procedures as in the previous section:[17]

$$\begin{aligned} \dot{r}(t) &= \frac{(\eta_{22} + \eta_{12} - \eta_{11} - \eta_{21})(\eta_{11} + \eta_{21})(\eta_{22} + \eta_{12})}{(\eta_{11}\eta_{22} - \eta_{12}\eta_{21})^2(\delta_2 - \delta_1)} \\ &\quad \times (r(t) - \rho_1)(r(t) - \rho_{12})(r(t) - \rho_2), \end{aligned} \tag{2.26}$$

$$c_1(t) = \frac{(\eta_{11} + \eta_{21})(r(t) - \rho_2)}{(\eta_{11} + \eta_{21} - \eta_{22} - \eta_{12})(r(t) - \rho_{12})} y, \tag{2.27}$$

$$c_2(t) = \frac{(\eta_{22} + \eta_{12})(r(t) - \rho_1)}{(\eta_{22} + \eta_{12} - \eta_{11} - \eta_{21})(r(t) - \rho_{12})} y, \tag{2.28}$$

where ρ_{12} represents

$$\rho_{12} = \left(\frac{\eta_{22} + \eta_{12}}{\eta_{22} + \eta_{12} - \eta_{11} - \eta_{21}}\right)\rho_1 + \left(\frac{-\eta_{11} - \eta_{21}}{\eta_{22} + \eta_{12} - \eta_{11} - \eta_{21}}\right)\rho_2. \tag{2.29}$$

To characterize these equilibrium dynamics, I reduce the consumption dynamics by substituting successively Eqs. (2.24), (2.25), (2.9), (2.10), and (2.14) into (2.8) as

---

[16] Case (d) represents a similar case in which ρ_2 is lower than ρ_1. However, in this case, social addiction occurs (see Table 1), so that agent 1 accumulates wealth to be eventually the dominant consumer.

[17] Eq. (2.26) is derived by substituting Eq. (2.14) into (2.24), differentiating logarithmically the result by t, and finally substituting (2.8). Eqs. (2.27) and (2.28) can be obtained by differentiating Eq. (2.14) by t and substituting Eq. (2.8) into the result.

$$\dot{c}_1(t) = \frac{c_1(t)\, c_2(t)\,(\delta_2 - \delta_1)}{\Omega\left[c_1(t), c_2(t)\right]}, \tag{2.30}$$

where

$$\Omega(c_1, c_2) = c_1(\eta_{22} + \eta_{12}) + c_2(\eta_{11} + \eta_{21}). \tag{2.31}$$

The steady state is thus given by $(c_1^*, c_2^*) = (y,\ 0)$ and $(0,\ y)$. As pointed out in Section 2.2, one of the agents finally becomes the dominant one who consumes all.

Eq. (2.30) reveals that the equilibrium consumption dynamics depend crucially on the sign of Ω, which in turn, from Eq. (3.11), depends on the signs of $\eta_{11} + \eta_{21}$ and $\eta_{22} + \eta_{12}$. Incorporating the positivity conditions for ρ_i into four possible combinations of $(\mathrm{sign}(\eta_{11} + \eta_{21}), \mathrm{sign}(\eta_{22} + \eta_{12}))$, I can distinguish the following cases:[18]

- case 1: $\eta_{11} + \eta_{21} > 0$ and $\eta_{22} + (\delta_2/\delta_1)\,\eta_{12} > 0$,
- case 2: $\eta_{11} + (\delta_1/\delta_2)\,\eta_{21} < 0$ and $\eta_{22} + \eta_{12} < 0$,
- case 3: $\eta_{11} + (\delta_1/\delta_2)\,\eta_{21} < 0$ and $\eta_{22} + (\delta_2/\delta_1)\,\eta_{12} > 0$,
- case 4: $\eta_{11} + \eta_{21} > 0$ and $\eta_{22} + \eta_{12} < 0$.

Case 1 represents the case in which $\Omega(c_1, y - c_1) > 0$ for any $c_1 \in [0,\ y]$. Case 2 is the opposite case in which $\Omega(c_1, y - c_1) < 0$ for any $c_1 \in [0,\ y]$. In cases 3 and 4, the sign of $\Omega(c_1, y - c_1)$ depends on the magnitude of c_1. As I shall show by turns, case 3 displays richer dynamics whereas in case 4 there is no equilibrium.

2.4.1 Case 1

Case 1 represents the case in which both agents are subjected to either a complementarity-inducing externality or a substitutability-inducing but sufficiently weak externality ($\eta_{21} > -\eta_{11}$, $\eta_{12} > -(\delta_1/\delta_2)\,\eta_{22}$). Since $\Omega > 0$, from Eq. (2.30), agent 1 monotonically increases consumption over time, and hence becomes the dominant consumer in the long run. Agent 2 shrinks over time. From Eq. (2.24), the interest rate approaches agent 1's induced time preference rate. Regarding the consumption and wealth dynamics, this case corresponds to cases (b) through (d) discussed in Section 2.3. [More specifically, if $\eta_{11}\eta_{22} - \eta_{12}\eta_{21} < 0$, social addiction occurs to both agents ($\pi_1, \pi_2 < 0$ from Eq. (2.9)). The resultant dynamics are the same as in case (d).]

2.4.2 Case 2

Contrarily, in case 2, two agents are subjected to sufficiently strong substitutability-inducing externalities ($\eta_{21} < -(\delta_2/\delta_1)\,\eta_{11}0$, $\eta_{12} < -\eta_{22}$). Since Ω is always negative, from Eq. (2.30), agent 2 continues to increase consumption to be finally

[18] For example, case 1 represents the case in which $\eta_{11} + \eta_{21} > 0$, $\eta_{22} + \eta_{12} > 0$, $\rho_1 > 0$, and $\rho_2 > 0$. From Eq. (2.10), the inequalities $\eta_{11} + \eta_{21} > 0$ and $\rho_2 > 0$ can be summarized as $\min(\eta_{11} + \eta_{21}, \eta_{11} + (\delta_1/\delta_2)\,\eta_{21}) > 0$; and the inequalities $\eta_{22} + \eta_{12} > 0$ and $\rho_1 > 0$ as $\min(\eta_{22} + \eta_{12}, \eta_{22} + (\delta_2/\delta_1)\,\eta_{12}) > 0$. These two inequalities reduce to the definition of case 1 in the text. The other cases are constructed in the same way.

the dominant consumer whereas agent 1 shrinks. From Eq. (2.24), the long run interest rate is given by agent 2's induced time preference. This case corresponds to case (a) in Section 2.3. [Indeed, since $\eta_{11}\eta_{22} - \eta_{12}\eta_{21} < 0$ in this case, from Eq. (2.9), social addiction does not occur to either agent ($\pi_1, \pi_2 > 0$).]

2.4.3 Case 3

In this case, an increase in c_1 has a negative external effect on agent 2's marginal felicity which is stronger than the internal effect on agent 1's ($\eta_{21} < -(\delta_2/\delta_1)\eta_{11} < -\eta_{11}$). An increase in c_2, on the contrary, exerts a positive or a negative but relatively weak effect on agent 2's marginal felicity ($\eta_{12} > -(\delta_1/\delta_2)\eta_{22} > -\eta_{22}$).

Letting $\underline{c}$ denote a value of c_1 which satisfies that $\Omega(c_1, y - c_1) = 0$, i.e.,

$$\underline{c} = \frac{(\eta_{11} + \eta_{21})\, y}{\eta_{11} + \eta_{21} - \eta_{22} - \eta_{12}},$$

it is valid that $\Omega(c_1, y - c_1) \gtreqless 0$ as $c_1 \gtreqless \underline{c}$. From Eq. (2.30), I thus obtain

$$\dot{c}_1 \gtreqless 0 \text{ as } c_1 \gtreqless \underline{c}\,. \tag{2.32}$$

This implies that there exist two "sinks of attraction," so that which steady state is attained depends on the initial consumption level: if the initial consumption for agent 1 is so large that $c_1(0) > \underline{c}$, then agent 1 continues to increase consumption over time and finally becomes the dominant consumer, as in case 1. In contrast, $c_1(0)$ which is smaller than $\underline{c}$ leads to subsequent monotonic decreases in c_1 and agent 2 consumes all the output in the long run, as in case 2.

To show typical dynamics, let me now focus on a case in which externalities are asymmetric: $\eta_{12} > 0$ (recall that $\eta_{21} < 0$ in case 3). Then, from Eqs. (2.9) and (2.10), the induced preference parameters satisfy

$$\pi_1 > 0, \pi_2 < 0, \text{ and } \rho_2 < \rho_{12} < \rho_1.$$

With these parameters, in turn, the interest rate given by Eq. (2.24) might be negative depending on the magnitude of c_1. Indeed, as proven in Appendix A.1, the interest rate is strictly positive if and only if agent 1's consumption rate satisfies:

$$c_1 > \underline{c} \text{ or } c_1 < \bar{c}, \tag{2.33}$$

where

$$\bar{c} = \frac{\{(\delta_2/\delta_1)\eta_{11} + \eta_{21}\}\, y}{(\delta_2/\delta_1)\eta_{11} + \eta_{21} - \eta_{22} - (\delta_2/\delta_1)\eta_{12}},$$

which is strictly positive in case 3.

Given these observations, Fig. 3 illustrates the equilibrium dynamics of (r, c_1, a_1). Panel (a) depicts the interest rate dynamics given by Eq. (2.26), where the coefficient of r^3 is negative in the present setting. Panel (b) illustrates relation (2.27) between c_1 and r.[19] As seen from Eq. (2.24), and indeed shown

[19] From Eq. (2.27), I can obtain $dc_1/dr = (\rho_2 - \rho_{12})\, c_1/(r - \rho_2)(r - \rho_{12})$, which is negative for $r \in (0, \rho_2) \cup (\rho_1, \infty)$ in the present case.

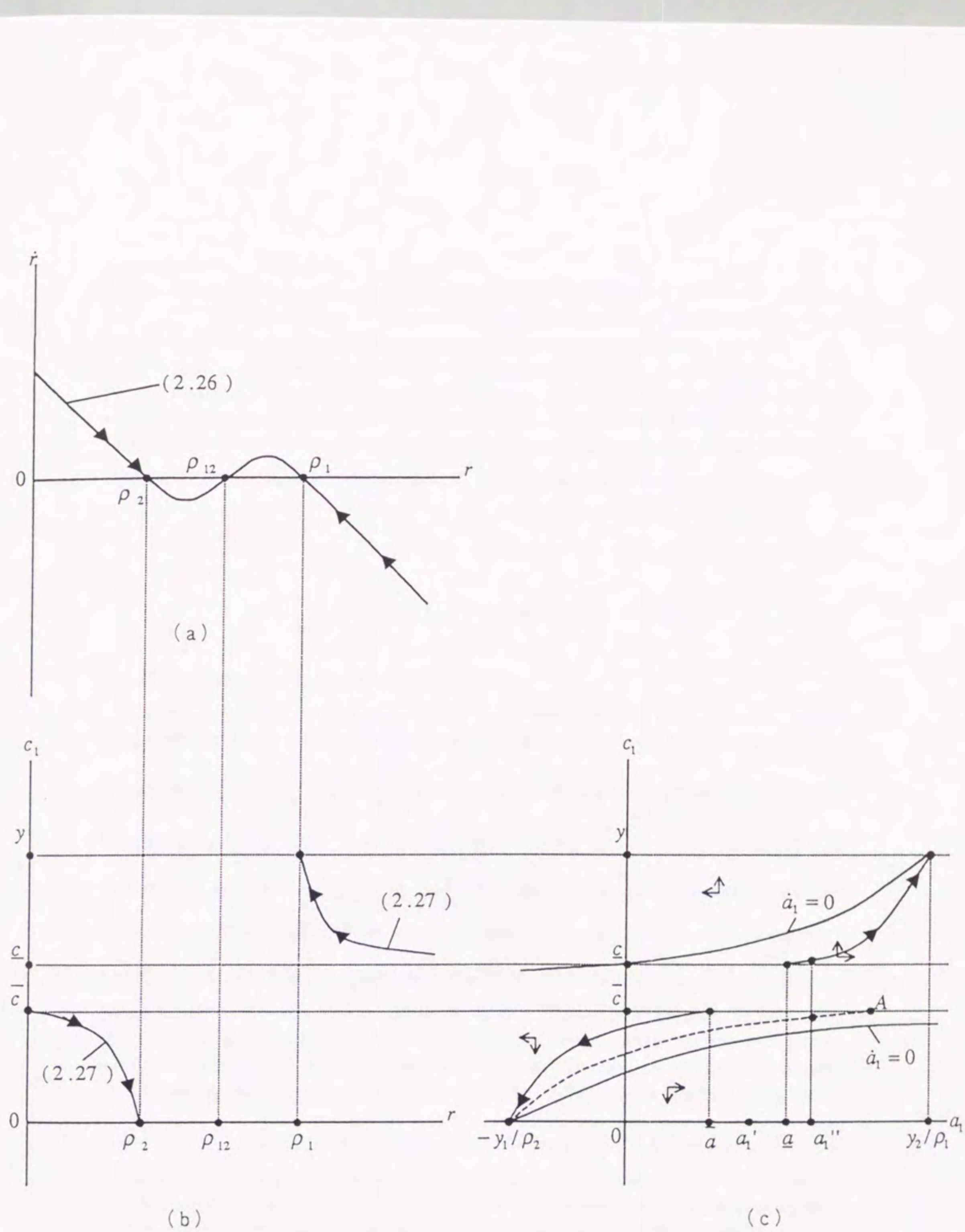


Figure 3. Equilibrium dynamics: case 3

in the panel, with positive c_1 and c_2 the interest rate satisfies either $r < \rho_2$ or $r > \rho_1$.[20] Panel (c) depicts the wealth dynamics generated by Eq. (2.2) with $w_i(t) = y_i$. The positively-sloping curves with arrows represent a saddle trajectory which is obtained from no-Ponzi-game condition (2.3) [see Appendix A.2 for derivation]. Given the initial condition, $a_1(0)$ =given, the equilibrium dynamics are determined on this trajectory. Note that the saddle trajectory is cut off between $\bar{a}$ and $\underline{a}$, where $\bar{a}$ and $\underline{a}$ represent the required asset stocks to satisfy no-Ponzi-game condition (2.3) when $c_1(0)$ approaches $\bar{c}$ and $\underline{c}$, respectively:

$$\begin{aligned}
\bar{a} &= \lim_{c_1(0)\uparrow\bar{c}} \int_0^\infty (c_1(\tau) - y_1) \exp\left(-\int_0^\tau r(s)\, ds\right) dt, \\
\underline{a} &= \lim_{c_1(0)\downarrow\underline{c}} \int_0^\infty (c_1(\tau) - y_1) \exp\left(-\int_0^\tau r(s)\, ds\right) dt.
\end{aligned}$$

Panel (c) illustrates the case in which $\underline{a} > \bar{a}$.

Regarding these typical dynamics, note the two following points. First, suppose that $\underline{a} > \bar{a}$, as depicted in panel (c). Then, there is a unique equilibrium for $a_1(0) \in \left(-\frac{y_1}{\rho_2}, \bar{a}\right) \cup \left(\underline{a}, \frac{y_2}{\rho_1}\right)$ whereas there exists no equilibrium otherwise [particularly for $a_1(0) \in \left(\bar{a}, \underline{a}\right)$, e.g., $a_1(0) = a_1'$ in panel (c)]. If agent 2 is initially so wealthy that $a_1(0) \in \left(-\frac{y_1}{\rho_2}, \bar{a}\right)$, then, as in case 2, agent 2's wealth and consumption continue to increase toward $(a_2, c_2) = \left(\frac{y_1}{\rho_2}, y\right)$ and agent 1 shrinks.[21] In contrast, if agent 1 is initially wealthy such that $a_1(0) \in \left(\underline{a}, \frac{y_2}{\rho_1}\right)$, he becomes eventually dominant as in case 1. Roughly speaking, in this case of asymmetric externalities, the initial wealth distribution determines which agent's externality is dominant and which agent becomes eventually dominant.

Secondly, when $\underline{a} < \bar{a}$ [as exemplified by the broken curve starting from point A in panel (c)], there are multiple equilibria for $a_1(0) \in \left(\underline{a}, \bar{a}\right)$, e.g., $a_1(0) = a_1''$ in panel (c). That is, if the initial asset holdings are fairly equal between agents 1 and 2, then the initial condition cannot determine which steady state is attained. Either equilibrium path is possible depending on beliefs or expectations.[22]

By definition, $\underline{a}$ and $\bar{a}$ depend on the entire time-path of (c_1, r). It is thus prohibitively difficult to examine qualitatively what determines the relative magnitudes of them. I instead give a numerical example below in which the order of the magnitudes of $\underline{a}$ and $\bar{a}$ change depending on output share y_1/y.[23]

[20] From Eq. (2.25), it is valid that $\omega(1-\omega) < 0$. Therefore, from Eq. (2.24), r takes values outside interval $[\rho_2, \rho_1]$.

[21] Although r is determined as lower than ρ_2, social addiction occurs to agent 2. From modified Keynes-Ramsey rule (2.8), he thus increases consumption and asset holdings over time.

[22] In finding the possibility of multiple equilibria, I am indebted to an anonymous referee for insightful suggestions.

[23] When $\eta_{12} < 0$, the equilibrium dynamics of r differ from those which are depicted by figure

Example 8: Set $(\eta_{11}, \eta_{12}, \eta_{22}, \eta_{21}, \delta_1, \delta_2, y) = (1, 1, 1, -6, 1, 4, 1)$, and hence $(\rho_1, \rho_{12}, \rho_2) = (0.4, 1, 2.5)$. Then, $\underline{a}$ and $\bar{a}$ can be obtained as the following functions of y_1, respectively: $\underline{a} = 0.16779 - 0.20268y_1$ and $\bar{a} = 0.22123 - 1.4778y_1$ (see Appendix A.3 for derivation). It follows that if y_1 is sufficiently large such that $y_1 \in (0.04191, 1(=y))$, $\underline{a}$ is larger than $\bar{a}$ and hence there is no equilibrium for $a_1(0) \in \left(\bar{a}, \underline{a}\right)$. If y_1 is smaller than critical value 0.04191, in contrast, $\underline{a}$ is smaller than $\bar{a}$, so that there are multiple equilibria for $a_1(0) \in \left(\underline{a}, \bar{a}\right)$.

2.4.4 Case 4

In case 4, no time-paths of the interest rate can be consistent with equilibrium. Intuitively, the reason can be demonstrated as follows. [For the formal proof, see Appendix A.4.] Case 4 can be characterized as the asymmetric case in which social addiction occurs only to the less patient agent. Since $\pi_1\pi_2 < 0$ in this case, from Eq. (2.24) the equilibrium interest rate must be higher than $\max(\rho_1, \rho_2)$ or lower than $\min(\rho_1, \rho_2)$, as in case 3. Suppose that $r(0) > \max(\rho_1, \rho_2)$. Then, by modified Keynes-Ramsey rule (2.8), the more patient agent increases the next instant consumption. This, in turn, *raises* r at that instant because, when $r > \max(\rho_1, \rho_2)$, the weight applied to the higher ρ_i in Eq. (2.24), which is larger than unity, is enlarged by an increase in the more patient agent's consumption. Through this unstable process, the more patient agent's consumption diverges and, sooner or later, exceeds the total output.

In the case of $r(0) < \min(\rho_1, \rho_2)$, the less patient agent increases consumption due to social addiction. From Eq. (2.24), it *lowers* r more. The resultant divergent process of the less patient agent's consumption necessarily violates the commodity market equilibrium condition.

2.5 Conclusion and Future Research

In this paper, I have examined the dynamic property of competitive equilibria under consumer interdependence. Key parameters in determining the optimal consumption plans of interdependent consumers are the induced elasticity of intertemporal substitution and the induced rate of time preference. Both the induced parameters of one agent depend on the other consumer's subjective preferences and consumption. The behavioral property embedded in the resulting modified Euler condition can rationalize several recent empirical studies in support of nonconcave utility functions and/or various time preference schedules. Given the modified Euler condition, the dynamic property of the Ramsey equilibrium under time preference differentials is substantially modified.

Further research is possible in various directions. First, the implication of consumption externalities for growth must be investigated. For example, there

3(a) if $\eta_{11}\eta_{22} - \eta_{12}\eta_{21}$ is strictly negative. In that case, the orders of the induced parameters are reversed such that $\pi_1 < 0$, $\pi_2 > 0$, and $\rho_1 < \rho_{12} < \rho_2$. The resultant equilibrium dynamics are those which would be obtained by interchanging agent index 1 with 2: A small (large) a_1 [and hence a small (resp. large) c_1] is associated with a high (resp. low) interest rate which is higher than ρ_2 (resp. lower than ρ_1).

is a belief or conjecture that a consumption externality caused through trade is an important factor affecting saving and growth. The present model is useful to analyze this issue. Second, the interpersonal dependence of preferences can be combined with intertemporal dependence of preferences by, for example, incorporating consumption externalities into the model of habit formation and/or endogenous time preferences. Third, the welfare implications of consumption externalities must be examined. See Ikeda (1995) for this issue. Fourth, the present analysis could be applied to examine implications of interdependent preferences toward risk for portfolio selections and asset pricing. Finally, it is necessary to test empirically the validity of interdependent consumer behavior described by the modified Keynes-Ramsey rule or the interdependent Euler equations. For example, Ikeda and Tsutsui (1996) report the empirical validity of consumer interdependence in asset pricing.

2.6 Appendix for Chapter 2

2.6.1 Appendix A.1: Derivation of (2.33)

This appendix proves that, in case 3 with $\eta_{12} > 0$, the necessary and sufficient condition for the interest rate to be strictly positive is given by Eq. (2.33).

Substituting successively Eqs. (2.25), (2.9), and (2.10) into (2.24), r is obtained as

$$r = \Gamma/\Omega,$$

where Ω is given by Eq. (3.11) and

$$\Gamma = \delta_1 \left\{ (c_1\eta_{22} + c_2\eta_{21}) + (\delta_2/\delta_1)(c_2\eta_{11} + c_1\eta_{12}) \right\}.$$

Noting that $\eta_{22} + (\delta_2/\delta_1)\eta_{12} > 0$, I thus obtain: $r > 0 \Leftrightarrow c_1/c_2$ satisfies either

$$\frac{c_1}{c_2} > \max\left[-\frac{\eta_{11}+\eta_{21}}{\eta_{22}+\eta_{12}},\ -\frac{(\delta_2/\delta_1)\eta_{11}+\eta_{21}}{\eta_{22}+(\delta_2/\delta_1)\eta_{12}}\right]$$

or

$$\frac{c_1}{c_2} < \min\left[-\frac{\eta_{11}+\eta_{21}}{\eta_{22}+\eta_{12}},\ -\frac{(\delta_2/\delta_1)\eta_{11}+\eta_{21}}{\eta_{22}+(\delta_2/\delta_1)\eta_{12}}\right].$$

Since

$$-\frac{\eta_{11}+\eta_{21}}{\eta_{22}+\eta_{12}} - \left\{-\frac{(\delta_2/\delta_1)\eta_{11}+\eta_{21}}{\eta_{22}+(\delta_2/\delta_1)\eta_{12}}\right\} = \frac{(\delta_2/\delta_1 - 1)(\eta_{11}\eta_{22} - \eta_{12}\eta_{21})}{(\eta_{22}+\eta_{12})\{\eta_{22}+(\delta_2/\delta_1)\eta_{12}\}} > 0$$

in the present setting, the above condition reduces to: $r > 0 \Leftrightarrow c_1/c_2$ satisfies either

$$\frac{c_1}{c_2} > -\frac{\eta_{11}+\eta_{21}}{\eta_{22}+\eta_{12}} \text{ or } \frac{c_1}{c_2} < -\frac{(\delta_2/\delta_1)\eta_{11}+\eta_{21}}{\eta_{22}+(\delta_2/\delta_1)\eta_{12}}.$$

Finally substituting $c_2 = y - c_1$ into these inequalities yields Eq. (2.33).□

2.6.2 Appendix A.2: Properties of the Saddle Trajectory in Case 3

This appendix derives the equilibrium saddle trajectory in case 3 of Section 2.4 (see Fig. 3). To define explicitly the equilibrium saddle trajectory, let $I(r_0)$

denote the required stock of financial assets for agent 1's consumption path [given by Eq. (2.27)] to satisfy Eq. (2.3) [together with (2.2)] when the initial interest rate is given by $r(0) = r_0$:

$$I(r_0) = \int_0^\infty (c_1(\tau) - y_1) \exp\left(-\int_0^\tau r(s)\, ds\right) dt \,|_{r(0)=r_0}\,. \tag{A1}$$

By changing integrating variables using Eq. (2.26), it can be rewritten as

$$I(r_0) = \int_{r_0}^{r_\infty(r_0)} (C_1(r) - y_1) \exp\left(-\int_{r_0}^{r} \frac{R}{f(R)} dR\right) \frac{dr}{f(r)}, \tag{A2}$$

where $r_\infty(r_0)$ represents $\lim_{t\to\infty} r(t)$ obtained from Eq. (2.26) when $r(0) = r_0$; $C_1(r)$ denotes the function on the right hand side of Eq. (2.27); and $f(r)$ represents the function to determine $\dot{r}$ in Eq. (2.26).

The equilibrium saddle trajectory in the (a_1, r) space is defined as the locus, $a_1 = I(r_0)$, which is rewritten in the (a_1, c_1) space as: $a_1 = I\left(C_1^{-1}(c_1)\right)$. Given this definition, I derive the following properties:

(P1) *Suppose that case 3 is the case and that externalities are asymmetric ($\eta_{12} > 0$). Then, for $r \in (0, \rho_2) \cup (\rho_1, \infty)$, the equilibrium saddle trajectory is strictly downward-sloping in r and hence strictly upward-sloping in c_1.*

Proof: Consider $r_0 \in (0, \rho_2)$. Then, $r_\infty(r_0) = \rho_2$ (see Fig. 3). I can obtain from Eq. (A2)

$$\frac{dI(r_0)}{dr_0}\,|_{\text{case }3} = -\frac{r_0}{f(r_0)}\left(\frac{C_1(r_0) - y_1}{r_0} - I(r_0)\right). \tag{A3}$$

Here note that for $r_0 \in (0, \rho_2)$ I have $f(r_0) > 0$ and hence $r(t)$ is increasing in t. It follows that $\frac{dI(r_0)}{dr_0}\,|_{\text{case }3} < 0\ \forall r_0 \in (0, \rho_2)$. This means that the equilibrium saddle trajectory is strictly decreasing in r for $r \in (0, \rho_2)$.

For $r_0 \in (\rho_1, \infty)$, noting that $r_\infty(r_0) = \rho_1$, the same expression as Eq. (A3) can be obtained. Since $f(r_0) < 0$ and thus $\frac{C_1(r_0) - y_1}{r_0} - I(r_0) < 0$, it is valid again that $\frac{dI(r_0)}{dr_0}\,|_{\text{case }3} < 0$ for any $r_0 \in (\rho_1, \infty)$. □

2.6.3 Appendix A.3: Computation in Example 8

Letting K represent the negative of the coefficient of r^3 in Eq. (2.26):

$$K = \frac{(\eta_{11} + \eta_{21} - \eta_{22} - \eta_{12})(\eta_{11} + \eta_{21})(\eta_{22} + \eta_{12})}{(\eta_{11}\eta_{22} - \eta_{12}\eta_{21})^2 (\delta_2 - \delta_1)},$$

the discount factor which appears in Eq. (A2) can be obtained from Eq. (2.26) as:

$$\exp\left(-\int_{r_0}^{r} \frac{R}{f(R)} dR\right) = \left|\frac{r - \rho_2}{r_0 - \rho_2}\right|^\alpha \left|\frac{r - \rho_{12}}{r_0 - \rho_{12}}\right|^\beta \left|\frac{r - \rho_1}{r_0 - \rho_1}\right|^\gamma,$$

where $\alpha = \frac{\rho_2}{K(\rho_1-\rho_2)(\rho_{12}-\rho_2)}$; $\beta = \frac{-\rho_{12}}{K(\rho_1-\rho_{12})(\rho_{12}-\rho_2)}$; $\gamma = \frac{\rho_1}{K(\rho_1-\rho_{12})(\rho_1-\rho_2)}$ (hence $\alpha+\beta+\gamma=0$). Thus, in case 3, $I(r_0)$ can be computed from Eq. (A2) as:

$$\begin{aligned}
\text{for } r_0 &\in (\rho_1, \infty), \\
I(r_0) &= L(r_0) \int_{\rho_1}^{r_0} \{C_1(r) - y_1\} (r-\rho_2)^{\alpha-1} (r-\rho_{12})^{\beta-1} (r-\rho_1)^{\gamma-1} dr;
\end{aligned}$$

$$\begin{aligned}
\text{for } r_0 &\in (0, \rho_2), \\
I(r_0) &= L(r_0) \int_{r_0}^{\rho_2} \{C_1(r) - y_1\} (\rho_2-r)^{\alpha-1} (\rho_{12}-r)^{\beta-1} (\rho_1-r)^{\gamma-1} dr,
\end{aligned}$$

where

$$L(r_0) = \frac{1}{K} \left| \frac{r_0-\rho_2}{r_0-\rho_{12}} \right|^{-\alpha} \left| \frac{r_0-\rho_1}{r_0-\rho_{12}} \right|^{-\gamma}.$$

Using these reduced forms, $\underline{a}$ and $\bar{a}$ in numerical example 8 can be obtained by computing $\underline{a} = \lim_{r_0\to\infty} I(r_0)$ and $\bar{a} = \lim_{r_0\to 0} I(r_0)$, respectively. The computation was conducted using *Maple* (Waterloo Maple Software) in *Scientific Workplace* Version 2.5 (TCI Software Research).

2.6.4 Appendix A.4: Proof of the Non-Existence of Equilibrium in Case 4

Consider first the case in which $\eta_{11}\eta_{22} - \eta_{12}\eta_{21} > 0$. Then, from Eqs. (2.10) and (2.29), I have $\pi_1 < 0$, $\pi_2 > 0$, and $\rho_2 < \rho_{12} < \rho_1$. From Eqs. (2.27) and (2.28), for $c_1(0)$ and $c_2(0)$ to be positive, it must be valid that either $r(0) > \rho_1$ or $r(0) < \rho_2$. Suppose that $r(0) > \rho_1$. Since the coefficient of r^3 in Eq. (2.26) is positive in the present case, this implies $r(t) > \rho_1$ $\forall t \geq 0$. Then, from modified Keynes-Ramsey rule (2.8), I obtain

$$\dot{c}_2(t)/c_2(t) = \pi_2(r(t) - \rho_2) > \pi_2(\rho_1 - \rho_2) = \text{a positve constant},$$

so that c_2 becomes larger than y within a finite period. It follows that $r(0) > \rho_1$ is not an equilibrium. Suppose instead that $r(0) < \rho_2$, and hence, from Eq. (2.26), that $r(t) < \rho_2$ $\forall t \geq 0$. Then, it is valid that

$$\dot{c}_1(t)/c_1(t) = \pi_1(r(t) - \rho_1) > \pi_1(\rho_2 - \rho_1) = \text{a positve constant},$$

implying that the divergence of c_1 necessarily violates market equilibrium condition (2.14) within a finite period.

When $\eta_{11}\eta_{22} - \eta_{12}\eta_{21} < 0$, I have $\pi_1 > 0$, $\pi_2 < 0$, and $\rho_1 < \rho_{12} < \rho_2$. I can apply the same argument to this case: The equilibrium interest rate must satisfy $r > \rho_2$ or $r < \rho_1$. When $r > \rho_2$, $c_1(t)$ exceeds y within a finite period. Similarly, when $r < \rho_1$, $c_2(t)$ sooner or later becomes larger than y.□

Bibliography

[1] Abel, A.B., 1990, Asset prices under habit formation and catching up with the Joneses, American Economic Review (Papers and Proceedings) 80, 38-42.

[2] Bakshi, G.S. and Z. Chen, 1996, The spirit of capitalism and stock-market prices, American Economic Review 86, 133-157.

[3] Becker, G.S., 1974, A theory of social interactions, Journal of Political Economy 82, 1063-1094.

[4] Becker, G.S. and K.M. Murphy, 1988, A theory of rational addiction, Journal of Political Economy 96, 675-700.

[5] Becker, R.A., 1980, On the long-run steady state in a simple dynamic model of equilibrium with heterogeneous households, Quarterly Journal of Economics 95, 375-382.

[6] Blanchard, O.J. and S. Fisher, 1989, Lectures on macroeconomics. The MIT Press, Cambridge.

[7] Blomquist, N. S., 1993, Interdependent behavior and the effect of taxes, Journal of Public Economics 51, 211-218.

[8] Browning, M., 1989, The intertemporal allocation of expenditure on non-durables, services, and durables, Canadian Journal of Economics 22, 22-36.

[9] Duesenberry, J.S., 1949, Income, saving, and the theory of consumer behavior. Harvard University Press, Boston.

[10] Eichenbaum, M.S., L.P. Hansen, and K.J. Singleton, 1988, A time series analysis of representative agent models of consumption and leisure choice under uncertainty, Quarterly Journal of Economics 103, 51-78.

[11] Epstein, L.G. and J.A. Hynes, 1983, The rate of time Preference and dynamic economic analysis, Journal of Political Economy 91, 611-635.

[12] Fisher, I., 1907, The Rate of Interest. Macmillan, London.

[13] Frenkel, J. and A. Razin, 1985, Government spending, debt, and international economic interdependence, Economic Journal 95, 619-639.

[14] Fukao, K. and K. Hamada, 1991, The Fisherian time preference and the evolution of capital ownership patterns in a global economy," NIRA Research Output 4, no.1, 86-97.

[15] Galí, J., 1994, Keeping up with the Joneses: consumption externalities, portfolio choice, and asset prices, Journal of Money, Credit, and Banking 26, 1-8.

[16] Giovannini, A., 1985, Saving and the real interest rate in LDCs, Journal of Development Economics 18, 195-217.

[17] Gylfason, T., 1993, Optimal saving, interest rates, and endogenous growth, Scandinavian Journal of Economics 95, 517-533.

[18] Hall, R.E., 1988, Intertemporal substitution in consumption, Journal of Political Economy 96, 339-357.

[19] Hayakawa, H. and Y. Venieris, 1977, Consumer interdependence via reference groups, Journal of Political Economy 85, 599-615.

[20] Ikeda, S., 1995, Time preference, intertemporal substitution, and dynamics under consumer interdependence, ISER Discussion Paper No.386, Osaka University.

[21] Ikeda, S. and Y. Ono, 1992, Macroeconomic dynamics in a multi-country economy: A dynamic optimization approach, International Economic Review 33, 629-644.

[22] Ikeda, S. and Y. Tsutsui, 1996, Asset pricing with consumption externalities: An empirical analysis, in: T. Tachibanaki and Y. Tsutsui (eds.), Japanese capital markets, Nippon-Hyoronsha Co., Tokyo, 29-42 (in Japanese).

[23] Kocherlakota, N., 1990, On the 'discount' factor in growth economies, Journal of Monetary Economics 25, 43-47.

[24] Kugler, P., 1988, Intertemporal substitution, taste shocks and cointegration, Economics Letters 26, 235-239.

[25] Leibenstein, H., 1950, Bandwagon, snob, and Veblen effects in the theory of consumer's demand, Quarterly Journal of Economics 64, 183-207.

[26] Mankiw, N., J. Rotemberg, and L. Summers, 1985, Intertemporal substitution in macroeconomics, Quarterly Journal of Economics 100, 225-251.

[27] Morgenstern, O., 1948, Demand theory reconsidered, Quarterly Journal of Economics 62, 165-201.

[28] Obstfeld, M., 1990, Intertemporal dependence, impatience, and dynamics, Journal of Monetary Economics 26, 45-75.

[29] Robson, A.J., 1992, Status, the distribution of wealth, private and social attitudes to risk, Econometrica 60, 837-857.

[30] Romer, P., 1986, Increasing returns and long run growth, Journal of Political Economy 94, 1002-1037.

[31] Ryder, H. E. and Heal, G. M., 1973, Optimum growth with intertemporally dependent preferences, Review of Economic Studies 40, 1-33.

[32] Shi, S. and L. Epstein, 1993, Habits and time preference, International Economic Review 34, 61-84.

[33] Takayama, A., 1985, Permanent income hypothesis: An optimal control theory approach to a classical macro thesis, Journal of Macroeconomics 7, 347-362.

[34] Uzawa, H., 1968, Time preference, the consumption function, and optimum asset holdings, in J.N. Wolfe (ed.), Capital and growth: Papers in honor of Sir John Hicks. Aldine, Chicago.

[35] Veblen, T., 1899, The theory of the leisure class, Allen & Unwin, London.

[36] Zou, H, 1995, The spirit of capitalism and savings behavior, Journal of Economic Behavior and Organization 28, 131-143.

Chapter 3

Habits, Costly Investment, and Current Account Dynamics

Abstract: Using a small country model with habit-forming consumers and costly investment, we analyze equilibrium dynamics of the economy and derive empirical and welfare implications. The model can mimic some stylized facts: (i) a temporary increase in fiscal spending always deteriorates the current account whereas a permanent increase in fiscal spending may have a weaker effect; (ii) permanent productivity shocks deteriorate the current account; and (iii) savings and investment tend to co-move upon macroeconomic shocks. Strong habit persistence causes sluggishness in welfare dynamics. Consequently, a beneficial fiscal policy may have a harmful hangover effect on the future welfare.

JEL Classification Numbers: F32, F41, E21.

Keywords: Habit, current account, investment, savings, fiscal policy, welfare.

3.1 Introduction

Being stimulated by recent developments in the current account experience, lots of empirical studies in openmacroeconomics have been conducted to report important stylized facts regarding the current account. They include that: (i) temporary increases in fiscal spending deteriorate the current account whereas permanent ones exert at most weaker negative effects on it [e.g., Ahmed (1986), Tanner (1994), and Obstfeld and Rogoff (1995)]; that (ii) an improvement in productivity have a detrimental effect on the current account [e.g., Glick and Rogoff (1992) and Elliot and Fatás (1996)]; and that (iii) savings and investment display a positive correlation in the short- and long-run [Feldstein and Horioka (1980) and Tesar (1991)]. To explain these findings, much attention in theoretical literature has been paid to the intertemporal aspects of savings, investment, and the current account [for surveys see Sen (1994) and Obstfeld and Rogoff (1995)]. However, there are still some gaps between the empirical findings and the theory.

We present a small country model in which current account dynamics are generated by consumption habits and costly investment. The purposes of this paper are: using the model, (a) to analyze equilibrium dynamics of consumption (savings), investment, and the current account, (b) to show that the resultant dynamics can mimic the above stylized facts, and (c) to derive welfare implications. To do so, we examine the effects of permanent as well as temporary changes in government spending, capital taxes, and productivity.

We incorporate habits and costly investment for both empirical and theoretical reasons. Empirically, these factors are fairly consistent with macroeconomic data. Particularly, habit-formation is often reported to be statistically significant in explaining consumers' behavior [e.g., Ferson and Constantinides (1991), Braun, Constantinides, and Ferson (1993), Naik and Moore (1996)]. Incorporating this intertemporally dependent preference would be prerequisite to model the saving behavior consistent with macroeconomic data.

As for a theoretical reason, habit formation and costly investment induce transition dynamics of the current account in response to demand and supply shocks, respectively. Although transition dynamics of savings can also arise in the overlapping-generation framework [e.g., Blanchard (1985)] and endogenous time-preference models,[1] these alternatives are at most empirically controversial for specifying the saving behavior.[2] We indeed find the habit-formation model more useful than these alternatives to explain the above stylized facts.

Furthermore, this specification allows us to model long-run persistent effects of temporary shocks by assuming constant discount rate equal to the world interest rate. The resultant 'zero root' property [Giavazzi and Wyplosz (1985)] is used to address effects of temporary shocks, especially stylized fact (i).

Our comparative dynamics produce the following results. (1) Temporary increases in fiscal spending always partially crowd out private consumption and deteriorate the current account. (2) Adverse productivity shocks and capital taxes improve the current account. (3) Both the non-human wealth and the capital stock tend to co-move in the short- and long-run. These results are consistent with empirical facts (i) through (iii), respectively. We also derive a welfare implication: (4) Consumers due to strong habits are present-oriented such that upon adverse shocks they sacrifice future spending and welfare in order to maintain their present standard of living. As a policy implication of this property, we show that under strong habit persistence even a beneficial fiscal policy may be accompanied by a negative hangover effect on the future welfare.

Several openmacroeconomic implications of habit formation are provided by Obstfeld (1992) and Mansoorian (1993, 1996). Our contribution differs from theirs in four points: We incorporate capital investment, consider temporary as well as permanent shocks, address empirical implications, and conduct a welfare analysis.

The remainder of the paper proceeds as follows: In Section 3.2, we present

[1] For the current account analysis using Blanchard (1985)'s overlapping-generation framework, see Matsuyama (1987) and Bovenberg (1994). For models with the Uzawa (1968)'s endogenous time-preference, see Penati (1987), Devereux and Shi (1991), and Karayalcin (1994).

[2] For example, Lawrance (1991) reported strong evidence against the Uzawa's (1968) time preference schedule. In Haug (1996)'s article, Blanchard's model of consumption is strongly rejected by quarterly Canadian data.

the model and derive equilibrium dynamics. The formulation of consumption habits is based on the seminal paper by Ryder and Heal (1973). Costly investment is formulated using the standard adjustment cost model à la Hayashi (1982). In Section 3.3, effects of macroeconomic disturbances are analyzed. Section 3.4 conducts a welfare analysis. Section 3.5 summarizes this paper with concluding remarks.

3.2 The Model

Consider a small open economy populated with infinitely-lived identical agents. There is a composite traded good that can be used for consumption and investment. The good is taken as numeraire. Let us specify the behavior of households, firms, and governments in turn, and then derive equilibrium dynamics.

3.2.1 Households

Given the market wage rate, w_t, households supply one unit of labor inelastically in each point in time t. They hold non-human wealth a_t in the form of bonds b_t and equity $q_t k_t$, where q_t denotes the equity price (Tobin's q) in terms of the consumption good; and k_t the number of equities held by domestic households. Bonds can be either purchased or issued freely at a constant interest rate, r, in the international market. By the no-arbitrage condition, the return on the equity equals r.

The preference of the households displays intertemporal dependence through habit formation. Let z_t represent the time-t *habit,* i.e., the average of past consumption rates defined by $z_t = \alpha \int_{-\infty}^{t} c_s \exp\left(-\alpha\left(t-s\right)\right) ds$, or equivalently

$$\dot{z}_t = \alpha\left(c_t - z_t\right), \tag{3.1}$$

where α represents the discount rate for past consumption rates; and c_t is the consumption rate. We specify consumers' lifetime utility as

$$U_0 = \int_0^{\infty} \{u(c_t, z_t) + v\left(g_t\right)\} \exp\left(-\theta t\right) dt, \tag{3.2}$$

where g_t represents government services. For brevity, we assume above that the felicity function is additively-separable between (c_t, z_t) and g_t. Function v is concave in g_t. We follow Ryder and Heal (1973) in assuming that function u satisfies the following regularity conditions: (C1) $u_c > 0$; (C2) $u_z \leq 0$; (C3) $u_c\left(c,c\right) + u_z\left(c,c\right) > 0$; (C4) u is concave in $\left(c,z\right)$; (C5) $\lim_{c\to 0} u_c\left(c,z\right) = \infty$ uniformly in z; and (C6) $\lim_{c\to 0}\left[u_c\left(c,c\right) + u_z\left(c,c\right)\right] = \infty$.

Intertemporal complementarities in consumption are characterized by using the following terminology coined by Ryder and Heal (1973): When $u_{cz}\left(c,c\right) + \frac{\alpha}{\theta+2\alpha} u_{zz}\left(c,c\right) <$ (resp. $>$)0, preferences are said to display *distant* (resp. *adjacent*) *complementarity,* meaning that the present consumption is complementary to the consumption in the distant (resp. adjacent) future. As shown later, the effects of macroeconomic shocks crucially depend on which complementarity takes place.

To ensure the steady state in this constant utility discounting model, let us assume

$$\theta = r. \tag{3.3}$$

This assumption will bring to the model the 'zero root' property that the steady state depends on the initial condition. As in Turnovsky and Sen (1991) and Sen and Turnovsky (1991), we will utilize this property to address the effects of temporary macroeconomic shocks.[3]

Given the initial values (b_0, k_0, z_0), consumers choose $C_0 = \{c_t, a_t\}_{t=0}^{\infty}$ so as to maximize (3.2) subject to the flow budget constraint,

$$\dot{a}_t = ra_t + w_t - c_t - x_t, \tag{3.4}$$

and the wealth constraint, $a_t = b_t + q_t k_t$. In equation (3.4) x_t denotes lump-sum tax payments to the government. Construct the Hamiltonian function with constraints (3.1) and (3.4) as:

$$H = u(c_t, z_t) + v(g_t) + \lambda_t (ra_t - c_t - x_t) + \xi_t \alpha (c_t - z_t),$$

where λ_t (≥ 0) is the shadow price of saving; and $\xi_t (\leq 0)$ is the shadow price of habit formation. The optimality conditions are given by

$$H_c = 0: \; u_c(c, z) = \lambda_t - \alpha \xi_t, \tag{3.5}$$

$$\dot{\xi}_t - \theta \xi_t = -H_z: \; \dot{\xi}_t = (\theta + \alpha) \xi_t - u_z(c_t, z_t), \tag{3.6}$$

together with (3.1), (3.4), and the transversality conditions for a_t and z_t. The resulting optimal consumption behavior will be summarized later by a saddle arm in the (z, c) space.

3.2.2 Firms

Given a costly installation technology of capital, a firm chooses time-profiles of labor demand l_t, the rate of net investment I_t, and hence capital stock k_t so as to maximize the present value of its future net cash flow. Supposed that the government levies a tax on capital stocks at constant rate $\tau \geqq 0$, the optimal behavior is described as

$$V_0 = \max \int_0^{\infty} \{\beta F(k_t, l_t) - w_t l_t - I_t(1 + \phi(I_t/k_t)) - \tau q_t k_t\} \exp(-rt) dt,$$

subject to

$$\dot{k}_t = I_t,$$

where $\beta F(k_t, l_t)$ represents the production (or revenue) function which is linearly homogeneous in k_t and l_t with productivity parameter β ; $I_t \phi(I_t/k_t)$ denotes the capital adjustment (or installation) cost function satisfying $\phi(0) = 0$, $\phi'(I_t/k_t) > 0$, and $2\phi'(I_t/k_t) + \phi''(I_t/k_t) > 0$.

[3] The zero root property may be somewhat controversial, although (3.3) is a standard setup in the literature [Obstfeld (1992) and Mansoorian (1993 and 1996)]. Obstfeld (1992) shows that this property disappears in the overlapping-generation framework. The property can be also ruled out by endogenizing utility-discounting [Shi and Epstein (1993)].

The equilibrium behavior satisfies

$$\beta F_l\left(k_t, 1\right) = w_t,$$

$$1 + \phi(I_t/k_t) + (I_t/k_t)\,\phi'(I_t/k_t) = q_t,$$

$$\dot{q}_t = (r+\tau)\,q_t - \left\{\beta F_k\left(k_t, 1\right) + (I_t/k_t)^2 \phi'(I_t/k_t)\right\},$$

together with the transversality condition for k_t, where we use the fact that the equilibrium labor supply can be normalized as unity and that the shadow price for k_t equals market stock price q_t. The optimal investment rule is summarized as $I_t = \eta\left(q_t\right) k_t$, where $\eta\left(1\right) = 0$ and $\eta'\left(q_t\right) > 0$. The equilibrium capital accumulation is thus determined by

$$\dot{k}_t = \eta\left(q_t\right) k_t, \tag{3.7}$$

$$\dot{q}_t = (r+\tau)\,q_t - \left\{\beta F_k\left(k_t, 1\right) + \eta\left(q_t\right)^2 \phi'\left(\eta\left(q_t\right)\right)\right\}. \tag{3.8}$$

Later this system will reduce to a saddle trajectory in the (q, k) space.

3.2.3 Governments

The government follows the balanced budget principle. Its fiscal spending g_t equals tax revenues from lump-sum taxes and capital taxes:

$$g_t = x_t + \tau_t q_t k_t. \tag{3.9}$$

In what follows, given exogenously government spending g_t and capital tax τ_t, lump-sum tax x_t is assumed to be determined by constraint (3.9).

3.2.4 Equilibrium Dynamics and Steady State

We focus on the dynamic system linearized in the neighborhood of a steady state. As stated earlier, however, owing to assumption (3.3) the steady state equilibrium depends on the initial condition for (b_0, k_0, z_0) and hence on the linear saddle surface. Avoiding a circular argument, we first assume that there exists a steady state and linearize the equilibrium dynamic system around the steady state. Then equilibrium saddle trajectories are obtained from the linearized system. Finally, the steady state point is chosen such that it is indeed connected with the initial point by the trajectories.

From equations (3.1), (3.5), (3.6), (3.7), and (3.8), the dynamic system for (z_t, ξ_t, k_t, q_t) linearized around a steady state is given by

$$\begin{pmatrix} \dot{\hat{z}}_t \\ \dot{\hat{\xi}}_t \\ \dot{\hat{k}}_t \\ \dot{\hat{q}}_t \end{pmatrix} = \begin{pmatrix} -\alpha\left(1 + \frac{u^*_{cz}}{u^*_{cc}}\right) & -\frac{\alpha^2}{u^*_{cc}} & 0 & 0 \\ \frac{(u^*_{cz}) - u^*_{cc} u^*_{zz}}{u^*_{cc}} & \theta + \frac{\alpha(u^*_{cc} + u^*_{zc})}{u^*_{cc}} & 0 & 0 \\ 0 & 0 & 0 & \eta'\left(1\right) k^* \\ 0 & 0 & -\beta F^*_{kk} & r+\tau \end{pmatrix} \begin{pmatrix} \hat{z}_t \\ \hat{\xi}_t \\ \hat{k}_t \\ \hat{q}_t \end{pmatrix}, \tag{3.10}$$

where $\hat{y}$ denotes deviations of variable y from its steady state value y^*: $\hat{y}_t \equiv y_t - y^*$; and the coefficient matrix is evaluated at the steady state point.

The linear dynamic system has a block structure, four characteristic roots of which are obtained separably from the two submatrices of two dimensions. For the $\left(\hat{z}, \hat{\xi}\right)$ block, the smaller root is given by

$$\omega \equiv \frac{\theta - \sqrt{(\theta+2\alpha)^2 + \frac{4\alpha(\theta+2\alpha)}{u^*_{cc}}\left(u^*_{cz} + \frac{\alpha}{\theta+2\alpha}u^*_{zz}\right)}}{2}. \tag{3.11}$$

If adjacent complementarity is too strong ($u^*_{cz} + \frac{\alpha}{\theta+2\alpha}u^*_{zz}$ is too large), ω can be positive. Our interest is, however, not in this destabilizing property of addictive habits. Following the existing literature [e.g., Obstfeld (1992) and Mansoorian (1993, 1996)], ω is assumed to be strictly negative.

Stable root ω specifies a saddle trajectory for the optimal consumption dynamics. In particular, since $\dot{\hat{z}}_t = \omega \hat{z}_t$ on the trajectory, the $\hat{c}$-$\hat{z}$ saddle trajectory is obtained by substituting it into (3.1) as

$$\hat{c}_t = \left(\frac{\omega+\alpha}{\alpha}\right)\hat{z}_t. \tag{3.12}$$

From equation (3.11), in the case of distant complementarity ($u^*_{cz} + \frac{\alpha}{\theta+2\alpha}u^*_{zz} < 0$), $\omega + \alpha$ is negative and thus $(\hat{c}, \hat{z})$-trajectory (3.12) is negatively sloping, whereas under adjacent complementarity ($u^*_{cz} + \frac{\alpha}{\theta+2\alpha}u^*_{zz} > 0$) the trajectory has a positive slope with positive $\omega + \alpha$. This can be understood as follows. Under adjacent complementarity, today's consumption is complementary to today's habits, but a substitute for those in the future steady state. Thus, if habits increase over time, consumption will also increase over time. A positively-sloping saddle arm under adjacent complementarity reflects this property. In contrast, under distant complementarity, today's consumption is complementary to the future steady state habits but a substitute for today's. It follows that increasing habits are accompanied by decreasing consumption, as represented by a negatively-sloping saddle arm.

For the $\left(\hat{k}, \hat{q}\right)$ block, the stable root is given by

$$\chi \equiv \frac{r+\tau-\sqrt{(r+\tau)^2 - 4\eta'(1)\,k^*\beta F^*_{kk}}}{2}\,(<0).$$

The resulting saddle trajectory is:

$$\hat{k}_t = \frac{\eta'(1)\,k^*}{\chi}\hat{q}_t = \frac{r+\tau-\chi}{\beta F^*_{kk}}\hat{q}_t. \tag{3.13}$$

Given the above dynamics for consumption and investment, the dynamics of the net foreign asset are described by the following balance of payments identity obtained from equations (3.4), (3.7), (3.8), and (3.9):

$$\dot{b}_t = rb_t + \beta F(k_t, 1) - c_t - g_t - (1+\phi_t)\,I_t, \tag{3.14}$$

which can be linearized around the steady state as

$$\dot{\hat{b}}_t = -k^*\eta'(1)\,\hat{q}_t + (r+\tau)\hat{k}_t - \hat{c}_t + r\hat{b}_t. \tag{3.15}$$

Using the saddle trajectories, (3.12) and (3.13), we can express $\hat{b}$ in terms of two nonjumpable state variables $\hat{z}$ and $\hat{k}$. This relationship can be obtained by, first, setting $\hat{b} = \varkappa_1 \hat{z} + \varkappa_2 \hat{k}$ and, next, looking for $\varkappa_1$ and $\varkappa_2$ which validate equation (3.15). Indeed, substituting (3.12), (3.13), and the trial solution for $\hat{b}$ into (3.15) yields

$$\{\varkappa_2 (r - \chi) + r + \tau - \chi\} \hat{k}_t + \left\{ \varkappa_1 (r - \omega) - \frac{\omega + \alpha}{\alpha} \right\} \hat{z}_t = 0,$$

so that $\varkappa_1$ and $\varkappa_2$ are obtained as

$$\varkappa_1 = \frac{\omega + \alpha}{\alpha (r - \omega)} \text{ and } \varkappa_2 = -\frac{r + \tau - \chi}{r - \chi},$$

respectively. It follows that the saddle surface for the net foreign asset is given by

$$\hat{b}_t = \frac{\omega + \alpha}{\alpha (r - \omega)} \hat{z}_t - \frac{r + \tau - \chi}{r - \chi} \hat{k}_t. \tag{3.16}$$

When this equation is evaluated at $t = 0$, it characterizes the steady state equilibrium since initial values b_0, k_0, and z_0 are exogenously given. Incorporating this relation, the steady state equilibrium, $(b^*, c^*, k^*, q^*, z^*, \xi^*, \lambda^*)$, is determined by the following equations:

$$\begin{aligned}
q^* &= 1, & (3.17)\\
\beta F_k (k^*, 1) &= r + \tau, & (3.18)\\
c^* &= z^*, & (3.19)\\
\xi^* &= \frac{1}{\theta + \alpha} u_z (c^*, z^*), & (3.20)\\
u_c (c^*, z^*) &= \lambda^* - \alpha \xi^*, & (3.21)\\
r b^* + \beta F (k^*, 1) &= c^* + g, & (3.22)\\
b_0 - b^* + \frac{r + \tau - \chi}{r - \chi} (k_0 - k^*) &= \frac{\omega + \alpha}{\alpha (r - \omega)} (z_0 - z^*), & (3.23)
\end{aligned}$$

where equation (3.17) represents $\dot{k} = 0$ from (3.7); (3.18) represents $\dot{q} = 0$ from (3.8); (3.19) represents $\dot{z} = 0$ from (3.1); (3.20) represents $\dot{\xi} = 0$ from (3.6); (3.22) represents $\dot{b} = 0$ from (3.14); and (3.23) comes from (3.16) evaluated at $t = 0$.

From equation (3.18), the steady state capital stock, k^*, is determined by productivity β and capital tax τ. We denote this capital stock by

$$k^* = K (\beta, \tau), \tag{3.24}$$

where $K_\beta = -F_k (k^*, 1) / \beta F_{kk} (k^*, 1) > 0$ and $K_\tau = 1/\beta F_{kk} (k^*, 1) < 0$. By substituting this into equations (3.22) and (3.23), we obtain

$$b^* = z^*/r + g/r - \beta F [K (\beta, \tau), 1] / r, \tag{3.25}$$

$$b^* = \frac{\omega + \alpha}{\alpha (r - \omega)} (z^* - z_0) + \frac{r + \tau - \chi}{r - \chi} \{k_0 - K (\beta, \tau)\} + b_0. \tag{3.26}$$

These two equations jointly determine b^{*} and z^{*}. Given the z^{*} value and equation (3.19), in turn, (3.20) gives ξ^{*} and (3.21) determines λ^{*}.

The transition dynamics are described by the following autonomous system:

$$
\begin{aligned}
\dot{\hat{b}}_t &= \omega \hat{b}_t + \tfrac{(\omega-\chi)(r+\tau-\chi)}{r-\chi} \hat{k}_t, \\
\dot{\hat{k}}_t &= \chi \hat{k}_t,
\end{aligned} \tag{3.27}
$$

where the first equation is obtained by differentiating (3.16) by t and substituting $\dot{\hat{k}}_t= \chi \hat{k}_t$, $\dot{\hat{z}}_t= \omega \hat{z}_t$, and (3.16) successively into the result. This two-dimensional dynamic system can produce nonmonotonic time-profiles of the current account, as will be shown in the next section.

Figures 1(a) and (b) depict the determination of the equilibrium dynamics derived above. The two figures correspond to the cases of adjacent and distant complementarities, respectively. In the (z, b) plane of the figures, curve BB' represents equation (3.25) and curve ZZ' denotes (3.26). Curve ZZ' is positively sloping under adjacent complementarity and negatively sloping under distant one. Intersection E of the two schedules gives (z^{*}, b^{*}). Given this steady state point, the (k, b) plane depicts the equilibrium dynamics around point $K(k^{*}, b^{*})$, where k^{*} is given by $K(\beta, \tau)$. The $\dot{b} = 0$ schedule can be either positively- or negatively-sloping depending on the sign of $\omega - \chi$. The figures assume $\omega - \chi > 0$ and thus the $\dot{b} = 0$ schedule is positively-sloping. In the (z, c) plane, saddle trajectory (3.12) is represented by schedule SS'. As aforementioned, the slope of the schedule can either positive or negative depending on whether adjacent or distant complementarity is the case. The steady state point, (z^{*}, c^{*}), is given by intersection C of schedule SS' and the 45-degree line representing $\dot{z} = 0$.

3.3 Effects of Macroeconomic Disturbances

Using the above equilibrium dynamics, this section examines effects of changes in government spending g, capital taxes τ, and productivity β. The shocks considered are unanticipated permanent and temporary changes of these variables. They take place at time zero. When the shocks are temporary, they last only until $t = T$ (> 0).

3.3.1 Government Spending

Begin with a permanent increase in government spending g which is financed by a lump-sum tax. The steady state effect of the shock can be derived from equations (3.19), (3.24), (3.25), and (3.26) as

$$
\begin{aligned}
\frac{dk^{*}}{dg} &= 0, \\
\frac{dz^{*}}{dg} = \frac{dc^{*}}{dg} &= \frac{\alpha(r-\omega)}{\omega(\alpha+r)} < 0, \\
\frac{db^{*}}{dg} = \frac{\omega+\alpha}{\omega(\alpha+r)} &\begin{cases} > 0 & \text{for distant complementarity,} \\ < 0 & \text{for adjacent complementarity.} \end{cases}
\end{aligned} \tag{3.28}
$$

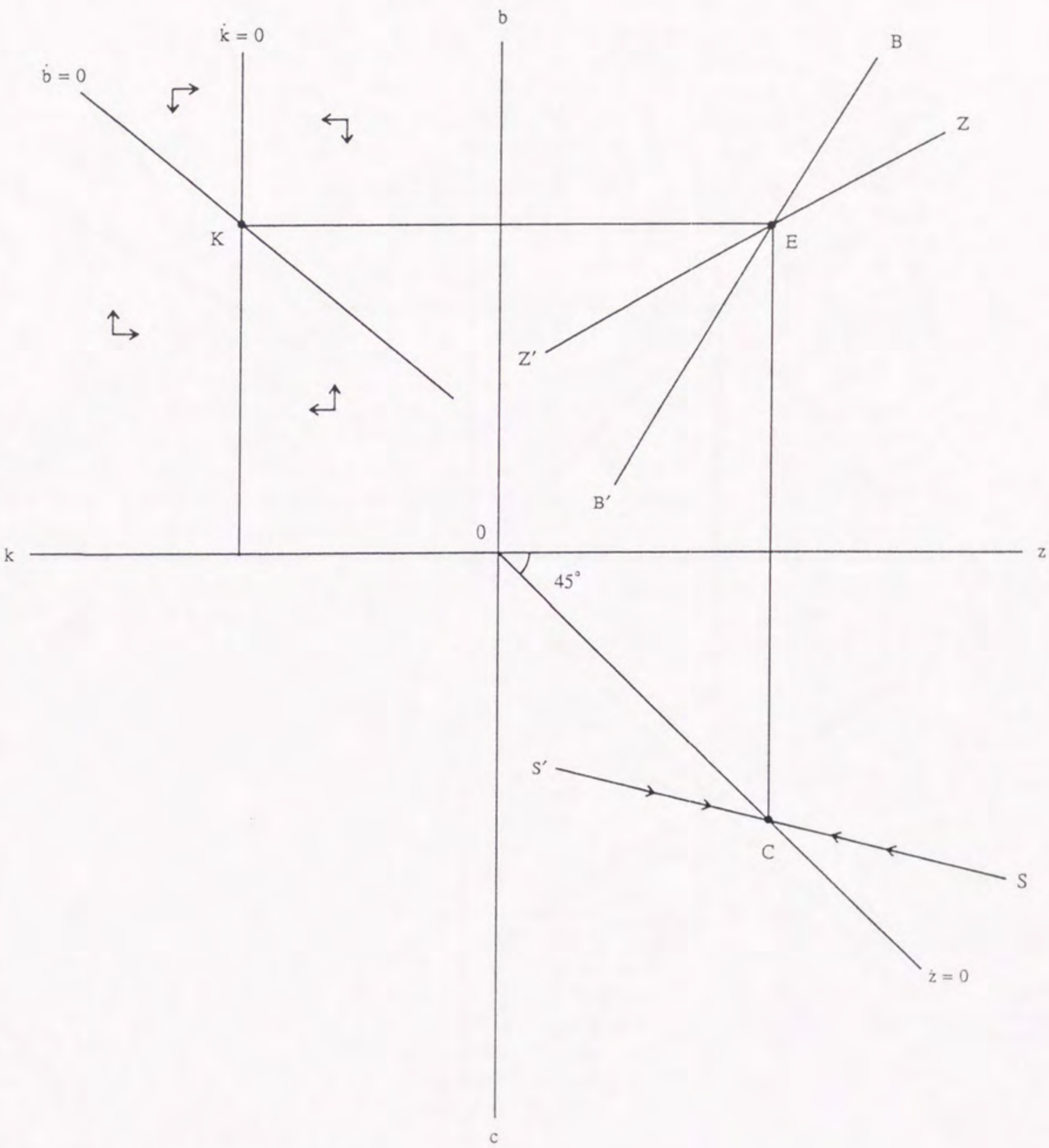


Figure 1(a) Adjacent Complementarity

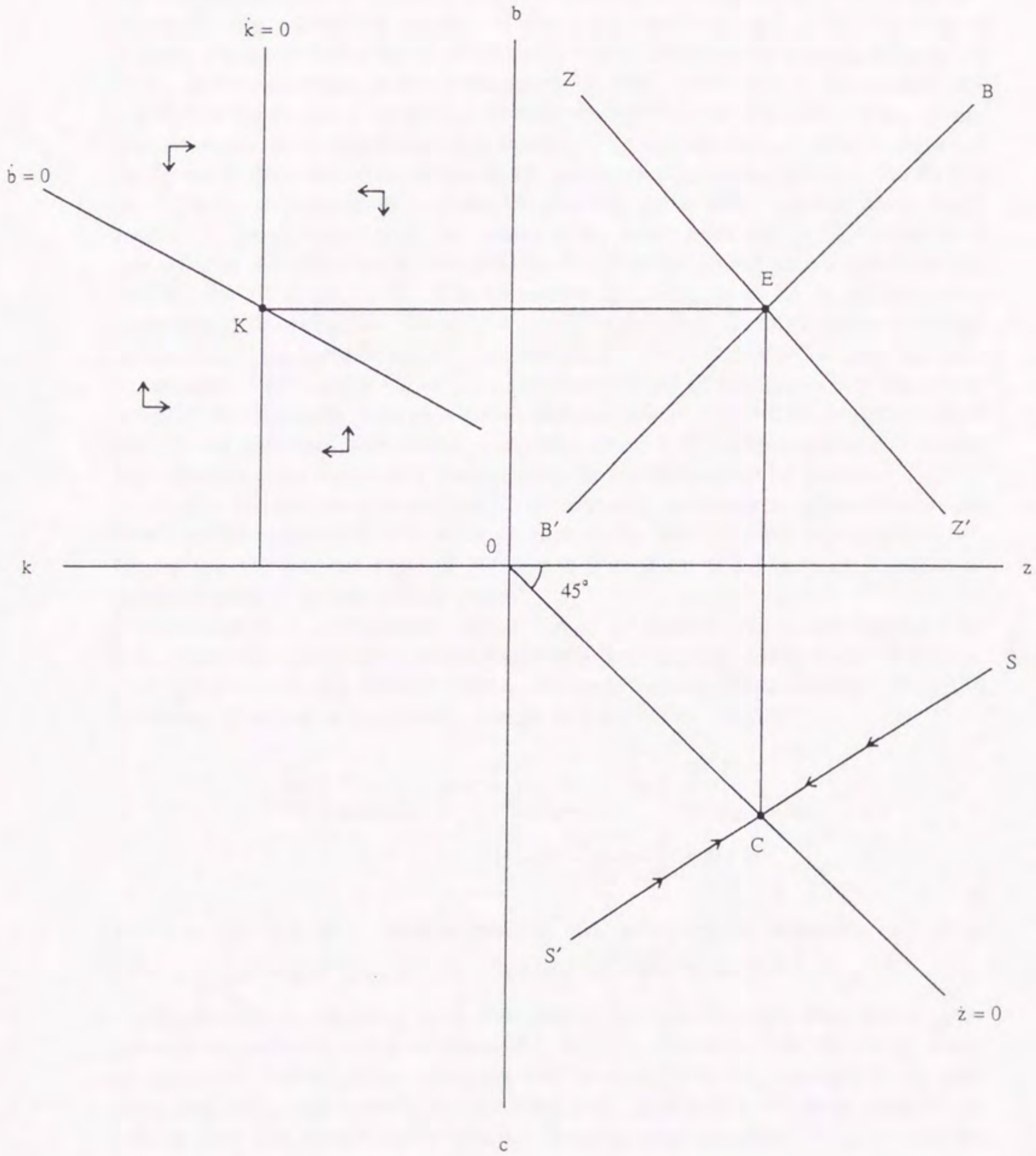


Figure 1(b) Distant Complementarity

Equations (3.28) show that an increase in government spending g reduces the long-run level of consumption c^* and habit z^*. It does not affect the steady state capital stock k^*. The effect on net foreign asset b^* crucially depends on whether adjacent or distant complementarity is the case: under adjacent (resp. distant) complementarity, b^* decreases (resp. increases) in response to the shock. Because of lump-sum tax financing, an increase in g implies a decrease in the permanent income of the representative agent. In the case of adjacent complementarity, he or she will not decrease consumption initially so much as the decrease in the permanent income. This causes the current account deficits in the interim run, thereby reduces the steady state stock of net foreign asset, as in the Keynesian theory. This adjustment process is depicted by figure 2. The initial equilibrium is represented by points C_0 and E_0. In the (z, b) panel, a permanent increase in g shifts curve BB' upward from B_0B_0' to B_1B_1', thereby changing the steady state point from E_0 to E_1. Since k_t is not affected by this shock, schedule Z_0Z_0' given by (3.26) also represents the saddle arm given by (3.16). The transition dynamics from E_0 to E_1 thus take place along the schedule. Given the steady state effect, in turn, trajectory SS' in the (z, c) panel shifts upward to, say, S_1S_1'. The consumption rate initially jumps down from point C_0 to C_{01}. However, owing to adjacent complementarity, this crowding-out is so small that the increase in g dominantly deteriorates the current account. Thereafter, along trajectory S_1S_1', consumption and habits decreasingly approach a new steady state level represented by point C_1.

Under distant complementarity, in contrast, consumers substantially cut down initial consumption upon a decline in the steady state consumption c^*, generating the current account surplus in transition and hence an increase in external asset b^* in the steady state.[4]

Consider next a temporary increase in g. No matter which complementarity is the case, the shock always has negative effects on the steady state consumption, habits, and net foreign assets. Indeed, tedious computation yields the following effects of a temporary change in g on c^*, z^* , and b^*:

$$\begin{aligned} \left.\frac{dc^*}{dg}\right|_{dg=\text{temp.}} &= \left.\frac{dz^*}{dg}\right|_{dg=\text{temp.}} = r\left.\frac{db^*}{dg}\right|_{dg=\text{temp.}} \\ &= (1-\Omega)\frac{\alpha(r-\omega)}{\omega(\alpha+r)} < 0, \end{aligned}$$

where ω' denotes the unstable root for the consumption dynamics and $\Omega = (\omega' - \omega) / \left[\omega' - \omega + \{1 - \exp((r - \omega')T)\}\frac{r(\omega'-\omega)}{\omega'-r}\right] \in (0, 1)$.[5]

Intuitively, in response to a temporary increase in lump sum taxes, permanent income and hence consumption initially decrease. The resulting lower consumption habits reduce the long run consumption even though lump sum taxes and hence permanent income return to their initial levels at time T. It follows that the steady state interest revenue must decrease so as to restore balance to the current account.

---

[4] These effects of a permanent decrease in real income on external debt positions are pointed out by Obstfeld (1992) and Mansoorian (1993, 1996).

[5] For the procedure of deriving the effect of temporary shocks, see Sen and Turnovsky (1990) and Turnovsky and Sen (1991).

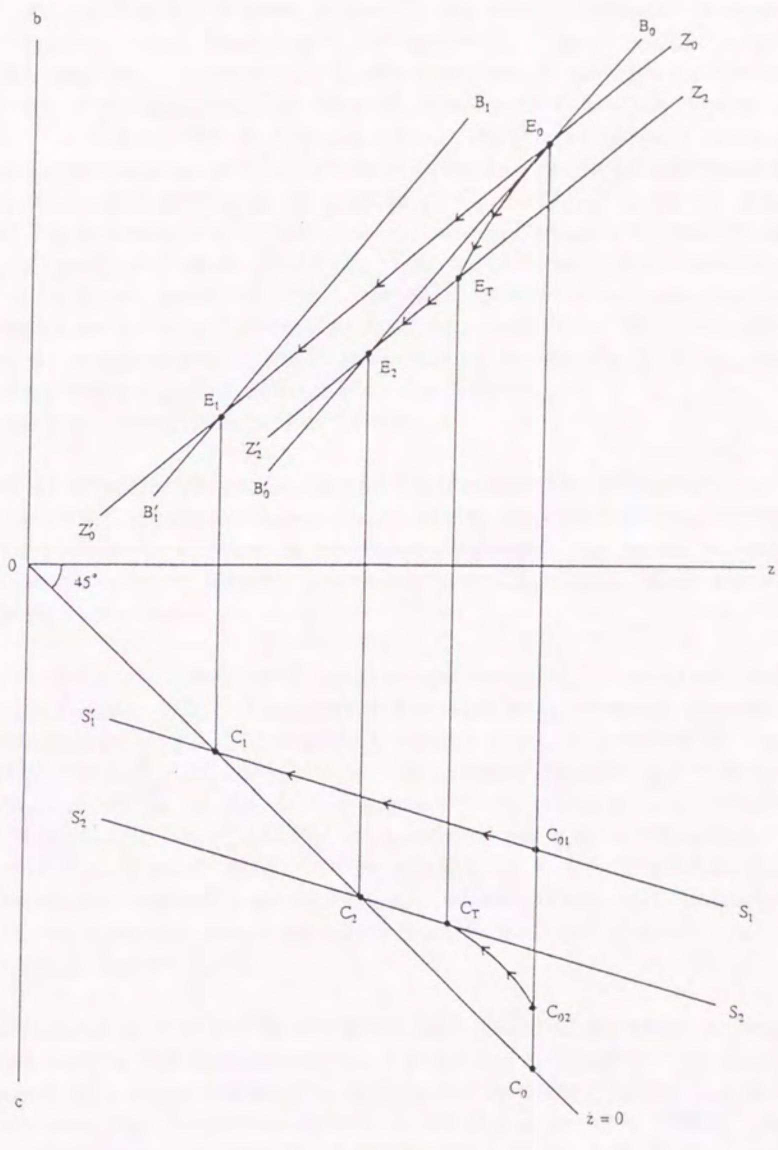


Figure 2 Fiscal Spending

Figure 2 depicts this adjustment process in the case of adjacent complementarity. Upon the shock, saddle arm SS' in the (z, c) panel shifts upward to S_2S_2', which lies below trajectory S_1S_1' corresponding to permanent shocks. The initial jump of consumption rate stops at some point below the new arm, say point C_{02}. This imperfect crowding-out effect on the initial consumption brings about current account deficits. The resultant asset decumulation reduces both c and z as depicted by the arrows from point C_{02} to C_T. Just when time t reaches T, at which g reverts to its initial level, the equilibrium arrives at point C_T in the (z, c) panel and point E_T in the (z, b) panel. Recall that trajectory ZZ' in the (z, b) panel, given by (3.26), depends on the initial asset stocks. The decumulated stocks of the net foreign asset and habits at time T, as a new 'initial' state, make schedule Z_0Z_0' shift downward up to, say, Z_2Z_2'. This gives new steady state point E_2, with smaller b^* and $z^*(= c^*)$.

These results can be summarized as follows:

Proposition 1: *Consider the model presented in Section 3.2. Then, (i) a temporary increase in government spending always deteriorates the current account, whereas (ii) a permanent increase in government spending can either deteriorate or improve the current account, depending on whether adjacent or distant complementarity is the case.*

Remark 1: *In the proposition, the Keynesian-type result (i) is consistent with the oft-reported negative effect of temporary increases in government spending on the current account [e.g., Ahmed (1986), Tanner (1994), and Obstfeld and Rogoff (1995)]. Property (ii) could account for another stylized fact that the effect of permanent shocks on the current account is insignificant [e.g., Ahmed (1986) and Obstfeld and Rogoff (1995)] or significant but smaller than that of temporary ones [e.g., Tanner (1994)]. For example, if $\omega + \alpha$ is positive but sufficiently small, the negative current account-government expenditure linkage is stronger in the case of a temporary shock than in the case of a permanent one, as reported by Tanner (1994).*

Note that the endogenous time preference model cannot mimic these stylized facts. In that setting the long-run consumption level is fixed by the world interest rate and permanent increases in government spending always improve the current account [e.g., Devereux and Shi (1991) and Karayalcin (1994)]. As an alternative specification, the overlapping-generation model [e.g., Matsuyama (1987) and Zou (1994)] cannot explain the negative effect of temporary shocks on the steady state external assets.

3.3.2 Capital Taxes

From equations (3.19), (3.24), (3.25), and (3.26), we can derive the steady state effect of a permanent increase in capital tax $\tau \gtreqless 0$ as follows:

$$\frac{dk^*}{d\tau} = K_\tau < 0,$$

$$\frac{dz^*}{d\tau} = \frac{dc^*}{d\tau} = \tau\left(1 - \frac{r}{r - \chi}\right) K_\tau \lesseqgtr 0, \tag{3.29}$$

$$\frac{db^*}{d\tau} = \frac{\alpha\,(r-\omega)\,\beta F_k^* K_\tau}{\omega\,(\alpha+r)}\left(\frac{r+\tau-\chi}{(r+\tau)\,(r-\chi)} - \frac{\omega+\alpha}{\alpha\,(r-\omega)}\right) \gtreqless 0$$
$$\text{as } \frac{r+\tau-\chi}{(r+\tau)\,(r-\chi)} \gtreqless \frac{\omega+\alpha}{\alpha\,(r-\omega)}.$$

That is, in the long run, an increase in τ reduces capital stock k^* and production. When capital taxation is initially in effect (i.e., $\tau > 0$), the policy negatively affects c^* and z^*. Although the effect on b^* can take either sign, it is positive if initial tax τ is small enough.[6]

Taking this case as an example, figure 3 depicts the adjustment to an increase in capital taxes. The shock makes the $\dot{k} = 0$ schedule shift to the right. This reduces k^*. From equations (3.25) and (3.26), both schedule BB' and ZZ' in the (z, b) panel then shift upward like $B_0B_0' \to B_1B_1'$ and $Z_0Z_0' \to Z_1Z_1'$, respectively. If initial tax τ is small enough, the shift of schedule ZZ' is dominant, thereby increases b^*. Under capital taxation ($\tau > 0$), the reduction in k^*, which decreases the steady state production, dominantly lowers c^* and z^*. These steady state effects are depicted by the movements from point K_0 to K_1, E_0 to E_1, and C_0 to C_1.

The transition dynamics are represented by the arrow-attached paths connecting K_0 to K_1 and C_{01} to C_1. In particular, along path K_0K_1 the current account adjusts nonmonotonically over time: it initially runs a surplus thereafter falls into deficits. The possibility of this 'overshooting' adjustment process of the current account is also derived by Matsuyama (1987). The key assumption behind this is that the capital stock adjustment is faster than the wealth adjustment. In contrast to his model, where the wealth adjustment speed is given by the effective discount rate, the wealth adjustment speed is determined by ω, which reflects the degree of habitual persistence in consumption.[7]

In a similar way to that in Subsection 3.3.1, we can examine the effect of a temporary increase in τ, although detailed discussions entail too tedious computation and are skipped here. From (3.18), the shock does not affect k^*. A new steady state in the (z, b) panel is determined at the intersection of the initial BB'-schedule, B_0B_0', and a new ZZ'-schedule. One interesting possibility is that the new ZZ'-schedule can lie above the initial schedule, Z_0Z_0', owing to the accumulated stock of b_T. The resulting steady state can be characterized by increases in b^*, c^*, and z^*. In this case, the steady state lifetime utility is enhanced by the temporary increase in capital taxes.

3.3.3 Productivity Shocks

From equations (3.25) and (3.26), we can conjecture that effects of negative productivity shocks are qualitatively similar to those of capital taxation. Indeed, from equations (3.19), (3.24), (3.25), and (3.26), we can derive

[6] Indeed, if there is initially no tax distortion, $\tau = 0$, we have $\frac{r+\tau-\chi}{(r+\tau)(r-\chi)} - \frac{\omega+\alpha}{\alpha(r-\omega)} = -\frac{\omega(r+\alpha)}{r\alpha(r-\omega)} > 0$ and hence $\frac{db^*}{d\tau} > 0$.

[7] The overshooting adjustment of the current account cannot take place when the capital adjustment is slower than the wealth adjustment, $\chi > \omega$. In that case the J-curve adjustment may occur as in Karayalcin (1994). Furthermore, in the present model, an increase in capital taxes can reduce the steady state external asset. The adjustment pattern in that case may substantially differ from that of the previous models.

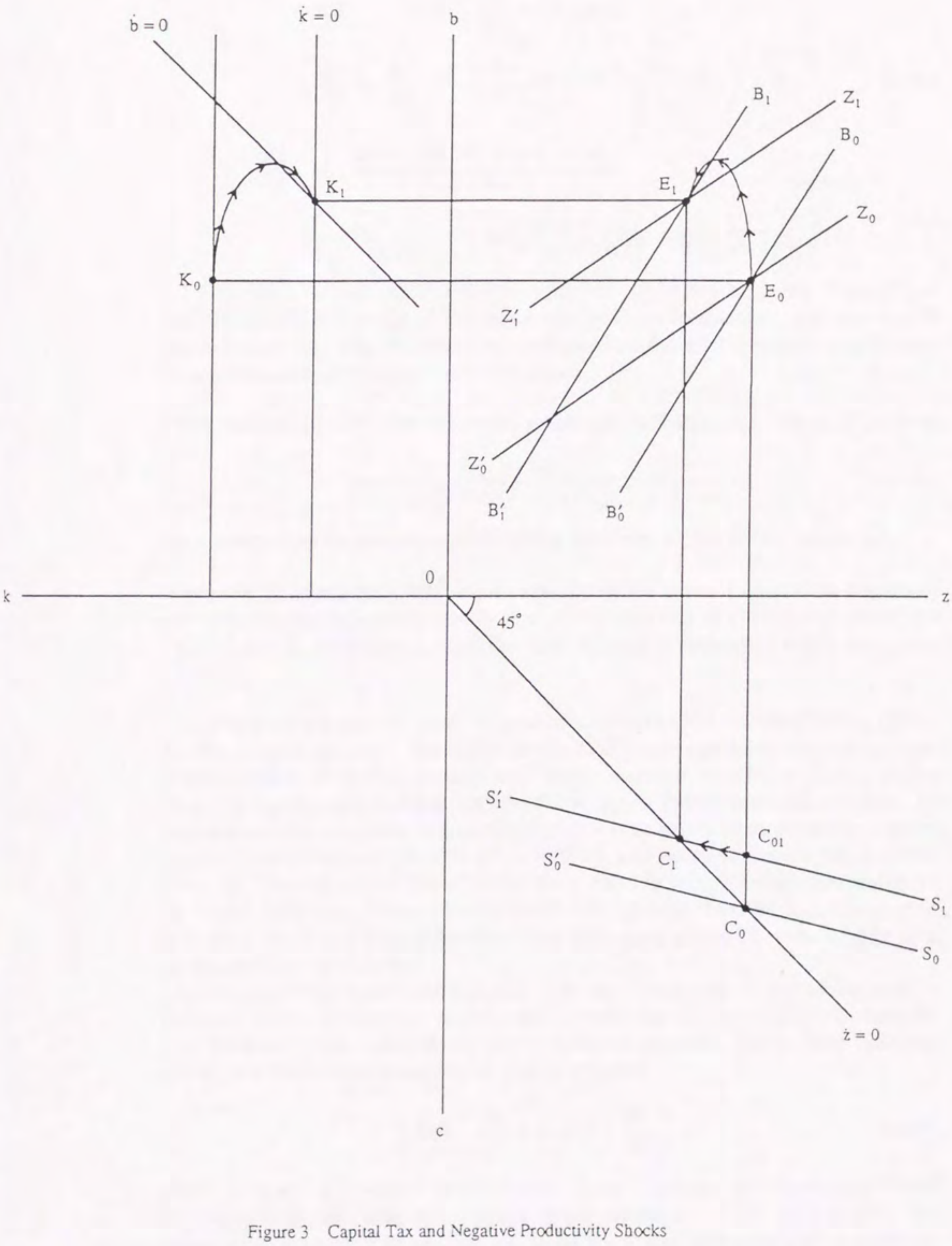


Figure 3 Capital Tax and Negative Productivity Shocks

$$\frac{dk^*}{d\beta} = K_\beta > 0,$$

$$\frac{dz^*}{d\beta} = \frac{dc^*}{d\beta} = -\frac{\alpha (r-\omega)}{\omega (\alpha + r)} \left(F^* - \frac{\tau\chi}{r-\chi} K_\beta \right) > 0, \tag{3.30}$$

$$\begin{aligned} \frac{db^*}{d\beta} &= \frac{\alpha (r-\omega) \{F^* + (r+\tau) K_\beta\}}{\omega (\alpha + r)} \\ &\quad \times \left(\frac{(r+\tau-\chi) K_\beta}{(r-\chi)\{F^* + (r+\tau) K_\beta\}} - \frac{\omega + \alpha}{\alpha (r-\omega)} \right) \gtreqless 0. \end{aligned}$$

Therefore, as seen by comparing equations (3.30) with (3.29), the signs of the above derivatives are almost the same as those in the case of a decrease in capital taxation. Figure 3 can be interpreted as describing overall adjustment to a permanent adverse productivity shock.

Proposition 2: *Consider the model presented in Section 3.2. Then, if we have*

$$\frac{(r+\tau-\chi) K_\beta}{(r-\chi)\{F^* + (r+\tau) K_\beta\}} > \frac{\omega + \alpha}{\alpha (r-\omega)}, \tag{3.31}$$

then permanent increases in productivity deteriorate the current account.

Remark 2: *Many empirical studies report that the current account is negatively correlated with productivity shocks [e.g., Glick and Rogoff (1992) and Elliot and Fatás (1996)]. Proposition 2 implies that this can be mimicked under inequality (3.31).*

Intuitively an improvement in productivity exerts two countervailing effects on the current account: the negative effect of encouraging investment and the positive effect of raising income and hence savings. Condition (3.31) claims that the investment-encouraging effect dominates the saving-raising effect. To understand this, consider a simple case where the production function is given by the Cobb-Douglas type, $F = k^\gamma$, $\gamma \in (0,1)$, and where no capital tax is levied ($\tau = 0$). The left hand side of (3.31) then reduces to γ. For this inequality to be valid, therefore, γ must be sufficiently large compared with ω. Note that a large γ implies a strong investment-encouraging effect whereas a large ω a strong saving-raising effect.

The analytical results, (3.30), also give some potentials of explaining empirical co-movements between savings and investment. Suppose that τ is initially negligible and that adjacent complementarity is the case. Then, from equation (3.23) and the second equation of (3.30), we have

$$\text{sign}\left(\frac{da^*}{d\beta}\right) = \text{sign}\left(\frac{dk^*}{d\beta}\right). \tag{3.32}$$

That is, upon permanent productivity shocks, savings and investment tend to co-move in the long run. Since, from equations (3.28) and (3.29), the other macroeconomic shocks cannot generate either co-movement or counter-movement between them (either of a^* and k^* is unchanged to these shocks), this mimics the oft-reported long run tendency of co-movement in savings and investment [e.g., Tesar (1991) and Coakley, Kulasi, and Smith (1996)].[8]

Given this long-run tendency, the shorter-run savings-investment co-movement can be explained by using the transition dynamics in our model. Since both $b_t + k_t$ and k_t move monotonically if τ is initially negligible,[9] relation (3.32) implies that

$$\operatorname{sign}\left(\dot{b}_t + \dot{k}_t\right) = \operatorname{sign}\left(\dot{k}_t\right). \tag{3.33}$$

Although $\dot{b}+\dot{k}$ differs from savings $\dot{a}\left(= \dot{b} + \dot{k} + \dot{q}k\right)$, capital gains $\dot{q}k$ are usually omitted from the statistical data of savings. Relation (3.33) thus explains the short-run co-movements of savings and investment which are empirically reported [e.g., Finn (1990) and Tesar (1991)].

Regarding these results, three comments are in order. First, adjacent complementarity plays a key role in producing it. Recall that, as demonstrated in Subsection 3.2.4, consumption and habits co-move under adjacent complementarity. This causes co-movements between a^* and z^* as demonstrated in (3.23). Secondly, adjacent complementarity is supported by many empirical studies [e.g., Constantinides (1980)]. Thirdly, the saving-investment co-movement obtained takes place only for permanent shocks. If the shocks are temporary, the underlying relations such as equation (3.23) are violated and hence this property may not hold true.

3.4 Welfare Implications

Using the above equilibrium path of the economy, let us discuss welfare implications. We shall first analyze the effect on welfare (the initial lifetime utility) of permanent macroeconomic shocks, and next examine the saddle dynamics of welfare.

3.4.1 Welfare Effects

To obtain the welfare effects of permanent shocks, linearize the instantaneous utility u around the steady state to obtain: $u(c_t, z_t) = u(c^*, z^*) + u_c^* \hat{c}_t + u_z^* \hat{z}_t$. Next, substituting equation (3.12) and $\hat{z}_t = \hat{z}_0 \exp(\omega t)$ (recall that $\dot{\hat{z}}_t = \omega \hat{z}_t$ along the saddle trajectory) successively into this equation yields: $u(c_t, z_t) = u(z^*, z^*) + \left\{\frac{u_c^*(\omega+\alpha)}{\alpha} + u_z^*\right\} \hat{z}_0 \exp(\omega t)$. By substituting this expression into (3.2), the lifetime utility at time zero is obtained as

$$U_0 = \frac{u(z^*, z^*) + v(g)}{\theta} + \frac{u_c^*(\omega+\alpha)/\alpha + u_z^*}{\theta - \omega}(z_0 - z^*). \tag{3.34}$$

[8] When capital taxes are initially in effect ($\tau > 0$), a^* and k^* may respond in different directions to an increase in τ. Indeed, we can obtain

$$\frac{da^*}{d\tau} = \frac{\tau K_\tau \{\chi(\alpha-\omega) - \omega(\alpha - r)\}}{\omega(\alpha + r)(r - \chi)},$$

which, together with (3.29), implies that $\operatorname{sign}\left(\frac{da^*}{d\tau}\right) = \operatorname{sign}\left(\frac{dk^*}{d\tau}\right)$ if and only if $\chi < \left(\frac{\alpha - r}{\alpha - \omega}\right)\omega$.

[9] The time-path of $b_t + k_t$ is monotonic if τ is initially zero because equation (3.16) connects $b_t + k_t$ with z_t, the time-path of which is monotonic along saddle trajectory (3.12).

It follows that the welfare effect of a permanent increase in variable $y = \tau$, β is given by

$$\frac{dU_0}{dy} = -\frac{\omega\left\{\left(1+\frac{\theta}{\alpha}\right)u_c^* + u_z^*\right\}}{\theta(\theta-\omega)}\frac{dz^*}{dy},$$

which, from regularity condition (c3), implies

$$\operatorname{sign}\left(\frac{dU_0}{dy}\right) = \operatorname{sign}\left(\frac{dz^*}{dy}\right) \quad \left(\equiv \operatorname{sign}\left(\frac{dc^*}{dy}\right)\right) \quad \text{for } y = \tau, \beta. \tag{3.35}$$

Therefore, whether a permanent change in τ and β is beneficial or harmful crucially depends on whether it enhances or weakens the steady state habits. From equations (3.29), (3.30), and (3.35), negative productivity shocks and increases in capital taxes are all harmful to the economy.

As for government spending, equation (3.34) implies

$$\frac{dU_0}{dg} = \frac{v'(g)}{\theta} - \frac{\omega\left\{\left(1+\frac{\theta}{\alpha}\right)u_c^* + u_z^*\right\}}{\theta(\theta-\omega)}\frac{dz^*}{dg}. \tag{3.36}$$

Thus the welfare effect of dg is composed of the direct utility-generating effect (the first term) and the indirect effect (the second term). From (3.28), the indirect effect is negative, reflecting that an increase in g raises lump-sum tax payments. The net welfare effect is determined by the relative magnitudes of these countervailing effects. Specifically, the optimal policy is characterized as one which equates the two effects.[10]

3.4.2 Welfare Dynamics

The transition dynamics of the lifetime utility can be derived by replacing time 0 with time t in equation (3.34) as

$$U_t - U^* = \frac{u_c^*(\omega+\alpha) + \alpha u_z^*}{\alpha(\theta-\omega)}(z_t - z^*), \tag{3.37}$$

where $U^* = \{u(z^*, z^*) + v(g)\}/\theta$.

The adjustment pattern of the resulting welfare dynamics largely depends on the property of intertemporal complementarity. When preferences display distant, or at most weakly adjacent, complementarity (i.e., when $\omega + \alpha$ is sufficiently small), the above trajectory is downward-sloping. This implies that the lifetime utility initially overshoots its steady-state level. Figure 4(a) illustrates this adjustment upon a harmful shock (e.g., an adverse productivity shock), where the schedule, $U^* = \{u(z^*, z^*) + v(g)\}/\theta$, represents the steady state relationship between the lifetime utility and habits. In contrast, in the case of strong adjacent complementarity (i.e., sufficiently large $\omega + \alpha$), the adjustment proceeds along the upward-sloping trajectory. As depicted by figure 4(b), the slope of the trajectory is smaller than that of $U^* = \{u(z^*, z^*) + v(g)\}/\theta$.[11]

[10] To be precise, the policy is constrained-optimal because the time path of g is restricted to be flat.

[11] The slope of $U^* = \{u(z^*, z^*) + v(g)\}/\theta$ is $\frac{u_c^* + u_z^*}{\theta}$. The difference between this and (3.37) can be computed as

$$\frac{u_c^* + u_z^*}{\theta} - \frac{u_c^*(\omega+\alpha) + \alpha u_z^*}{\alpha(\theta-\omega)} = -\frac{\omega}{\alpha\theta(\theta-\omega)}\{(\alpha+\theta)u_c^* + \alpha u_z^*\} > 0.$$

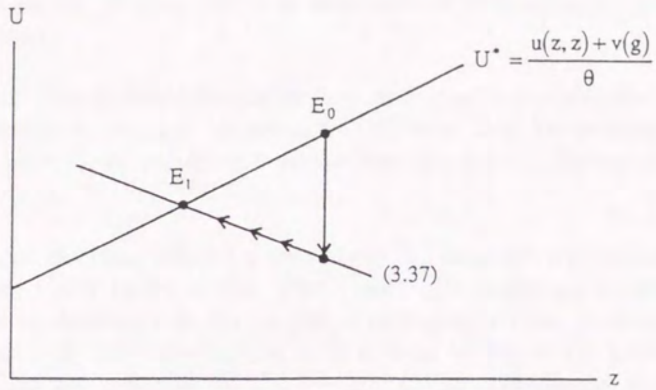


Figure 4(a) $u_c^*(\omega+\alpha)+\alpha u_z^* < 0$

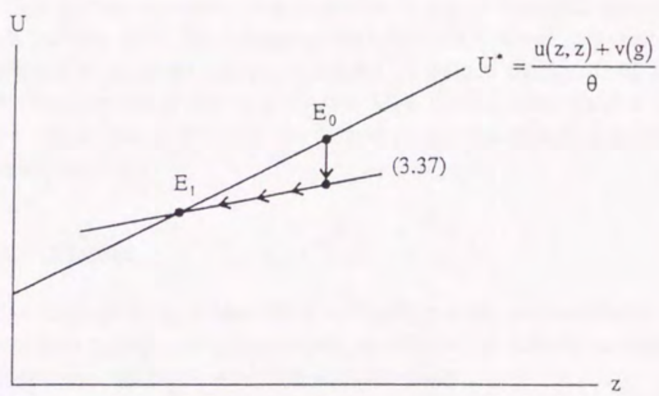


Figure 4(b) $u_c^*(\omega+\alpha)+\alpha u_z^* > 0$

Initial impacts on the lifetime utility in this case are thus always smaller than steady state effects:

Proposition 3: *Strong habit persistence (i.e., strong adjacent complementarity) causes sluggishness in welfare dynamics in the sense that the present lifetime utility is less sensitive to permanent shocks than the future lifetime utility [see figure 4(b)].*

The above proposition reflects a propensity by habitual consumers to give priority to their short range needs. The resulting sluggish adjustment has a sharp contrast to dynamics in the model of endogenous time preferences. In that model, the long run consumption level is fixed by the world interest rate, so that the short-run response of the lifetime utility overshoots its long-run response [e.g., Karayalcin (1994)]. In this habit model, contrastingly, the present consumption and hence the present lifetime utility are anchored to the initial habit. Instead the consumption in the distant future must bear the burden of adjustment against disturbances, which results in high sensitivity of the future welfare and low sensitivity of the present one. Which of the two adjustment patterns is actually the case is purely an empirical problem. As mentioned in Section 3.1, the cumulative empirical literature on intertemporally dependent preferences seems to support the habit model.

One policy implication of proposition 3 is that consumers due to strong habits will sacrifice future consumption in order to maintain their present standard of living. It follows that, for example, even if a permanent increase in fiscal spending has a beneficial effect on the (present) lifetime utility, that is, if the right hand side of equation (3.36) is positive, that policy may have a negative hangover effect in that the welfare in the future (e.g., the steady state welfare) is depressed upon the shock.[12]

3.5 Conclusion

In this paper, we have examined the effect of fiscal policy, capital taxation, and productivity shocks on key macroeconomic variables including consumption, investment, the current account, and the social welfare.

Comparative dynamics have produced several results with theoretical and empirical implications. They can mimic recent stylized facts regarding the current account. The key parameters are the degree of habit persistence and adjustment costs of capital. The stylized facts could be consistently explained by choosing appropriately magnitudes of these two parameters. As a welfare implication, intertemporal complementarities play a crucial role in welfare dynamics. In particular, consumers under strong habit persistence will persist in keeping the present welfare unchanged. As a result, even a beneficial fiscal policy may have a negative hangover effect on the future welfare.

An interesting possible extension is to endogenize the labor supply by incorporating the leisure and labor choice. This will enable us to extend our analysis in two directions. First, it brings to the model interactions between consumption and investment dynamics. Secondly, leisure habits can be incorporated.

[12] Of course, this does not mean that there is a conflict of interests among generations in this dynasty of households.

Bibliography

[1] Ahmed, S., 1986, Temporary and permanent government spending in an open economy, Journal of Monetary Economics 17, 197-224.

[2] Becker, G. and K. Murphy, 1988, A theory of rational addiction, Journal of Political Economy 96, 675-700.

[3] Blanchard, O., 1985, Debt, deficits, and finite horizon, Journal of Political Economy 93, 223-247.

[4] Bovenberg, A. L., 1994, Capital taxation in the world economy, in: F. van der Ploeg ed., The Handbook of International Economics (Blackwell Publishers, Massachusetts).

[5] Braun, P., G. Constatinides, and W. Ferson, 1993, Time nonseparability in aggregate consumption: International evidence, European Economic Review 37, 897-920.

[6] Coakley, J., F. Kulasi, and R. Smith, 1996, Current account solvency and the Feldstein-Horioka puzzle, The Economic Journal 106, 620-627.

[7] Constantinides, G., 1990, Habit formation: A resolution of the equity premium puzzle, Journal of Political Economy 98, 519-543.

[8] Devereux, M. and S. Shi, 1991, Capital accumulation and the current account in a two-country model, Journal of International Economics 30, 1-25.

[9] Elliot, G and A. Fatás, 1996, International business cycles and the dynamics of the current account, European Economic Review 40, 361-387.

[10] Feldstein, M. and C. Horioka, 1980, Domestic saving and international capital flows, Economic Journal 90, 314-329.

[11] Ferson, W. and G. Constantinides, 1991, Habit persistence and durability in aggregate consumption: Empirical tests, Journal of Financial Economics 29, 199-240.

[12] Finn, M., 1990, On savings and investment dynamics in a small open economy, Journal of International Economics 29, 1-21.

[13] Giavazzi, F. and C. Wyplosz, 1985, The zero root problem: A note on the dynamic determination of the stationary equilibrium in linear models, Review of Economic Studies 52, 353-357.

[14] Glick, R. and K. Rogoff, 1992, Global versus country-specific productivity shocks and the current account, NBER Working Paper No. 4140.

[15] Haug, A., 1996, Blanchard's model of consumption: An empirical study, Journal of Business & Economics Statistics 14, 169-177.

[16] Hayashi, F., 1982, Tobin's marginal q and average q: A neoclassical interpretation, Econometrica 50, 213-224.

[17] Karayalcin, C., 1994, Adjustment costs in investment, time preference, and the current account, Journal of International Economics 37, 81-95.

[18] Lawrance, E., 1991, Poverty and the rate of time preference: Evidence from panel data, Journal of Political Economy 99, 54-77.

[19] Mansoorian, A., 1993, Habit persistence and the Harberger-Laursen-Metzler effect in an infinite horizon model, Journal of International Economics 34, 153-166.

[20] Mansoorian, A., 1996, On the macroeconomic policy implications of habit persistence, Journal of Money, Credit, and Banking 28, 119-166.

[21] Matsuyama, K., 1987, Current account dynamics in a finite horizon model, Journal of International Economics 23, 299-313.

[22] Naik, N. and M. Moore, 1996, Habit formation and intertemporal substitution in individual food consumption, Review of Economics and Statistics LXXVIII, 321-328.

[23] Obstfeld, M., 1992, International adjustment with habit-forming consumption: A diagrammatic exposition, Review of International Economics 1, 32-48.

[24] Obstfeld, M. and K. Rogoff, 1995, The intertemporal approach to the current account, in: G. Grossman and K. Rogoff eds., Handbook of International Economics, Vol. 3 (North-Holland Publishing Co., Amsterdam).

[25] Penati, A., 1987, Government spending and the real exchange rate, Journal of International Economics 22, 237-256.

[26] Ryder, H. E. and Heal, G. M., 1973, Optimal growth with intertemporally dependent preference, Review of Economic Studies 40, 1-33.

[27] Sen, P., 1994, Savings, investment, and the current account, in: F. van der Ploeg ed., The Handbook of International Economics (Blackwell Publishers, Massachusetts).

[28] Sen, P. and S.J. Turnovsky, 1991, Investment tax credit in an open economy, Journal of Public Economics 42, 277-299.

[29] Shi, S. and L.G. Epstein, 1993, Habits and time preference, International Economic Review 34, 61-84.

[30] Tanner, E., 1994, The effect of government spending on the current account, output, and expenditures: Evidence from Latin America, Journal of Development Economics 44, 287-310.

[31] Tesar, L., 1991, Savings, investment, and international capital flows, Journal of International Economics 31, 55-78.

[32] Turnovsky, S.J. and P. Sen, 1991, Fiscal policy, capital accumulation, and debt in an open economy, Oxford Economic Papers 43, 1-24.

[33] Uzawa, H., 1968, Time preference, the consumption function, and optimal asset holdings, in: J. Wolfe, ed., Value capital, and growth: Paper in Honor of Sir John Hicks (Aldine, Chicago).

[34] Zou, Heng-Fu, 1994, Short-run analysis of fiscal policy and the current account in a finite horizon model, Journal of Macroeconomics 16, 347-357.

Part II

Bubbles

Chapter 4

Fundamentals-Dependent Bubbles in Stock Prices

Abstract: In a continuous-time model of stock prices with dividends growing stochastically we examine bubbles which depend on market fundamentals. The fundamentals dependency stabilizes bubble dynamics. They can be stochastically stable, saddlepoint-stable, or unstable. Stock prices with these bubbles can be less volatile than fundamental prices. These bubbles exhibit various transition patterns, such as nonmonotonic movements and monotonic shrinkage in magnitude and volatility. The sign of their correlation with market fundamentals is time-varying. We introduce crash risks, permitting bubbles to crash partially and display various stochastic process switching. Crash risks affect the stochastic stability of bubbles.

JEL Classification Number: E44, G12.

Keywords: Fundamentals-dependent bubbles, stochastic stability, price volatility, partial crashes, stochastic process switching.

4.1 Introduction

When the current market price of an asset is determined based on rational expectations of its future price changes, a price bubble can occur with price diverging from market fundamentals in response to arbitrary, self-fulfilling expectations.[1] In a typical formulation [e.g., Blanchard (1979), Flood and Garber (1980), and Diba and Grossman (1983)] an asset price is, explicitly or implicitly, regarded as a function of time, so that price bubbles result from the indeterminacy of this function which arises under self-confirming speculation. As is well-known, these time-driven bubbles (i) are dynamically unstable, (ii) are independent of market fundamentals and as a result market prices are more volatile in the presence of

[1] The empirical validity of the absence (or presence) of price bubbles is still controversial. For example, Blanchard and Watson (1982), West (1987), Froot and Obstfeld (1991b), and Dezhbakhsh and Demirguc-Kunt (1990) provide empirical support for the presence of speculative bubbles. Diba and Grossman (1988a) derive an empirical result against the existence of rational bubbles. Evans (1991) questions Diba and Grossman's empirical findings.

bubbles, and (iii) exhibit monotonic patterns of dynamics until a market crash occurs.

We have another possible scenario which describes speculative bubbles. Suppose that market fundamentals change randomly and their current values provide information for investors to form expectations of future market conditions. Then, an asset price may well be a function of the current value of market fundamentals, and the indeterminacy problem may also arise pertaining to the mapping from market fundamentals to the asset price. Price bubbles would then depend on market fundamentals. This might change the basic features of bubbles.

The purpose of this paper is to provide a theoretical formulation of the above scenario: In a linear arbitrage model of stock price determination we analyze qualitatively fundamentals-dependent bubbles, which are defined in our model as dividend-dependent, rationally-expected price deviations from the expected present value of the dividend stream. The important feature of the model is that dividends follow a continuous Markov process with the result that their current quantities as information affect market prices through expectations. In addition, a free-disposal constraint is imposed explicitly on a stock price, ruling out the existence of negative price. Within this setting, systematic analyses are provided pertaining to the stochastic stability, volatility, and transition patterns of fundamentals-dependent bubbles.

Several researchers have already made successful attempts on the same topic, though they left some theoretical problems. Froot and Obstfeld (1991b) focus on "intrinsic" bubbles, which exclusively depend on market fundamentals, providing an empirical support for the presence of these bubbles in the U.S. stock market. Deriving a bubble solution composed of stable and unstable components, they assume away the stable component on the ground that a stock price should go to zero as dividends go to zero. However, it might be hard to justify this assumption a priori since bubbles represent deviations from fundamentals by definition.[1] Miller and Weller (1990) and Buiter and Pesenti (1990) examine the effects of fundamentals-dependent bubbles on exchange rate dynamics, using log-linear models, in which the consideration of a free-disposal or price-positivity constraint is not required. In models specified in terms of natural numbers (rather than logs), however, some of their dramatic results might be ruled out by this constraint.[2] Resolving these issues, we provide a more general treatment of fundamentals-dependent bubbles.

Furthermore, we formulate a stock price as a function of both time and market fundamentals. The resulting bubble solution will bridge a gap between time-driven bubbles [Flood and Garber (1980) et al.] and bubbles exclusively depending on fundamentals [Froot and Obstfeld (1991b)], including these two solutions as special cases. It will turn out that depending on a parameter which decides relative degrees of fundamentals dependency and time dependency, the bubble solution obtained exhibits various dynamic properties which cannot be derived by combining linearly the two special solutions.

Focusing on this larger class of bubble solutions, we derive the following propositions. (i) In general the fundamentals dependency stabilizes bubble dy-

[1] As an extreme case, Tirole defines bubbles as assets which have positive prices even though they pay no dividends [Tirole (1985), p.1075].

[2] See Diba and Grossman (1987, 1988b) for the implication of a free-disposal constraint on price bubbles.

namics. Indeed, the dynamics of fundamentals-dependent bubbles can be stochastically stable, saddlepoint-stable, or unstable. Owing to this stabilizing effect, fundamentals-dependent bubbles generally satisfy a stochastic version of the transversality condition. (ii) Stock prices with fundamentals-dependent bubbles can be *less* volatile than fundamental prices (i.e., the discounted present values of dividends). Thus, these bubbles may not be precluded by the variance bounds tests introduced by Shiller (1981a, b). (iii) These bubbles exhibit various transition patterns, such as nonmonotonic movements and monotonic shrinkage in magnitude and volatility. Furthermore, the sign of their correlation with market fundamentals is time-varying.

We extend the model by introducing crash risk. The model permits bubbles to crash partially. It is shown that fundamentals-dependent bubbles display various stochastic process switchings in response to partial crashes: The volatility of a price bubble may be instantaneously enlarged; the sign of the correlation between a bubble and market fundamentals can switch instantaneously. The effect of crash risk on the stochastic stability of bubble dynamics is also examined.

The paper is structured as follows. In Section 4.2 we present a stochastic model of stock price determination and derive explicit solutions for the fundamental price and bubbles. Section 4.3 examines the dynamic properties of fundamentals-dependent bubbles. Section 4.4 introduces the possibilities of market crashes. In Section 4.5 are the conclusions.

4.2 Bubbles and Fundamental Prices in a Random Dividend Model

4.2.1 The Basic Model

Let us consider a stock share which yields dividends $D(t)$ at time $t[t \in [0, \infty)]$. These dividends follow a geometric Brownian motion with positive drift:[3]

$$\mathrm{d}D(t) = gD(t)\mathrm{d}t + \sigma D(t)\mathrm{d}z(t), \tag{4.1}$$

$$D(0) = D_0, \quad g - \sigma^2/2 > 0, \quad \sigma > 0,$$

where D_0 is some initial value of D. Constants g and σ are, respectively, the expected value and the standard deviation of the instantaneous rate of dividend growth. $\mathrm{d}z$ is an independent increment of a standard Wiener process, z, with the initial condition $z(0) = 0$. For the purpose of exposition we assume that $g - \sigma^2/2 > 0$. Since $\ln D$ follows a normal distribution $N\{\ln D_0 + (g - \sigma^2/2)t, \sigma^2 t\}$ from (4.1), this assumption implies that the time series of dividend payments have a positive trend.[4] However, this does not affect our main analytical results presented below.

[3] As for stochastic processes and calculus in continuous-time settings, see Malliaris and Brock (1982).

[4] Note the difference between the meaning of $(g - \sigma^2/2)$ and that of g. Application of Ito's Lemma to $\ln D$ yields:

$$\mathrm{d}\ln D = (g - \sigma^2/2)\,\mathrm{d}t + \sigma \mathrm{d}z,$$

implying that $(g - \sigma^2/2)$ represents the expected growth rate of *log*-dividends. Next integrate this equation. Then, after some manipulations, we obtain:

$$D(t) = D_0 \exp\{(g - \sigma^2/2)t + \sigma z(t)\}.$$

Stochastic dividend-payment process (4.1) is the only source of randomness and the filtration generated by this process, $\Omega = \{\Omega_t\}_{t=0}^{\infty}$, specifies the information process of investors. Supposing risk-neutrality of investors and free disposability of the stock, we assume that the (cum-dividend) stock price, P, is determined by the following two conditions:

$$\mathrm{E}[\mathrm{d}P(t)|\Omega_t]/\mathrm{d}t + D(t) = rP(t), \qquad r > 0, \tag{4.2}$$

$$P(t) \geq 0, \quad \forall t \in [0, \infty), \quad \text{w.p.1}, \tag{4.3}$$

where $\mathrm{E}[\cdot|\Omega_t]$ represents mathematical expectations conditional on Ω_t, and parameter r denotes the riskless interest rate which is assumed to be constant.

Eq. (4.2) represents the standard arbitrage condition, requiring that the expected return on the stock (the expected instantaneous capital gains plus dividend payments) equals the riskless interest. Eq. (4.3) is the price positivity condition. The model is fully described by eqs. (4.1) through (4.3), as being a stochastic version of Williams's (1938)-Gordon's (1962) model of growing stock prices and, at the same time, the continuous-time variety of Froot and Obstfeld's (1991b).

The rational expectations stochastic process of the stock price is obtained by solving nonhomogeneous partial differential eq. (4.2), subject to the dividend payment process (4.1) and the price positivity condition (4.3). Viewing D as a Markov process from (4.1) such that its current value completely describes the information at that time, we assume that the solution to this problem is a twice-differentiable deterministic function of both market fundamentals D and time t :

$$P(t) = P(D, t).$$

As in the case of deterministic rational expectations models, the solution of eq. (4.2) with respect to $P(D,t)$ can be expressed as the sum of the forward-looking particular solution, $F(D,t)$, and the general solution of the homogeneous counterpart of (4.2), $B(D,t)$:

$$P(D,t) = F(D,t) + B(D,t). \tag{4.4}$$

Given this decomposition, we call the forward-looking solution of (4.2), $F(D,t)$, a *fundamental price process* if it satisfies the positivity condition,

$$F(\cdot, t) \geq 0, \quad \forall t \in [0, \infty), \quad \text{w.p.1}.$$

Provided that a fundamental price process exists, a *price bubble process* is defined as the general solution of the homogeneous counterpart of (4.2), $B(D,t)$, which satisfies the price positivity condition, (4.3).[5] We now derive conditions under which these processes exist and present their explicit solutions.

By applying Arnold (1973)'s Lemma 8.4.4 (p.138), the conditional expectation of $D(t)$ is computed as:

$$\mathrm{E}[D(t)|\Omega_0] = D_0 \cdot \exp(gt),$$

which reveals that g denotes the expected growth rate of the *level* of dividends.

[5] These definitions of fundamental prices and rational price bubbles are the standard ones. For example, see Diba and Grossman (1988b) and Froot and Obstfeld (1991b).

4.2.2 The Fundamental Price Process

By definition, function F is obtained from (4.1) as:

$$F(D,t) = \mathrm{E}\left[\int_t^\infty \exp\{-r(s-t)\} \cdot D(s)\mathrm{d}s|\Omega_t\right]. \tag{4.5}$$

Note that $\mathrm{E}[D(s)|\Omega_t] = D(t) \cdot \exp\{g(s-t)\}(s \geq t)$ from (4.1).[6] Therefore, from eq. (4.5), a fundamental price process exists if and only if:

$$r > g. \tag{4.6}$$

Assume that this inequality holds. Then, the fundamental price process is uniquely determined as:

$$F(D,t) = D(t)/(r-g). \tag{4.7}$$

More explicitly, the stochastic process of the fundamental price is obtained by applying Ito's Lemma to (4.7) and substituting (4.1) into the result:

$$\mathrm{d}F = gF\mathrm{d}t + \sigma F\mathrm{d}z. \tag{4.8}$$

Thus, the proportional rate of rises in the fundamental price, $\mathrm{d}F/F$, equals that in the market fundamentals (i.e., dividends), $\mathrm{d}D/D$, with probability one.

4.2.3 Bubble Processes

By construction, function B satisfies the following homogeneous equation:

$$\mathrm{E}[\mathrm{d}B(D,t)|\Omega_t]/\mathrm{d}t = rB(D,t). \tag{4.9}$$

Applying Ito's Lemma to function B to obtain $\mathrm{E}[\mathrm{d}B|\Omega_t]/\mathrm{d}t$, eq. (4.9) is rewritten as:

$$(1/2)\sigma^2 D^2 B_{DD}(D,t) + gDB_D(D,t) + B_t(D,t) = rB(D,t), \tag{4.10}$$

where the subscripts on B denote its partial derivatives, e.g., $B_D = \partial B/\partial D$, $B_{DD} = \partial^2 B/\partial D^2$ etc. When we define $q = \ln D$ and set:

$$B(D,t) = b(q,t),$$

eq. (4.10) reduces to a differential equation with constant coefficients:

$$(1/2)\sigma^2 b_{qq}(q,t) + (g - \sigma^2/2)b_q(q,t) + b_t(q,t) = rb(q,t). \tag{4.11}$$

This partial differential equation is solved by the separation-of-variables method.[7] First of all, set function $b(q,t)$ equal to the product of a fundamentals-dependent factor, $x(q)$, and a time-dependent factor, $y(t)$:

$$b(q,t) = x(q) \cdot y(t). \tag{4.12}$$

---

[6] See footnote 5.

[7] See Farlow (1983) for analytical methods of solving partial differential equations.

Next, substitute this into eq. (4.11). The resulting equation can be decomposed into the following two ordinary differential equations with respect to $x(q)$ and $y(t)$, respectively:

$$(\sigma^2/2)x''(q) + (g - \sigma^2/2)x'(q) = kx(q), \tag{4.13}$$

$$y'(t) = (r - k)y(t), \tag{4.14}$$

where k is an arbitrary constant. These equations are easily solved as:

$$x(q) = G_1 \exp(\lambda_1 q) + G_2 \exp(\lambda_2 q), \tag{4.15}$$

$$y(t) = J \cdot \exp\{(r - k)t\}, \tag{4.16}$$

where G_i $(i = 1, 2)$and J are arbitrary constants and λ_i denote characteristic roots for (4.13):

$$\begin{aligned} \lambda_1 &= \tfrac{-(g-\sigma^2/2)+\{(g-\sigma^2/2)^2+2\sigma^2 k\}^{1/2}}{\sigma^2}, \\ \lambda_2 &= \tfrac{-(g-\sigma^2/2)-\{(g-\sigma^2/2)^2+2\sigma^2 k\}^{1/2}}{\sigma^2}. \end{aligned} \tag{4.17}$$

Finally, by substituting eqs. (4.15) and (4.16) into (4.12) and setting $G_i J = A_i$, we obtain a solution to (4.11) as:

$$b(q, t) = A_1 \exp\{(r - k)t + \lambda_1 q\} + A_2 \exp\{(r - k)t + \lambda_2 q\}.^8 \tag{4.18}$$

As can be seen from eqs. (4.13) and (4.14), parameter k represents the expected rate of changes in fundamentals-dependent factor $x(q)$, and $(r - k)$ in time-dependent factor $y(t)$. Since factor y is deterministic and hence its expected changes equal actual ones, time t drives y at the evolution rate, $(r-k)$. From eq. (4.17), on the other hand, parameters λ_i, which characterize the stochastic dynamics of $x(q)$ [and hence of $b(q, t)$], are non-linear functions of its expected evolution rate k. Fig. 1 depicts this relation between λ_i and k. As is shown in the next section, it is this nonlinearity that causes price bubbles to exhibit various dynamic properties depending on k.

Bubble processes are derived by restricting constants A_i and k such that function (4.18) satisfies the price positivity condition (4.3). When we define

$$\underline{k} = -\frac{(g - \sigma^2/2)^2}{2\sigma^2},$$

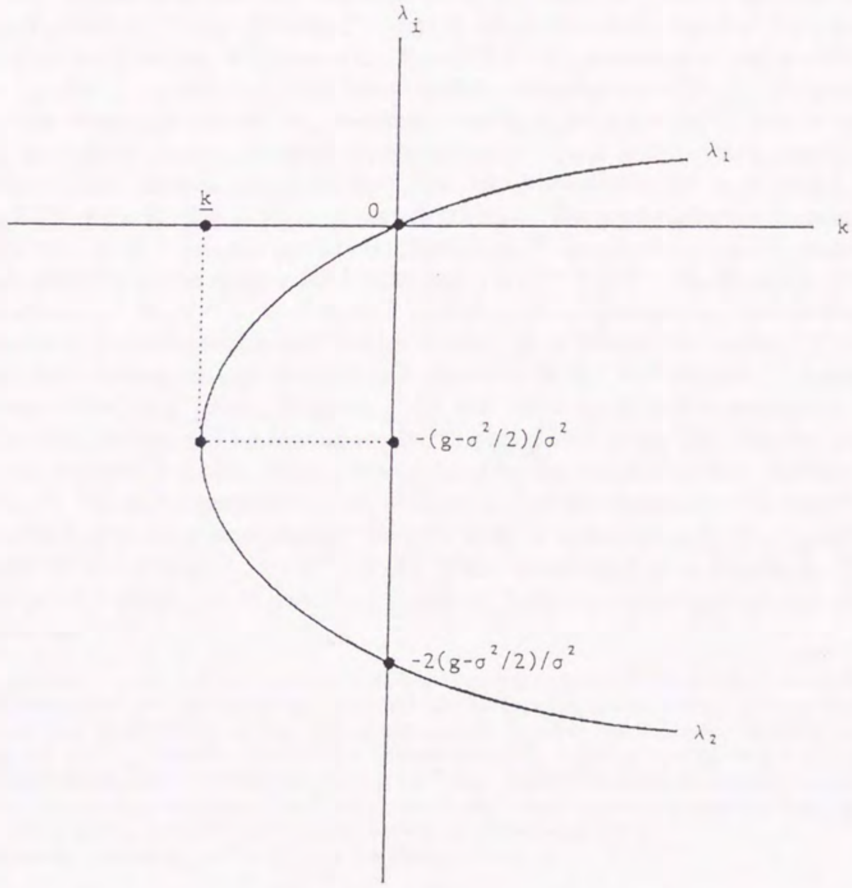
this restriction amounts to the following:

Proposition 1: *Suppose that process b [given by (4.18)] is nontrivial in the sense that it is not always equal to zero. Then, the necessary and sufficient condition for this process to satisfy the price positivity condition (4.3), and hence to be a price bubble, is*

$$k \geq \underline{k} \text{ and } (A_1, A_2) > 0, \tag{4.19}$$

where $(\cdot, \cdot) > 0$ represents that the two elements are both positive and at least one of them is strictly positive.

Proof: See Appendix A.1.□



λ_i
λ_1
$\underline{k}$
0
k
$-(g-\sigma^2/2)/\sigma^2$
$-2(g-\sigma^2/2)/\sigma^2$
λ_2

Fig. 1

From eq. (4.17) or fig. 1 the λ_i values are real numbers if $k \geq \underline{k}$, otherwise they are imaginary numbers. Thus, the above proposition says that a price bubble exists if and only if the roots λ_i are real numbers and, at the same time, the A_i values are both positive. If the λ_i are imaginary numbers and if function b is not trivially equal to zero, process b displays a cycle around zero which is perturbed by Brownian motion z, so that it becomes strictly negative with non-zero probability. However small the amplitude of the cycle may be, perturbation z, which follows a normal distribution, causes b to deviate downwards from this cyclical trend such that process $F + b$ as a whole becomes negative with non-negligible probability. On the other hand, if the λ_i values are real numbers and if either A_1 or A_2 [and thus one of $b's$ two components in (4.18)] is strictly negative, there exist values for parameter q which make this negative component large in absolute value such that processes b and $P(= F + b)$ become negative.

Given that inequalities (4.19) hold true, our bubble solution is provided by eq. (4.18), with A_i and k being given arbitrarily.[12] By construction it generally depends on both time and market fundamentals.[13] As shown in fig. 1, when k equals zero, λ_1 becomes zero [and λ_2 equals $-2\,(g - \sigma^2/2) < 0$]. In one special case where $k = 0$ $(\lambda_1 = 0)$ and $A_2 = 0$, therefore, price bubbles are independent of fundamentals and exclusively driven by time, as in Flood and Garber (1980). In another special case in which $k = r$, bubbles exclusively depend on fundamentals from (4.18). Since, from eq. (4.7), the fundamental price component F is also fully determined by fundamentals, this single-state-variable solution represents 'intrinsic' bubbles, which are examined by Froot and Obstfeld (1991b).[14] (Actually, by setting additionally $A_2 = 0$, we can obtain essentially the same solution that they have focused on.) We next analyze systematically the dynamic properties of our hybrid solution, (4.18), from the viewpoints of stochastic stability, price volatility, and transition patterns, bridging a gap between the two special cases.[15]

[12] Constants $A_i (i = 1, 2)$ and k would be endogenously determined if some preannouncement concerning future process switching of market fundamentals were introduced. For a model in which process switching is specified as conditional on time, see, for example, Gray and Turnovsky (1979). Models of stochastic process switching in which the process of market fundamentals switches conditionally on some particular state of the market include Flood and Garber (1983, 1991); Krugman (1987, 1991); Froot and Obstfeld (1991a); Miller, Weller, and Williamson (1989); Svensson (1991); and Bertola and Svensson (1990).

[13] However, note that, if $\sigma^2 = 0$, equation (4.13) reduces to:

$$g \cdot x' = k \cdot x.$$

Since $q(t) = q_0 + gt$ in this case, this equation is solved as:

$$x(q) = K \cdot \exp\{(k/g)q(t)\} = L \cdot \exp(kt),$$

where K and L are arbitrary constants. Thus, the fundamentals-dependent factor is described completely in terms of time here. This is because in a deterministic setting each value of market fundamentals has a one-to-one correspondence with time, so that the current state contains exactly the same information as time.

[14] Froot and Obstfeld (1991b) also analyze another special case corresponding to $k = g$ and $A_2 = 0$.

[15] As is easily seen from the above discussion, our bubble solution (4.18) does not satisfy McCallum's (1983) 'minimal-state-variable' criterion. Although this criterion might facilitate focusing on fundamental prices, it does not seem to provide a rational restriction on speculative price dynamics such as bubbles, sunspots, and chaotic price motions.

4.3 The Stability, Volatility, and Dynamics of Price Bubbles

4.3.1 Stochastic Stability

Let us first examine the stochastic (or almost sure) stability of price bubble b. Bubble process $b(t)$ is defined as stochastically (or almost surely) asymptotically stable (in the large) if it converges *with probability one* as $t \to \infty$ [for any $b(0) \in [0, \infty)$].[16] Note that time t evolves a price bubble in two ways: by driving time-dependent factor y directly and by evolving fundamentals-dependent factor x through changing q randomly. As is easily seen from the previous discussion on eq. (4.18), the stability of $y(t)$ is determined by $(r-k)$ while the stochastic stability of $x(q)$ by characteristic roots λ_i, which are non-linear in k. Because of this non-linearity, the stochastic stability of bubble dynamics has a large spectrum depending on k, as we shall show now.

Since from eq. (4.1) $q(t)(= \ln D)$ is expressed as:

$$q(t) = q_0 + (g - \sigma^2/2)t + \sigma z(t), \tag{4.20}$$

where $q_0 = \ln D_0$, eq. (4.18) can be rewritten as:

$$b(q,t) = \sum_i^2 A_i (D_0)^{\lambda_i} \cdot \exp[\{r - \psi_i(k)\}t + \lambda_i \sigma z(t)], \tag{4.21}$$

where:

$$\psi_i(k) = k - \lambda_i (g - \sigma^2/2), \quad i = 1, 2. \tag{4.22}$$

In eq. (4.21) we have $\lim_{t\to\infty}\{z(t)/t\} = 0$ (w.p.1) by the strong law of large numbers [see Arnold (1973), p.46]. Thus, it implies that the stochastic stability of b crucially depends on the relative magnitudes of r and ψ_i :

Proposition 2 (*stochastic stability of price bubbles*): *Suppose that the inequalities in (4.19) hold, and $A_1 > 0$ and $A_2 > 0$. Then, (i) if $r < \psi_1(k)$, price bubble process $\{b(q(t),t)\}_{t=0}^{\infty}$ $(= \{B(D(t),t)\}_{t=0}^{\infty})$ is stochastically asymptotically stable; (ii) if $\psi_1(k) \le r < \psi_2(k)$, it is stochastically saddle-point stable; and (iii) if $\psi_2(k) \le r$, it is stochastically unstable.*[17]

Fig. 2 illustrates this proposition by drawing the loci of ψ_i in $k - \psi_i$ space. As is verified in Appendix A.2, ψ_1 is represented by a U-shaped curve which has a turning point at the origin. The ψ_2 is monotonically increasing in k. The gradients of both curves become infinite (in absolute values) at point $\underline{k}$ and approach one as $k \to +\infty$. Given that the vertical axis measures the interest

[16] It is well known that the stability of a stochastic system can be defined differently depending on how stochastic convergence is defined (e.g., in terms of 'with probability one' or the n-th moment). However, the definition of stochastic stability, with which we are concerned here, is a direct and natural extension of the deterministic definition. For details concerning the stability of stochastic dynamics, see Arnold (1973), Chap.11. See also Turnovsky and Weintraub (1971) and Kiernan and Madan (1989) for stochastic-stability analyses of economic models.

[17] Proposition 2 assumes that both A_1 and A_2 are strictly positive. When one of the two equals zero, stochastic process b must be either stable or unstable depending on the relative magnitudes of r and ψ_i which is associated with nonzero A_i.

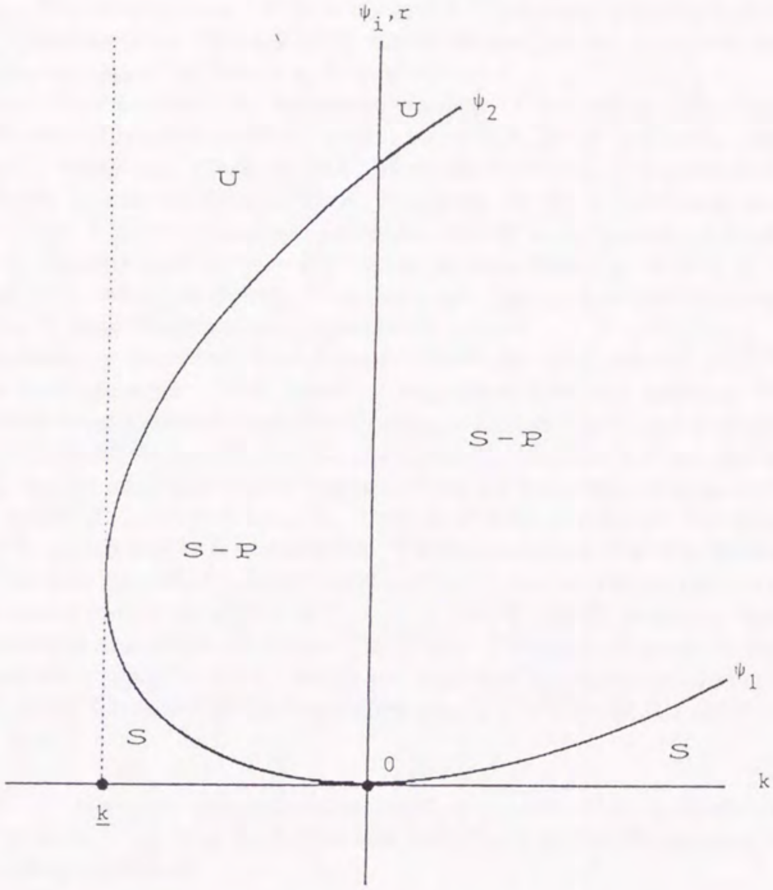


Fig. 2. Stochastic stability of bubbles. Region S represents the stochastically stable area; Region S-P denotes the stochastically saddle-point stable area; and Region U is the stochastically unstable area.

rate r, region S represents the set of points (k, r) under which bubble dynamics are stochastically asymptotically stable, region $S-P$ denotes the stochastically saddlepoint-stable area, and region U the stochastically unstable area. It can easily be seen from this figure that, for a given k, the larger r becomes, the less stable are bubble dynamics. This is because a large r implies a high evolution rate of time-dependent factor y while the λ_i values, the characteristic roots of fundamentals-dependent factor x, do not rely on r.

The relation between the stochastic stability of bubbles and the expected evolution rate of fundamentals-dependent factor x, k, is not so simple, reflecting that the λ_i values are nonlinear in k. However, note that if a price bubble is independent of market fundamentals, it evolves at the growth rate equal to r while from (4.21) the (average) evolution rate of a fundamentals-dependent bubble is characterized by $(r - \psi_i)$. Here, as seen from fig. 2, the ψ_i values are positive for any $k \in [\underline{k}, \infty)$. Therefore, the dynamics of price bubbles are stabilized by their fundamentals dependency.

This result is consistent with Arnold's (1973, p. 185) remark 11.2.17 and Kiernan and Madan's (1989) result. They show that the addition of heteroscedastic shocks tends to stabilize dynamics. Kiernan and Madan provide an intuitive explanation, which may be applicable to interpret the stochastic stability of fundamentals-dependent bubbles: Consider a discrete-time counterpart of our model: $\Delta b_{i+1} = (r + \alpha z_{i+1})b_i$ where z_i denotes a noise on fundamentals in the i-th period and α is a multiplier. These shocks can stabilize bubble dynamics, as may be seen by observing the effect of two shocks of the same size but different signs, $(r + \alpha z)(r - \alpha z) = \{r^2 - (\alpha z)^2\}$, which is smaller than r^2, the exclusively time-dependent case. Intuitively, the same argument is valid in our model since positive shocks, which are generated by Brown motion z, occur with the same frequency as corresponding negative shocks of the same size in infinite time.[18]

Remark 1: *Owing to this stabilizing effect of fundamentals dependency, for any nonzero* $k \in [\underline{k}, \infty)$, *a price bubble satisfies a stochastic version of the transversality condition,*

$$ac\text{-}\lim_{t\to\infty} \exp(-rt)\{F(D,t) + b(q,t)\} = 0, \tag{4.23}$$

where ac represents 'with probability one'. This can be seen by deriving from (4.21) the process for discounted bubble $\exp(-rt)b(q,t)$:

$$\exp(-rt)b(q,t) = \sum_i^2 A_i(D_0)^{\lambda_i} \cdot \exp[\{-\psi_i(k)\}t + \lambda_i \sigma z(t)].$$

As noted above, $ac\text{-}\lim_{t\to\infty}\{z(t)/t\} = 0$ *by the strong law of large numbers. Thus, the evolution rate of a discounted bubble is characterized by* $-\psi_i$, *which are strictly negative for any nonzero* $k \in [\underline{k}, \infty)$ *as shown in fig. 2. Thus, a discounted bubble (as well as the discounted value of the fundamental price*

[18] See Kiernan and Madan (1989) for more detailed discussions.

component) converges to zero with probability one as $t \longrightarrow \infty$.[19,20]

4.3.2 Volatility

The volatility of price bubble b can be measured in terms of the absolute value of partial derivative b_q. This can be seen if we obtain bubble dynamics by applying Ito's Lemma to $b(q,t)$ and substituting (4.11) into the result as:

$$\mathrm{d}b = rb\mathrm{d}t + b_q\sigma\mathrm{d}z, \tag{4.24}$$

which implies:

$$\mathrm{var}(\mathrm{d}b) = (b_q)^2\sigma^2, \tag{4.25}$$

where $\mathrm{var}(\cdot)$ represents variance per unit time. Thus the volatility of bubbles is proportional to that of dividend payments and $(b_q)^2$, so that a large $|b_q|$ implies large volatility of bubbles relative to that of the market fundamentals.

Furthermore, b_q determines the sign of the correlation between b and F : From (4.8), (4.24), and (4.25), the instantaneous correlation coefficient between $\mathrm{d}b$ and $\mathrm{d}F$, $\Gamma(\mathrm{d}b,\mathrm{d}F)$, is obtained as:

$$\Gamma(\mathrm{d}b, \mathrm{d}F) = \mathrm{sign}(b_q) \cdot (+1).$$

That is, $\mathrm{d}b$ and $\mathrm{d}F$ are perfectly correlated with each other and the sign of the correlation is the same as that of b_q.

In short, the volatility of price $P(= F + b)$ depends on b_q in two ways: the absolute value of b_q affects the price volatility by changing the volatility of bubbles; and the sign of b_q affects it by determining the sign of the correlation between the bubble component and the fundamental price component. Therefore, the stock price may be less volatile in the presence of bubbles than in the absence of them if b_q is negative and not so large in absolute value. In fact, since eqs. (4.4), (4.8), and (4.24) imply:

$$\mathrm{var}(\mathrm{d}P) = \{(b_q/F) + 1\}^2\, \mathrm{var}(\mathrm{d}F),$$

we obtain the following result:

---

[19] As for the discounted value of the fundamental price component, $\exp(-rt)F(D,t)$, we obtain from equation (4.8):

$$\begin{aligned} ac\text{-}\lim_{t\to\infty} \exp(-rt)F(D,t) &= ac\text{-}\lim_{t\to\infty} \exp(-rt)F(D_0,0)\exp\{(g-\sigma^2/2)t + \sigma z(t)\} \\ &= 0, \end{aligned}$$

since $r > (g - \sigma^2/2)$ from (4.6).

[20] In general, mean stability implies stochastic stability, but the converse is not true. See Arnold (1973), Chap. 11. Indeed, from eq. (4.9), $B(= b)$ is expected to diverge at the positive rate, r, so that it is always unstable in mean even if it is stochastically stable. As can be seen from Froot and Obstfeld's (1991b, p.1197) discussion, a stable bubble will be far from zero relatively rarely in large samples, but when it is, it diverges by an amount large enough to equalize the mean growth rate of the bubble to the interest rate. Reflecting this, price bubbles do not satisfy the transversality condition (TVC) if, replacing (4.23), we follow many studies [e.g., Froot and Obstfeld (1991b)] to specify the TVC in terms of the mean as:

$$\lim_{t\to\infty} \exp(-rt)E[F(D,t) + b(q,t)|\Omega_0] = 0.$$

Imposing some conditions, Brock (1982) proves that the TVC in terms of the mean is a necessary condition for optimality.

Proposition 3 (*volatility of bubbly prices*): *The volatility of price P and that of the fundamental price component F satisfy:*

$$\mathrm{var}(\mathrm{d}P) \gtreqless \mathrm{var}(\mathrm{d}F) \Leftrightarrow |(b_q/F)+1| \gtreqless 1. \tag{4.26}$$

Thus, the volatility of bubbly prices can be smaller than that of the fundamental price. This phenomenon occurs when $-2F < b_q < 0$. In this case, a bubble is negatively correlated with the fundamental price while the volatility of the bubble itself is not too large, so that the price volatility as a whole becomes smaller in the presence of bubbles.

Remark 2: *This possibility implies that, if the fundamental-dependent bubbles presented here take place, they may not be precluded by the variance bounds tests introduced by Shiller (1981a, b).*

4.3.3 Dynamics: Two Special Cases

It is difficult to depict the dynamic transition patterns of price bubbles qualitatively since in our formulation they are driven by two forces, t and q (or, equivalently, D). Concentrating on clarifying the effect of fundamentals-dependency on bubble motions, we here compare two special cases; the case of fundamentals-independent bubbles and that of time-independent intrinsic bubbles.

As is pointed out in Section 4.2, fundamentals-independent bubbles are provided by setting $k = 0$ and $A_2 = 0$. In this case process b, given by (4.18), reduces to:

$$b(t) = A_1 \cdot \exp(rt).$$

This bubble exhibits a monotonic divergent motion as in Flood and Garber (1980) and others.

Next consider the case in which $k = r$. Given the current value of the market fundamentals, b is independent of time t from (4.18). As seen from fig. 1 or 2, these intrinsic bubbles are stochastically saddlepoint-stable.[21] To focus on this variation structure, note that in this case function b is the solution of the second-order ordinary differential equation which is obtained by setting $b = b(q)$ in (4.11):

$$(1/2)\sigma^2 b''(q) + (g - \sigma^2/2)b'(q) = rb(q). \tag{4.27}$$

Letting $v(q)$ denote $b'(q)$, we decompose this equation into:

$$\begin{aligned} v'(q) &= -\{2(g-\sigma^2/2)/\sigma^2\}v(q) + (2r/\sigma^2)b(q), \\ b'(q) &= v(q). \end{aligned} \tag{4.28}$$

These simultaneous equations determine the variations of v and b caused by fluctuations in market fundamentals q. By applying the usual phase analysis to eqs. (4.28), we can depict the saddle-shaped variation structure of intrinsic bubbles in fig. 3. The arrow attached to each trajectory indicates the direction in which each point moves along the trajectory with parametric increases in q. Although the value of q is not given parametrically but generated by stochastic process (4.20), fig. 3 is informative in the following two senses. First, given

[21] In fig. 2 intrinsic bubbles can be depicted by the 45-degree line (a linear space representing $r = k$).

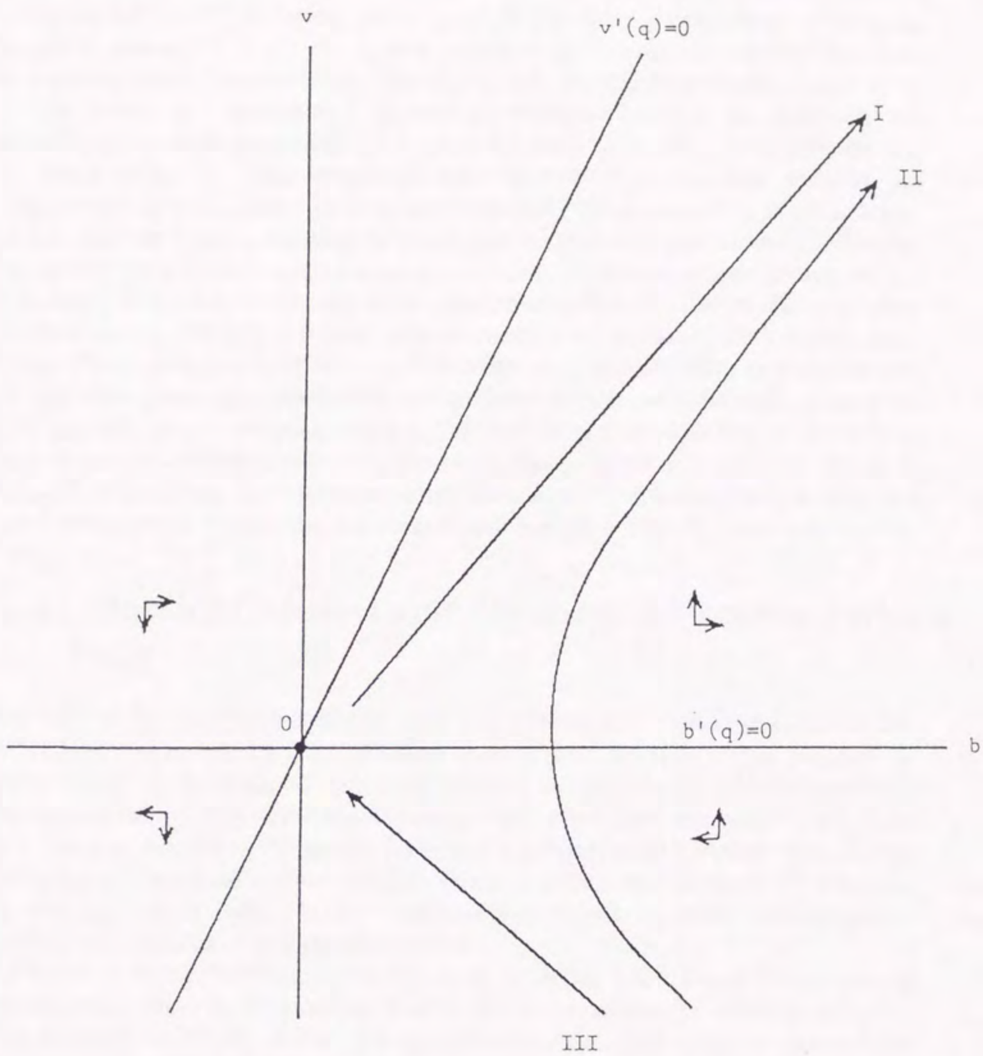


Fig. 3. Dynamics of intrinsic bubbles. Trajectory I represents the case in which $A_1>0$ and $A_2=0$; trajectory II denotes the case, $A_1>0$ and $A_2>0$; and trajectory III denotes the case, $A_1=0$ and $A_2>0$.

constants A_1 and A_2 in (4.18), each trajectory can be regarded as a transition trail on which a price bubble and its volatility are determined depending on realized values of state variable q. [From (4.25), the volatility of a bubble is proportional to v^2.] Secondly, since $g - \sigma^2/2 > 0$ by assumption, q becomes larger on average.[22] Thus, the arrows drawn in fig. 3 indicate average directions in which intrinsic bubbles move over "randomly passing time" represented by q.

As shown by trajectories I, II, and III, intrinsic bubbles can exhibit three transition patterns since both of A_1 and A_2 must be positive from proposition 1. Along trajectory I the magnitude and the volatility of intrinsic bubbles become larger as q increases. In contrast, trajectory III represents a stochastically stable bubble: b and v stochastically asymptotically converge to zero.[23] Finally, trajectory II represents a nonmonotonic bubble: b is decreasing on average while q is small, but turns increasing after q becomes sufficiently large. Furthermore, along this trajectory v monotonically increases and crosses the horizontal axis. Since the absolute value of $v(= b_q)$ is a measure of the volatility of bubbles and its sign determines how the bubble is correlated with market fundamentals q (or D), this monotonic increase in v implies, first, that the volatility of the intrinsic bubble goes down to zero, but turns increasing after q becomes sufficiently large, and secondly, that its correlation with market fundamentals is negative until the process crosses the horizontal axis, but after that it turns positive.

4.4 Partial Crashes and Stochastic Process Switching

We have so far implicitly assumed away the possibilities of market crashes. Let us extend the model by incorporating crash risks. Recalling that function b, given by (4.18), is obtained by setting it equal to the product of fundamentals-dependent factor $x(q)$ and time-dependent factor $y(t)$, we introduce crash risks in a factor-wise manner: Instead of remaining forever, at any instant factors $x(q)$ and $y(t)$ are assumed to face instantaneous constant probabilities of crashing, ρ_x and ρ_y, respectively. In our continuous-time setting, these instantaneous probabilities can take any positive values.

For b to be a price bubble, the expected evolution rate of each factor during the duration must be higher than that in the no-crash case in order to compensate for each crash risk. Without crash risks the expected evolution rate is k for $x(q)$, and $(r - k)$ for $y(t)$, as pointed out in Subsection 4.2.3. Bubble processes under crash risks are thus obtained by replacing $(r - k)$ with $(r + \rho_y - k)$ and k with $(k + \rho_x)$ in (4.17) and (4.18):

$$\begin{aligned} b(q, t; \rho_x, \rho_y) &= A_1 \exp\{(r + \rho_y - k)t + \theta_1 q\} \\ &\quad + A_2 \exp\{(r + \rho_y - k)t + \theta_2 q\}, \end{aligned} \tag{4.29}$$

where the θ_i values are the characteristic roots defined by using $k + \rho_x$ instead

[22] From eq. (4.20), as $t \to \infty$, $q(t) \to \infty$ with probability one.

[23] This case is similar to "the asymptotically bubbleless equilibrium" which Tirole (1985) has presented.

of k :[24]

$$\begin{aligned}\theta_1 &= \frac{-(g-\sigma^2/2)+\{(g-\sigma^2/2)^2+2\sigma^2(k+\rho_x)\}^{1/2}}{\sigma^2},\\ \theta_2 &= \frac{-(g-\sigma^2/2)-\{(g-\sigma^2/2)^2+2\sigma^2(k+\rho_x)\}^{1/2}}{\sigma^2}.\end{aligned} \tag{4.30}$$

The expected evolution rate of b, implied by (4.29) and (4.30), is $(r+\rho_x+\rho_y)$ while the probability that it remains for more than τ unit time is $\exp(-\rho_x\tau)\cdot\exp(-\rho_y\tau)=\exp\{-(\rho_x+\rho_y)\tau\}$. It follows that:

$$\begin{aligned}\mathrm{E}[b(t+\tau)|\Omega_t] &= \exp\{-(\rho_x+\rho_y)\tau\}\exp\{(r+\rho_x+\rho_y)\tau\}b(t)\\ &= \exp(r\tau)b(t).\end{aligned}$$

This reveals that random processes implied by (4.29) actually satisfy the definition of a price bubble, (4.9).

A remarkable feature of the present model is that the crashes of fundamentals-dependent factor x may be partial ones: When $b_i (i=1,2)$ denote two components of b in eq. (4.29); i.e., $b=b_1+b_2$, where $b_i=A_i\exp\{(r+\rho_y-k)t+\theta_i q\}$; components b_1 and b_2 do not necessarily crash simultaneously at a crash of x, but generally one of the two components remains until the second crash takes place.[25] At partial crashes, price bubbles display various stochastic process switching. To show this briefly, we now assume $\rho_y=0$ and consider the case of intrinsic bubbles, i.e., $k=r$.

Bubble dynamics are now described by equations which are obtained by replacing r with $r+\rho_x$ in (4.28). The introduction of crash risk ρ_x makes the $v'=0$ schedule and saddle trajectories in fig. 3 steeper. Viewing this, fig. 4 depicts the variation structure of intrinsic bubbles in the presence of crash risks. When a partial crash occurs on trajectory II$'$, the bubble process switches to that along saddle trajectory I$'$ or III$'$. If unstable component b_1 crashes first, the intrinsic bubble process switches to a stable one which is governed by stable root θ_2. In contrast, at the first crash of stable component b_2, the bubble process switches to an unstable one which is governed by unstable root θ_1. For example, suppose that a partial crash takes place at point C_1. If this crash pertains to unstable component b_1, the bubble instantaneously jumps from C_1 to some point on stable saddle trajectory III$'$, say point C_2, and thereafter, as q increases on average, the bubble monotonically decreases in magnitude and volatility along that trajectory until the second crash, which takes place with respect to b_2. On the other hand, if stable component b_2 crashes at point C_1, the crash destabilizes the bubble dynamics in the sense that after the crash the magnitude and volatility of the bubble follow a divergent process along trajectory I$'$ until the second crash occurs.

[24] Formally, eq. (4.29) is derived by solving the following two ordinary differential equations for $x(q)$ and $y(t)$, respectively:

$$(\sigma^2/2)x''+(g-\sigma^2/2)x'=(k+\rho_x)x, \qquad y'=(r+\rho_y-k)y.$$

[25] This can be seen by noting that the motion of component b_1 and that of b_2, which are expressed by using Ito's Lemma as:

$$\mathrm{d}b_i=(r+\rho_x+\rho_y)b_i\mathrm{d}t+\theta_i b_i\sigma\mathrm{d}z, \qquad i=1,2,$$

are both autonomous in the sense that they do not interact with each other.

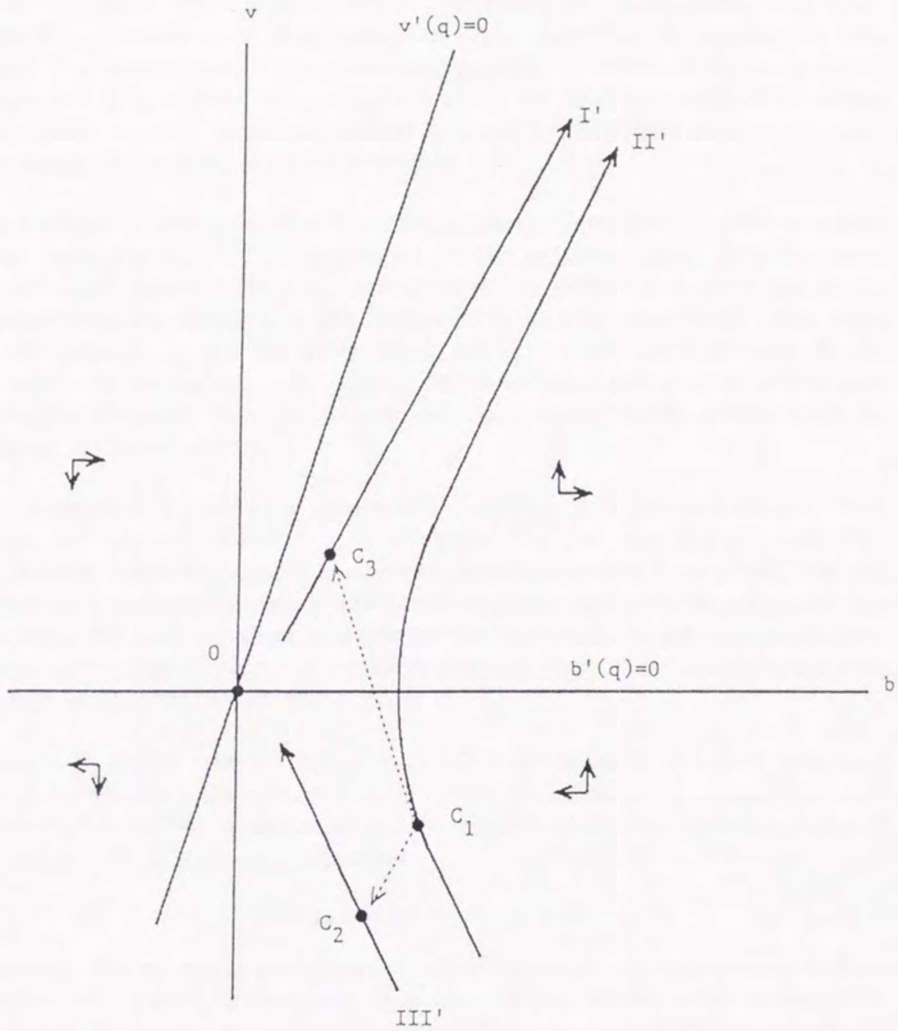


Fig. 4. Partial crashes and stochastic process switching.

In order to clarify the impact effects of partial crashes (on intrinsic bubbles), recall that the absolute value of first derivative $v(=b_q)$ determines the volatility of a bubble and its sign decides that of the correlation of the bubble and dividends. Note that v is expressed as the sum of components $\theta_1 b_1$ and $\theta_2 b_2$, which have different signs since from (4.30) $\theta_1 > 0$ and $\theta_2 < 0$ when $k = r$. The impact of a partial crash on the stochastic process of bubbles depends on which component, $\theta_1 b_1$ or $\theta_2 b_2$, crashes at that time. Let us refer to one of these two components which is larger in absolute value as the dominant component of v. It is straightforward to show the following:

Proposition 4 (*impact effects of partial crashes*): *Consider an intrinsic bubble under crash risk ρ_x. Let A_1 and A_2 be strictly positive. Then, (i) if the non-dominant component of $v(=b_q)$ crashes first, the volatility of the price bubble is instantaneously enlarged by this partial crash, as illustrated by the jump from C_1 to C_2 in fig. 4; On the other hand, (ii) if the first crash pertains to the dominant component of v, the sign of the correlation between the bubble and market fundamentals (i.e., dividends) switches instantaneously at this crash, as depicted by the jump from C_1 to C_3.*

Even when $k \neq r$ and $\rho_y \neq 0$, results (i) and (ii) in proposition 4 are also valid insofar as $k + \rho_x > 0$ since $\theta_1 b_1 > 0$ and $\theta_2 b_2 < 0$ in this case. If $\underline{k} \leq k + \rho_x < 0$, on the other hand, $\theta_1 b_1$ and $\theta_2 b_2$ are both strictly negative from (4.30). Partial crashes thus necessarily reduce the price volatility and have no effect on the correlation between a bubble and market fundamentals. In general, stochastic process switchings provided by proposition 4 take place under a large crash risk of fundamentals-dependent factor x.

Remark 3: *Crash risks ρ_x and ρ_y also affect the stability of bubble dynamics during the duration. In the same way as in Subsection 4.3.1, we can see that the stochastic stability under crash risks is decided by the relative magnitudes of $(r + \rho_y)$ and $\beta_i(k, \rho_x)(i = 1,\ 2)$, where*

$$\beta_i(k, \rho_x) = k - \theta_i(g - \sigma^2/2).$$

Naturally, $\beta_i(k, 0)$ equals $\psi_i(k)$. Fig. 5 depicts the stability map in the presence of crash risks by drawing the loci of β_i in $k-\beta_i$ space. As can easily be seen from this figure, the crash risk for time-dependent factor ρ_y destabilizes the stochastic dynamics as does r in the absence of crash risks (see fig. 2). The introduction of crash risk into fundamentals-dependent factor x enlarges the region in which bubble dynamics are stochastically saddlepoint-stable.[26]

4.5 Conclusions

Using a stochastic dividend-growth model, we have provided a general analysis of fundamentals-dependent bubbles in stock prices. Given that dividends follow a continuous Markov process, a stock price is specified as a function of dividends (i.e., market fundamentals) as well as of time. We have derived a partial differential equation with respect to this price function from an arbitrage

[26] Within a model of time-driven bubbles, Evans (1991) shows that certain partial crashes stabilize bubble dynamics.

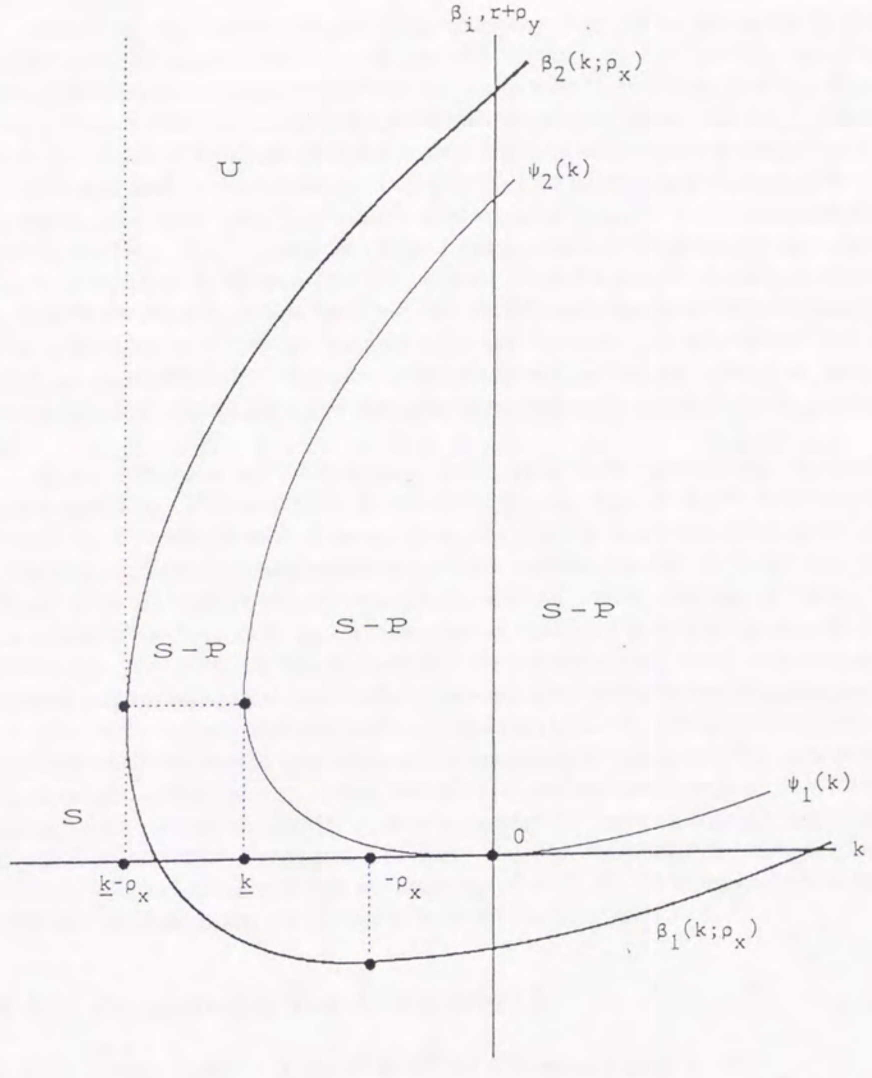


Fig. 5. Stochastic stability under crash risks.

equation. Provided that a free-disposal condition is satisfied, a fundamental price process is defined as the forward-looking particular solution of this equation, and a price bubble as the general solution of the corresponding homogeneous equation.

Reflecting the indeterminacy of the mapping from dividends to stock prices under rational expectations, price bubbles depend on the market fundamentals. This fundamentals-dependency crucially affects the basic feature of bubbles. First of all, the fundamentals-dependency stabilizes bubble dynamics: The dynamics of fundamentals-dependent bubbles can be stochastically stable, saddlepoint-stable, or unstable. Owing to this stabilizing effect, fundamentals-dependent bubbles generally satisfy a stochastic version of the transversality condition. Secondly, stock prices with fundamentals-dependent bubbles can be less volatile than fundamentals. This occurs when the price volatility is reduced by the negative correlation between the bubble and fundamental components. These bubbles may not be precluded by the variance bounds tests. Thirdly, fundamentals-dependent bubbles exhibit various transition patterns, such as nonmonotonic movements and monotonic shrinkage in magnitude and volatility.

As an extension we have incorporated crash risks, permitting bubbles to crash partially. The volatility of the price bubble may be instantaneously enlarged by a partial crash. Furthermore, the sign of the correlation between a bubble and market fundamentals can switch instantaneously at a partial crash. Crash risks also affect the stochastic stability of bubble dynamics during the duration. The crash risk for the time-driven factor of a bubble destabilizes its dynamics. The introduction of crash risk for the fundamentals-driven factor enlarges the possibility that bubble dynamics are stochastically saddlepoint-stable.

We finally suggest some directions of future research. First, our analysis can easily be extended to the case of the other asset prices. Ikeda and Shibata (1992), for example, construct a model of speculative exchange rate dynamics with fundamentals uncertainty. Secondly, it is necessary to test empirically the validity of our fundamentals-dependent bubbles. Thirdly, dropping the assumption of riskneutrality of investors, the consistency of fundamentals-dependent bubbles with the rational behavior of investors must be examined.

4.6 Appendix for Chapter 4

4.6.1 Appendix A.1: Proof of Proposition 1

Assuming that process b [given by (4.18) in the text] is nontrivial, we here prove proposition 1 by providing two lemmata. The first lemma constrains the range of k :

Lemma A.1: *Process $P(= F + b)$ becomes negative with non-zero probability when characteristic roots $\lambda_i (i = 1, 2)$ are imaginary numbers. That is,*

$$k < \underline{k} \Rightarrow \text{Prob}[P(\cdot, t) > 0,\ \forall t \in [0, \infty) | \Omega_0] < 1.$$

Proof: When $k < \underline{k}$, the λ_i, defined by (4.17) in the text, are expressed as:

$$\lambda_1 = -\pi + \varepsilon i,$$

$$\lambda_2 = -\pi - \varepsilon i,$$

where $\pi = (g - \sigma^2/2)/\sigma^2 (> 0)$ and $\varepsilon = \{|(g - \sigma^2/2)^2 + 2\sigma^2 k|^{1/2}\}/\sigma^2 (> 0)$. By using these expressions, stochastic process b [given by (4.18)] can be rewritten as:

$$b(q,t) = \exp\{(r-k)t - \pi q\}(a_1 \cos \varepsilon q + a_2 \sin \varepsilon q), \tag{4.31}$$

where $a_i (i = 1, 2)$ are arbitrary constants. Since fundamental price F [given by (4.7)] is expressed in terms of q as $F = \alpha \cdot \exp(q)$, where $\alpha = 1/(r-g)$ [> 0 from eq. (4.6)], process $P(= F + b)$ is derived from (4.31) as:

$$P(t) = \alpha \exp\{(r-k)t - \pi q\}[\exp\{(k-r)t + (\pi+1)q\} - (d_1 \cos \varepsilon q + d_2 \sin \varepsilon q)], \tag{4.32}$$

where $d_i = -a_i/\alpha (i = 1,\ 2)$. (4.32) implies that $P(t)$ is positive iff:

$$\exp\{(k-r)t + (\pi+1)q\} \geq d_1 \cos \varepsilon q + d_2 \sin \varepsilon q. \tag{4.33}$$

Assume that, for some $(d_1, d_2) \neq 0$, process P is always positive with probability one, that is, inequality (4.33) is valid for all $t \in [0, \infty)$ with probability one. Note that the R.H.S. of (4.33) represents a cycle around zero, and therefore, for any t, say t_0, one can choose q such that the R.H.S. is strictly positive. We denote this value by q_0. Now fix q at q_0 and change t parametrically from t_0 to infinity. Then, since $k - r < (\underline{k} - r <)0$ and hence the L.H.S. of (4.33) monotonically decreases to zero as $t \to \infty$ while the R.H.S. keeps constant, there exists some finite time t^* which violates (4.33) for $q = q_0$. This and the continuity of the R.H.S. of (4.33) with respect to q imply that, for $t = t^*$, there exists a range of q at the neighborhoods of q_0 such that (4.33) is not valid. This is a contradiction. □

Therefore, inequality $k \geq \underline{k}$ must hold true for (nontrivial) process b to be positive. Next, the following lemma rules out the possibilities of negative A_i :

Lemma A.2: *When characteristic roots $\lambda_i (i = 1, 2)$ are distinct real numbers, process $P(= F + b)$ is always positive with probability one if and only if the A_i values are both positive. That is, when $k > \underline{k}$,*

$$(A_1,\ A_2) \geq 0 \Leftrightarrow \text{Prob}[P(\cdot, t) \geq 0,\ \forall t \in [0, \infty) | \Omega_0] = 1.$$

Proof: Assume that $k > \underline{k}$. From (4.7) and (4.18) in the text, process $P(= F + b)$ is expressed as:

$$P(D,t) =$$
$$\alpha \cdot \exp\{(r-k)t + q\}[\exp\{(k-r)t\} + (A_1/\alpha)\exp\{(\lambda_1 - 1)q\} + (A_2/\alpha)\exp\{(\lambda_2 - 1)q\}].$$

Therefore, P is positive if and only if:

$$\exp\{(k-r)t\} \geq -(A_1/\alpha)\exp\{(\lambda_1 - 1)q\} - (A_2/\alpha)\exp\{(\lambda_2 - 1)q\}. \tag{4.34}$$

Since $(A_1,\ A_2) \geq 0$ is obviously a sufficient condition for price positivity, we now verify that if either A_1 or A_2 is strictly negative, (4.34) is not valid with non-zero probability for some $t \in [0, \infty)$.

Consider the case in which $A_1 > 0$ and $A_2 < 0$. In this case inequality (4.34) is violated for any t by sufficiently small q's (i.e., the q values which are negative and large in absolute value) since $\lambda_2 - 1 < 0$. (Note that when $\lambda_1 - 1 < 0$, $|\lambda_2 - 1| > |\lambda_1 - 1|$.)

Next, suppose that inequality (4.34) is valid with probability one for some $A_1 < 0$ and $A_2 > 0$. Obviously, if $\lambda_1 - 1 > 0$, (4.34) is violated by a sufficiently large q since $\lambda_1 - 1 > \lambda_2 - 1$ from (4.17). Therefore, we consider the case in which $(\lambda_2 - 1 <)\lambda_1 - 1 \leq 0$. In this case, for any $t \in [0, \infty)$, say t_1, one can choose a sufficiently large value for q such that the R.H.S. of (4.34) is strictly positive since $|\lambda_1 - 1| < |\lambda_2 - 1|$. Let q_1 denote this value. Now fix q at q_1 and change t parametrically from t_1 to infinity. Here note that on the L.H.S. of (4.34), $k - r < 0$ when $\lambda_1 - 1 \leq 0$. This is because from (4.17) $\lambda_1 - 1$ is increasing in k and equals zero at $k = g$ while $g < r$ from (4.6). Therefore, the L.H.S. of (4.34) monotonically decreases to zero as $t \to \infty$. It follows that there exists some finite time t^{**} which violates (4.34), implying that inequality (4.34) is not valid for $t = t^{**}$ and $q \geq q_1$. This is a contradiction. □

In lemma A.2 we do not explicitly consider the case in which $k = \underline{k}$. In this case the positivity of A_i is a sufficient, but not necessary condition for price positivity. Indeed, if $k = \underline{k}$ and $A_1 = -A_2$, $\mathrm{Prob}[P \geq 0|\Omega_0] = 1$. However, this represents a case in which $b(t)$ is a trivial process. Therefore, lemmata A.1 and A.2 imply proposition 1.

4.6.2 Appendix A.2: The Shapes of the ψ_i-Curves

By the definition of ψ_i [eq. (4.22)] we have:

$$\psi_1'(k) = 1 - (g - \sigma^2/2)\{(g - \sigma^2/2)^2 + 2\sigma^2 k\}^{-1/2},$$

$$\psi_2'(k) = 1 + (g - \sigma^2/2)\{(g - \sigma^2/2)^2 + 2\sigma^2 k\}^{-1/2}.$$

Therefore, functions ψ_i satisfy the following:

$$\begin{aligned}
\psi_1(k): \quad & \text{(i)} && \lim_{k \downarrow \underline{k}} \psi_1'(k) = -\infty, \\
& \text{(ii)} && \lim_{k \uparrow +\infty} \psi_1'(k) = 1, \\
& \text{(iii)} && \psi_1(0) = 0, \\
& \text{(iv)} && \operatorname*{argmin}_{k > \underline{k}} \psi_1(k) = 0, \\
\psi_2(k): \quad & \text{(i)} && \lim_{k \downarrow \underline{k}} \psi_2'(k) = +\infty, \\
& \text{(ii)} && \lim_{k \uparrow +\infty} \psi_2'(k) = 1, \\
& \text{(iii)} && \psi_2''(k) < 0, \forall k \in [\underline{k}, \infty).
\end{aligned}$$

The equations above imply that the ψ_i-curves can be expressed in k-ψ_i space as drawn in fig. 2.

Bibliography

[1] Arnold, L., 1973, Stochastic differential equations: Theory and applications (John Wiley & Sons, New York).

[2] Bertola, G. and L.E.O. Svensson, 1990, Stochastic devaluation risk and empirical fit of target zone models, Seminar paper no.481 (Institute for International Economic Studies, Stockholm University, Stockholm).

[3] Blanchard, O.J., 1979, Speculative bubbles, crashes and rational expectations, Economics Letters 3, 387-389.

[4] Blanchard, O.J. and M. Watson, 1982, Bubbles, rational expectations and financial markets, in: P. Wachtel, ed., Crisis in the economic and financial structure (Lexington Books, Lexington, MA).

[5] Brock, W.A., 1982, Asset prices in a production economy, in: J.J. McCall, ed., The economics of information and uncertainty (University of Chicago Press, Chicago, IL), 1-46.

[6] Buiter, W.H. and P.A. Pesenti, 1990, Rational speculative bubbles in an exchange rate target zone, Discussion paper no.479 (Centre for Economic Policy Research, London).

[7] Dezhbakhsh, H. and A. Demirguc-Kunt, 1990, On the presence of speculative bubbles in stock prices, Journal of Financial and Quantitative Analysis 25, 101-112.

[8] Diba, B.T. and H.I. Grossman, 1983, Rational asset price bubbles, NBER working paper no.1059.

[9] Diba, B.T. and H.I. Grossman, 1987, On the inception of rational bubbles, Quarterly Journal of Economics 102, 697-700.

[10] Diba, B.T. and H.I. Grossman, 1988a, Explosive rational bubbles in stock prices? American Economic Review 78, 520-530.

[11] Diba, B.T. and H.I. Grossman, 1988b, The theory of rational bubbles in stock prices, Economic Journal 98, 746-754.

[12] Evans, G.W., 1991, Pitfalls in testing for explosive bubbles in asset prices, American Economic Review 81, 922-930.

[13] Farlow, S., 1983, Partial differential equations (John Wiley & Sons, New York).

[14] Flood, R.P. and P. Garber, 1980, Market fundamentals versus price level bubbles: The first tests, Journal of Political Economy 88, 745-770.

[15] Flood, R.P. and P. Garber, 1983, A model of stochastic process switching, Econometrica 51, 537-551.

[16] Flood, R.P. and P. Garber, 1991, The linkage between speculative attack and target zone models of exchange rates, Quarterly Journal of Economics 106, 1367-1372.

[17] Froot, K.A. and M. Obstfeld, 1991a, Exchange-rate dynamics under stochastic regīme shifts: A unified approach, Journal of International Economics 31, 203-229.

[18] Froot, K.A. and M. Obstfeld, 1991b, Intrinsic bubbles: The case of stock prices, American Economic Review 81, 1189-1214.

[19] Gordon, M., 1962, The investment, financing, and valuation of the corporation (Irwin, Homewood, IL).

[20] Gray, M.R. and S.J. Turnovsky, 1979, The stability of exchange rate dynamics under perfect myopic foresight, International Economic Review 20, 643-660.

[21] Ikeda, S. and A. Shibata, 1992, Fundamental uncertainty and exchange rates, Unpublished manuscript.

[22] Kiernan, E. and D.B. Madan, 1989, Stochastic stability in macro models, Economica 56, 97-108.

[23] Krugman, P.R., 1987, Trigger strategies and price dynamics in equity and foreign exchange markets, NBER working paper no.2459.

[24] Krugman, P.R., 1991, Target zones and exchange rate dynamics, Quarterly Journal of Economics 106, 669-682.

[25] McCallum, B.T., 1983, On non-uniqueness in rational expectations models: An attempt at perspective, Journal of Monetary Economics 11, 139-168.

[26] Malliaris, A.G. and W.A. Brock, 1982, Stochastic methods in economics and finance (North-Holland, Amsterdam).

[27] Miller, M. and P. Weller, 1990, Currency bubbles which affect fundamentals: A qualitative treatment, Economic Journal 100 (supplement volume), 170-179.

[28] Miller, M., P. Weller, and J. Williamson, 1989, The stabilizing properties of target zones, in: R.C. Bryant, D.A. Curie, J.A. Frenkel, P.R. Masson, and R. Portes eds., Macroeconomic policies in an interdependent world (Brookings Institution, Washington, D.C.).

[29] Shiller, J., 1981a, Do stock prices move too much to be justified by subsequent changes in dividends?, American Economic Review 71, 421-436.

[30] Shiller, J., 1981b, The use of volatility measures in assessing market efficiency, Journal of Finance 36, 291-304.

[31] Svensson, L.E.O., 1991, Target zones and interest rate variability, Journal of International Economics 31, 27-54.

[32] Tirole, J., 1985, Asset bubbles and overlapping generations, Econometrica 53, 1071-1100.

[33] Turnovsky, S.J. and R. Weintraub, 1971, Stochastic stability of a general equilibrium system under adaptive expectations, International Economic Review 12, 1971.

[34] West, K.D., 1987, A specification test for speculative bubbles, Quarterly Journal of Economics 102, 553-580.

[35] Williams, J.B., 1938, The theory of investment (Harvard University Press, Cambridge, MA).

Chapter 5

Fundamentals Uncertainty, Bubbles, and Exchange Rate Dynamics

Abstract: Using a monetary model of exchange rate determination, we study exchange rate dynamics with bubbles which depend on stochastic market fundamentals. These dynamics can be either stochastically stable or unstable; and either monotonic or non-monotonic (including cyclic). In an extreme case, they converge with probability one and exhibit cyclic movements. Implications for the analysis of time-dependent regime shifts are also explored. Exchange rates with bubbles are likely to appear less volatile than the fundamentals in finite samples. Both the variance bounds and cointegration tests might thus be ineffective in testing the absence of bubbles under fundamentals uncertainty.

JEL Classification Number: D84, E44, F31, G15.

Keywords: Exchange rates, fundamentals uncertainty, bubbles, cycles, stability, stochastic process switching, volatility, cointegration.

5.1 Introduction

This paper treats three topics concerning exchange rate dynamics in a stochastic environment.[1] First, the stochastic properties of speculative exchange rate dynamics caused by rational expectations under fundamentals uncertainty are examined. Second, the exchange rate process obtained is used to derive implications for the analysis of stochastic regime shifts. Third, the results raise questions about the power of empirical tests such as the variance bounds and cointegration tests for asset price bubbles.

Our research may be viewed as a logical next step to an extensive line of theory on stochastic regime shifts, for example, Flood and Garber (1983, 1992), Miller et al. (1989), Klein (1990), Krugman (1991), Froot and Obstfeld (1991a,b), and Svensson (1991a,b, 1992). These studies verify that, in the

[1] For recent stylized facts concerning the behavior of exchange rates, see Levich (1985), MacDonald (1988), Baillie and McMahon (1990) and Meese (1990).

presence of stochastic process switching of market fundamentals, the exchange rate deviates from its fundamental free-float level and the deviation depends on the current state of the fundamentals. This fundamentals-dependent deviation from the fundamental free-float rate, however, is not limited to the case of stochastic regime shifts: even when the exchange rate floats freely, it may also deviate from its fundamental value under fundamentals uncertainty, reflecting the indeterminacy of the mapping from market fundamentals to the exchange rate under rational expectations. By specifying the exchange rate as a function of market fundamentals as well as time, a larger class of rational expectations solutions describing erratic exchange rate dynamics can be considered.

Applying this basic idea to the standard monetary model of exchange rate determination, we characterize the erratic behavior of exchange rates under the free float by bubbly dynamics which are generated by market fundamentals uncertainty. It is shown that the free-float exchange rate is likely to appear less volatile than market fundamentals in the presence of bubbles which depend on stochastic fundamentals: the bubbles can be stochastically asymptotically stable, that is, they can converge with probability one; and the equilibrium free-float exchange rate can exhibit non-monotonic movements, especially cyclic movements. The basic model is extended to a multiple stochastic fundamentals case. We show that interactions among multiple fundamentals-dependent factors cause exchange rates to display richer dynamics; and that, in this case, even bubbles depending exclusively on market fundamentals, which correspond to Froot and Obstfeld's (1991c) intrinsic bubbles, can exhibit non-monotonic, especially cyclic movements.

By using the solutions obtained, we also present a general procedure to solve problems of time-dependent stochastic process switching. The procedure is applied to examine the effect of announcing a future shift from the free-float regime into some new currency regime such as a target zone.

As implications for empirical research, it is finally pointed out that, in finite samples, the standard cointegration test developed by Campbell and Shiller (1987) and Shiller's (1981) variance bound test may not be as powerful as previously thought for bubble detection: bubbles can be present even if the hypothesis that there is no cointegration between prices and fundamentals is rejected, although the acceptance of the hypothesis implies the presence of bubbles.

There are several studies which are based on ideas similar to ours. In a present value model of stock price determination, Froot and Obstfeld (1991c) analyze intrinsic bubbles, which depend exclusively on dividends, and provide empirical evidence for the existence of these bubbles in the US stock market. Ikeda and Shibata (1992) examine qualitatively a broader class of these dividends-dependent bubbles in stock prices.[2] The present paper can be differentiated from these two studies in three ways. First, because of the simple linear structure of their models, richer price dynamics are ruled out by the non-negativity constraint on asset prices. In contrast, our model is non-linear (but log-linear) in the exchange rate, so that more complex patterns (including cycles) of exchange rate dynamics can be derived even under the non-negativity constraint. Second, in the earlier studies intrinsic bubbles display only non-cyclical movements. In this paper it is shown that even intrinsic bubbles can

[2] See also Miller and Weller (1991) who provide graphical examples of non-monotonic currency bubbles which depend on stochastic fundamentals.

be cyclic if there are multiple fundamentals. Third, implications for the effect of stochastic regime shifts are discussed.

The organization of the paper is as follows. Section 5.2 presents the basic model and derives an explicit solution to the model. In Section 5.3, the stochastic properties of free-float exchange rate dynamics under a single fundamental are examined. Section 5.4 extends the model to the multiple-fundamental case. In Section 5.5, implications of our solution for the theory of stochastic process switching and for widely-used empirical tests are discussed. Section 5.6 contains some concluding remarks.

5.2 The Model and Its Solutions

5.2.1 The Model

We use the standard monetary model of exchange rate determination, see, for example, Frenkel (1978), Bilson (1978), and Mussa (1978). The log of the exchange rate at time t, $e(t)$, equals some random variable which affects the money market, $m(t)$, plus a term proportional to the expected percentage change in the exchange rate:

$$e(t) = m(t) + (1/\alpha)\, E\left[\mathrm{d}e(t) \mid \Omega_t\right]/\mathrm{d}t, \quad \alpha > 0. \tag{5.1}$$

In (5.1), $m(t)$ may represent the difference of logs of the domestic money supply and the foreign money supply or the logs of the relative velocities of money in two countries. This variable is referred to as a market fundamental. The case of multiple fundamentals is treated in Section 5.4. The parameter $1/\alpha$ denotes the semi-elasticity of money demand with respect to the expected percentage change in the exchange rate, $E\left[\mathrm{d}e(t) \mid \Omega_t\right]$, where $E\left[\cdot \mid \Omega_t\right]$ is the expectation operator conditional on the information available at time t, Ω_t, which will be specified later.

Assume that the fundamental, m, follows a Brownian motion with drift μ and instantaneous standard deviation σ:

$$\mathrm{d}m(t) = \mu \mathrm{d}t + \sigma \mathrm{d}z(t), \quad m(0) = m_0, \tag{5.2}$$

where $z(t)$ is a standard Wiener process. From (5.2) the distribution of $m(t)$ conditional on $m(0) = m_0$ is $N\left(m_0 + \mu t, \sigma^2 t\right)$. For simplicity, it is assumed that $\mu > 0$, but this assumption does not affect any of the results.

Eqs. (5.1) and (5.2) constitute a simple model of exchange rate determination.[3] The process for stochastic fundamentals, represented by (5.2), is the source of randomness in the model. The filtration generated by this process specifies the information process, $\Omega = \{\Omega_t\}_{t=0}^{\infty}$. From (5.2), m follows a Markov process, so that the probability distributions of its future values depend only on its current realization. Given this Markov property, it is natural to consider the exchange rate as a function of the market fundamental, m. To consider a broader class of solutions, we assume that the exchange rate depends on two variables, the fundamental and time:

$$e(t) = e(m, t). \tag{5.3}$$

[3] The argument can be applied to macro models which consist of more than two stochastic differential equations. For example, it is straightforward to apply our results to a small open economy model with two stochastic shocks along the same lines as Klein (1990).

Equilibrium exchange rate dynamics are derived as a function, $e(m,t)$, which satisfies the stochastic differential equation (5.1) under fundamental dynamics (5.2).[4] As in the case of deterministic rational expectations models, the solution of the model with respect to $e(m,t)$ can be expressed as the sum of the forward-looking particular solution, $f(m,t)$, and the general solution of the homogeneous counterpart, $b(m,t)$, that is

$$e(m,t) = f(m,t) + b(m,t). \tag{5.4}$$

The parts $f(m,t)$ and $b(m,t)$ are referred to as *a fundamental exchange rate* and *a bubble process,* respectively.[5]

5.2.2 Solutions for the Equilibrium Exchange Rate

First, consider the fundamental component of the exchange rate process. By definition, it is obtained as:[6]

$$\begin{aligned} f(m,t) &= \alpha E\left[\int_t^\infty \exp\{-\alpha(s-t)\}\, m(s)\,\mathrm{d}s \mid \Omega_t\right] \\ &= \alpha E\left[\int_t^\infty \exp\{-\alpha(s-t)\}\left\{m(t) + \int_t^s \mu \mathrm{d}\tau + \int_t^s \sigma \mathrm{d}z(\tau)\right\}\mathrm{d}s \mid \Omega_t\right] \\ &= m(t) + (1/\alpha)\,\mu, \end{aligned} \tag{5.5}$$

implying:

$$\mathrm{d}f(m,t) = \mathrm{d}m = \mu\mathrm{d}t + \sigma\mathrm{d}z. \tag{5.6}$$

In words, stochastic increments of the fundamental exchange rate almost surely equal those of the fundamental process $\{m(t)\}_{t=0}^{\infty}$.

Next, the bubble component of the exchange rate process is analyzed. From the definition of the bubble process, $b(m,t)$ is obtained by solving the following differential equation:[7]

$$E[\mathrm{d}b(m,t) \mid \Omega_t]/\mathrm{d}t = \alpha b(m,t). \tag{5.7}$$

Applying Ito's lemma to $b(m,t)$ and substituting the result into (5.7) yields:

$$(1/2)\,\sigma^2 b_{mm}(m,t) + \mu b_m(m,t) + b_t(m,t) = \alpha b(m,t), \tag{5.8}$$

where the subscripts represent partial derivatives of b, for example, $b_m = \partial b/\partial m, b_{mm} = \partial^2 b/\partial m^2$, etc. By solving this partial differential equation, the bubble process can be obtained explicitly.

[4] Note that the non-negativity constraint is not binding in our log-linear model, that is, a negative value of e cannot be ruled out by the assumption of free disposability, as pointed out by Kompas and Spotton (1989, p. 329). See also Diba and Grossman (1987) and Ikeda and Shibata (1992) for the implication of this constraint in the linear arbitrage model of stock prices.

[5] These definitions are the standard ones, for example, Diba and Grossman (1987) and Froot and Obstfeld (1991c). In contrast, only bubbles with crash risks are called "bubbles" by Miller and Weller (1990). Although throughout this paper the possibility of bubble crashes is ignored, it is easy to incorporate crashes into the model as in Ikeda and Shibata (1992).

[6] Froot and Obstfeld (1991a,b) call this process the saddlepath exchange rate.

[7] As can be seen from (5.7), we can also treat sunspots by specifying component b as including an extraneous variable, x, which follows some stochastic process. We follow Froot and Obstfeld (1991c) and Ikeda and Shibata (1992) in assuming away extraneous variables as driving forces in a similar spirit to McCallum (1983).

The separation-of-variables method is applicable to (5.8) (see Farlow, 1983). Let us specify the function $b(m,t)$ in the separable from:

$$b(m,t) = x(m) \cdot y(t), \tag{5.9}$$

where x and y are unknown functions which exclusively depend on m and t, respectively. Then (5.8) can be decomposed into the following two equations:

$$(1/2)\,\sigma^2 x''(m) + \mu x'(m) = kx(m), \tag{5.10}$$

$$y'(t) = (\alpha - k)\,y(t), \tag{5.11}$$

where k is an arbitrary constant. Linear ordinary differential equations (5.10) and (5.11) can be solved as, respectively,

$$x(m) = C_1 \exp(\lambda_1 m) + C_2 \exp(\lambda_2 m), \tag{5.12}$$

$$y(t) = C_3 \exp[(\alpha - k)\,t], \tag{5.13}$$

where $C_i\,(i = 1, 2, 3)$ are constants and λ_1 and λ_2 are the eigenvalues of Eq. (5.10), namely,

$$\lambda_1 = \frac{-\mu + \left(\mu^2 + 2\sigma^2 k\right)^{1/2}}{\sigma^2}, \tag{5.14}$$

$$\lambda_2 = \frac{-\mu - \left(\mu^2 + 2\sigma^2 k\right)^{1/2}}{\sigma^2}. \tag{5.15}$$

The processes satisfying (5.12) and (5.13) are called, respectively, the fundamentals-driven factor and the time-driven factor. As specified in (5.9), a bubble solution is given as the product of these two factors:

$$b(m,t) = A_1 \exp[(\alpha - k)\,t + \lambda_1 m] + A_2 \exp[(\alpha - k)\,t + \lambda_2 m], \tag{5.16}$$

where $A_1 (= C_1 C_3)$ and $A_2 (= C_2 C_3)$ are constants.[8] Eq. (5.16) is one possible solution. The solution to (5.8) is completely characterized in Section 5.5.

Note that, from (5.10) and (5.11), the constants k and $(\alpha - k)$ represent the expected rate of change in the fundamentals-driven factor x, and the expected rate of change in the time-driven factor y, respectively. From (5.7), the bubble b grows at the expected rate of α. Therefore, k determines how much of the total rate, α, is generated by the mean dynamics of the fundamentals-driven factor. On the other hand, the eigenvalues, λ_i, which characterize the stochastic dynamics of x and, hence b, are non-linearly related to k. As is shown in the following sections, it is this non-linearity that causes this bubble solution to exhibit a large spectrum of dynamic properties depending on k.

Solution (5.16) contains two special cases which are examined in the existing literature. From (5.14) and (5.16), when $k = 0$, that is, the expected growth rate of the fundamentals-driven factor is zero, and $A_2 = 0$, the bubble process becomes the familiar time-driven bubble presented by Flood and Garber (1980). Another benchmark case is obtained by setting $k = \alpha$ in (5.16). In this case, bubbles depend exclusively on fundamentals, which correspond to Froot and Obstfeld's (1991c) intrinsic bubbles. Note that the bubble solution (5.16) cannot be obtained by combining linearly these two special solutions.

[8] More detailed derivation of this equation is given in Appendix A.1.

5.3 Stochastic Dynamics of the Speculative Exchange Rate

The speculative deviations of the exchange rate from its fundamental process are now examined. As discussed earlier, the eigenvalues, λ_i, which characterize the stochastic dynamics of bubbles, depend in a non-linear way on the expected rate of change in the fundamentals-driven factor, k. In particular, λ_i can be imaginary numbers depending on k. Defining $\underline{k}$ as

$$\underline{k} = -\frac{\mu^2}{2\sigma^2} \ (< 0),$$

then, from (5.14) and (5.15), λ_i are imaginary numbers if $k < \underline{k}$, otherwise they are real numbers. These two cases are examined in turn.

5.3.1 The Case of Real Roots ($k \geqslant \underline{k}$)

First, the stochastic stability of bubble dynamics in the case of real characteristic roots is examined.[9] Suppose that $k \geqslant \underline{k}$ and $A_1 \cdot A_2 \neq 0$. Time, t, drives a price bubble in two ways; by driving directly the time-driven factor, y, and by evolving the fundamentals-driven factor, x, though changing m randomly. The stochastic stability of bubbles is determined by the relative magnitudes of these two forces.

Formally, since $m(t)$ is obtained from (5.2) as:

$$m(t) = m_0 + \mu t + \sigma z(t), \tag{5.17}$$

(5.16) can be rewritten as:

$$b(m,t) = \sum_{i=1}^{2} A_i \cdot \exp(\lambda_i m_0) \cdot \exp\left[\{(\alpha - k) - (-\mu\lambda_i) + \lambda_i \sigma z(t)/t\} t\right]. \tag{5.18}$$

As discussed in Section 5.2, $\alpha - k$ in (5.18) represents the growth rate of y while the other terms in the exponential denote the random growth rate of x. Note that $\lim_{t \to \infty}[z(t)/t] = 0$ (w.p.1) by the strong law of large numbers (see Arnold, 1973, p.46). The stochastic stability of (5.18) thus depends on the relative magnitudes of the two exponents; the (negative) exponent which represents the driving force of x generated by the trend dynamics of m, $(-\mu\lambda_i)$, and the exponent which denotes the driving force of y, $(\alpha - k)$.

Fig.1 depicts this property by drawing the loci of $g(k) = (\alpha - k)$ and $h_i(k) = -\mu\lambda_i \ (i = 1, 2)$, where interval S represents the set of the values of k for which the bubble dynamics are stochastically asymptotically stable; the interval $S-P$ represents the stochastically saddle-point stable region; and the interval U represents the stochastically unstable region.[10] From (5.14) and (5.15), the h_1-curve is downward sloping and passes through the origin whereas the h_2-curve

[9] If bubble $b(t)$ converges with probability one (w.p.1) as $t \to \infty$ [for any $b(0)$], then it is said to be stochastically (or almost surely) asymptotically stable (in the main). Note that mean stability implies stochastic stability but the converse is not true. See Arnold (1973, chap. 11).

[10] Stable bubbles may be counter-intuitive since by definition (5.7) the bubbles diverge in mean. As is discussed by Froot and Obstfeld (1991c), a skewness in the distribution of the bubble's realizations makes it possible: the stable bubble will be far from zero relatively rarely in large samples, but when it is, it diverges by an amount large enough to equalize the mean growth rate of the bubble to the semi-elasticity of money demand.

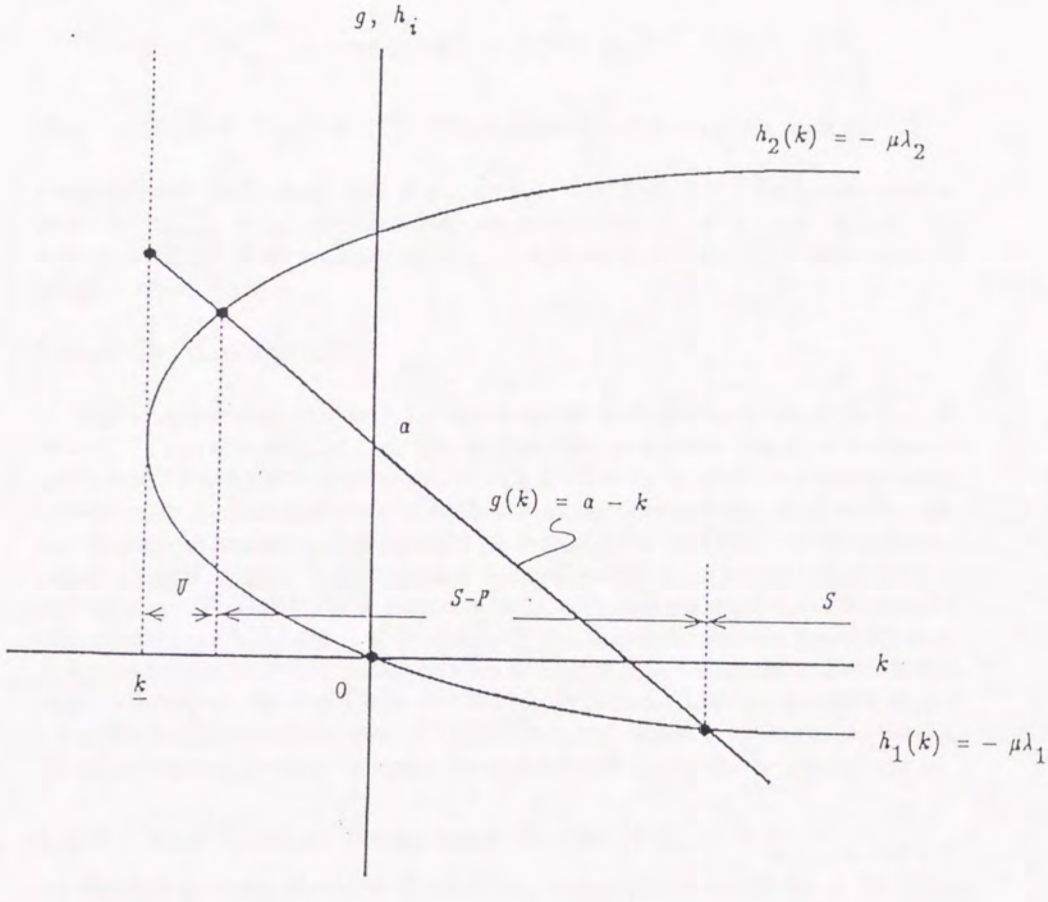


Fig. 1. Stochastic stability for given a.

has a positive gradient. These two curves are connected with each other at the point $k = \underline{k}$ where the gradients of both curves are infinite. Since the g-line and h_i-curves necessarily cross at two points as depicted in Fig.1, a large spectrum of stability is possible.

It is assumed in Fig.1 that $g(\underline{k}) > h_i(\underline{k})$ and hence $\alpha > \mu^2/\sigma^2$ by the definition of $\underline{k}$. If α is sufficiently small, there is no unstable region in this figure, implying that stochastic stability depends on the magnitude of α as well as k. In order to consider the stochastic stability in k-α space, (5.18) is rewritten as follows:

$$b(m,t) = \sum_{i=1}^{2} A_i \cdot \exp(\lambda_i m_0) \cdot \exp\left[\{\alpha - \psi_i(k) + \lambda_i \sigma z(t)/t\} t\right],$$

where $\psi_i(k) = k - \lambda_i \mu \; (i = 1, 2)$. This implies the following proposition:

Proposition 1: *Suppose that* $k \geqslant \underline{k}$ and $A_1 \cdot A_2 \neq 0$. *Then the bubble process* $\{b(m(t), t)\}_{t=0}^{\infty}$ *is* (i) *stochastically asymptotically stable if* $\alpha < \psi_1(k)$; (ii) *stochastically saddlepoint stable if* $\psi_1(k) \leqslant \alpha < \psi_2(k)$; *and* (iii) *stochastically unstable if* $\psi_2(k) \leqslant \alpha$.

Proof: See Appendix A.2.□

Fig.2 depicts this property by drawing the loci of functions ψ_i in $k - \psi_i$ space.[11,12] As is easily verified, the ψ_1 function is represented by a U-shaped curve which has a turning point at the origin. The ψ_2 function is monotonically increasing in k. The gradients of both curves become infinite (in absolute values) at point $\underline{k}$ and approach one as $k \to +\infty$. Given that the vertical axis also measures α, the region S represents the set of points (k, α) under which bubble dynamics are stochastically asymptotically stable, the region $S - P$ denotes the stochastically saddlepoint stable area, and the region U denotes the stochastically unstable area. It can easily be seen from this figure that, for a given k, the larger α becomes, the less stable are bubble dynamics. This is because a larger α implies a high evolution rate of time-dependent factor y while the λ_i values, the characteristic roots of fundamentals-dependent factor x, do not rely on α.

5.3.2 The Case of Imaginary Roots ($k < \underline{k}$)

Let us next examine the case of imaginary characteristic roots ($k < \underline{k}$). The roots given by (5.14) and (5.15) can be expressed as:

$$\begin{aligned} \lambda_1 &= -p + qi, \\ \lambda_2 &= -p - qi, \end{aligned}$$

[11] Even if one of the A_i equals zero, the argument holds with only slight modifications.
[12] The gradients of the ψ_i curves are:

$$\psi_1' = 1 - \mu / \left[\left(\mu^2 + 2\sigma^2 k\right)^{1/2} \right] \text{ and } \psi_2' = 1 + \mu / \left[\left(\mu^2 + 2\sigma^2 k\right)^{1/2} \right].$$

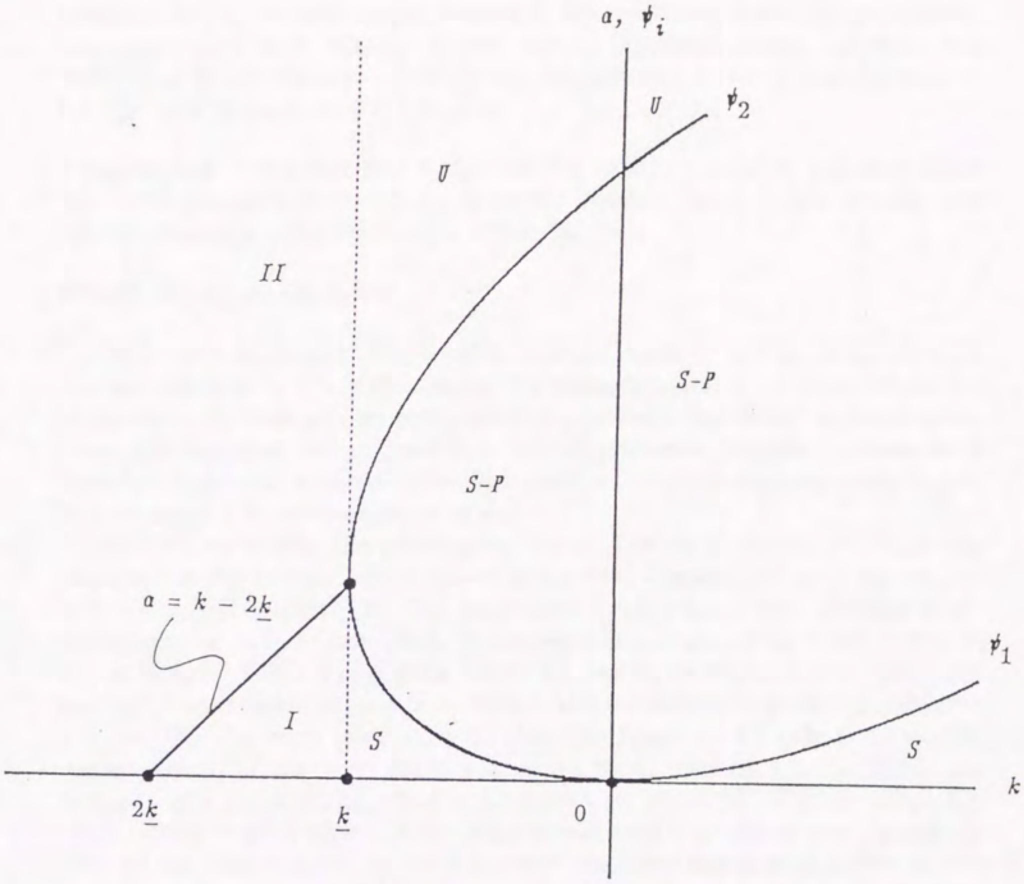


Fig. 2. Stochastic stability.

where $p = \mu/\sigma^2$ (> 0) and $q = \left[\left|\mu^2 + 2\sigma^2 k\right|^{1/2}\right]/\sigma^2$ (> 0). Using (5.17), (5.16) is thus rewritten as:

$$\begin{aligned} b(m,t) &= [a_1 \cos(qm) + a_2 \sin(qm)] \cdot \exp[(\alpha - k)t - pm] \\ &= [a_1 \cos(qm) + a_2 \sin(qm)] \cdot \exp(-pm_0) \\ &\cdot \exp[\{\alpha - (k + p\mu)\}t - p\sigma z(t)], \end{aligned} \tag{5.19}$$

where a_1 and a_2 are arbitrary constants.[13] By the strong law of large numbers, $\lim_{t\to\infty}[z(t)/t] = 0$ (w.p.1). In the case of imaginary roots, therefore, the stability of bubbles is determined by the relative magnitudes of α and $k + p\mu = k - 2\underline{k}$. This amounts to the following:

Proposition 2: *Suppose that* $k < \underline{k}$ *and that either* a_1 or a_2 is non-zero. *Then the bubble process exhibits* (i) *stochastically unstable cycles if* $\alpha \geqslant k - 2\underline{k}$; *and* (ii) *stochastically stable cycles if* $\alpha < k - 2\underline{k}$.

Proof: See Appendix A.3.□

These arguments are illustrated in regions I and II in Fig. 2 by drawing the line segment $\alpha = k - 2\underline{k}$. Region I represents pairs (k, α) which yield the cyclic stable bubble process and region II represents the cyclic unstable area. From this figure it can be seen that bubble processes become unstable as k decreases and/or α increases. This is because$\alpha - k$ represents the growth rate of y although p is independent of α and k.

In order to clarify the mechanism, three illustrative curves of (5.19) are presented in Fig.3. First, the shape of $a_1 \cos(qm) + a_2 \sin(qm)$ as a function of m is illustrated by curve (a). The term $a_1 \cos(qm) + a_2 \sin(qm)$ exhibits cyclic movements as m increases. Next, incorporate the term $\exp[(\alpha - k)t - pm]$ to the cycle as in (5.19). For a given value of t, say t_0, an increase in m stabilizes the cyclic movements since $p > 0$. This stable behavior for given t is depicted by curve (b). For each value of m, on the other hand, an increase in t enlarges the amplitude of the cycle since now $\alpha - k > 0$. Setting $t = t_1$ ($> t_0$), for example, this destabilizing effect is illustrated by curve (c). The stability of a cyclic bubble is determined by the relative magnitudes of the former stabilizing effect of the fundamentals-driven factor and the latter destabilizing effect of the time-driven factor. In case (ii) of Proposition 2, the former effect dominates the latter, so that the bubbles converge with probability one as time passes.

The exchange rate exhibits only monotonic movements in most of the monetary models with rational expectations see, for example, Frenkel (1978) and Mussa (1978). This property comes from the conventional formulation that the exchange rate depends exclusively on time. As shown here, if the exchange rate depends on both random fundamentals and time, its equilibrium behavior can be non-monotonic, and in particular cyclical. Moreover, although in rational expectations models bubble solutions are usually disposed of by requiring that the solution be bounded (e.g. Sargent and Wallace, 1973; Dornbusch, 1976), our fundamentals-dependent bubbles cannot be ruled out by this conventional requirement because the bubbles can converge with probability one.

[13] Note in (5.19) that the term $a_1 \cos(qm) + a_2 \sin(qm)$ varies within the range $\left[-\left(a_1^2 + a_2^2\right)^{1/2}, \left(a_1^2 + a_2^2\right)^{1/2}\right]$ since $[a_1 \cos(qm) + a_2 \sin(qm)] = \left(a_1^2 + a_2^2\right)^{1/2} \sin(qm + \beta)$, where β satisfies $\sin\beta = a_1/\left(a_1^2 + a_2^2\right)^{1/2}$.

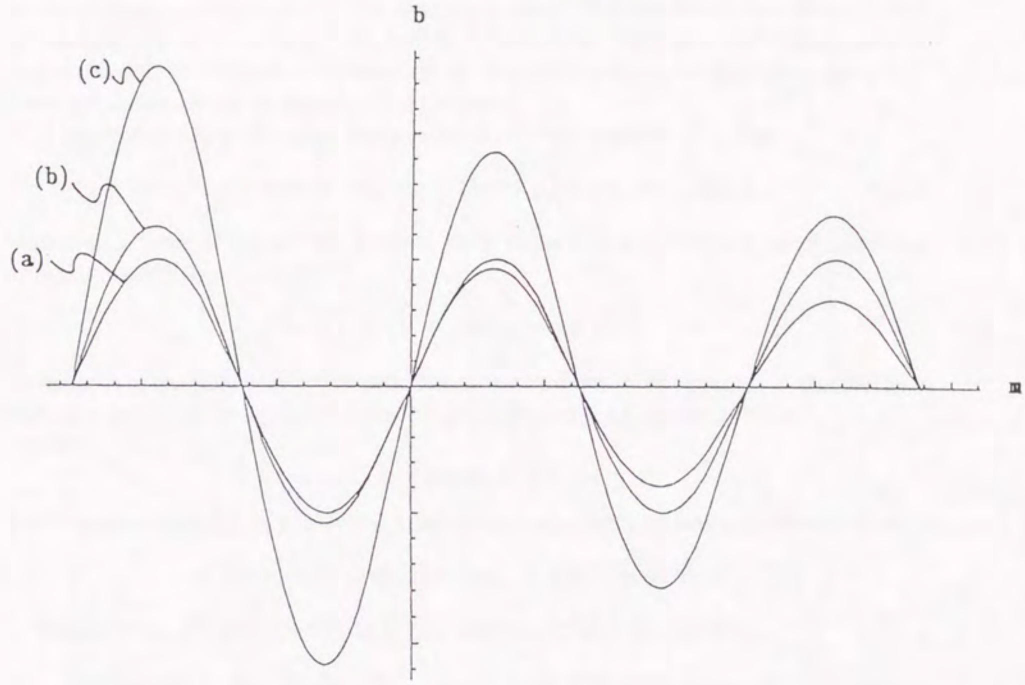


Fig. 3. Cyclic bubbles.

(a): $[a_1 cos(qm) + a_2 sin(qm)]\cdot\exp[(a-k)t_0]$

(b): $[a_1 cos(qm) + a_2 sin(qm)]\cdot\exp[(a-k)t_0-pm]$

(c): $[a_1 cos(qm) + a_2 sin(qm)]\cdot\exp[(a-k)t_1-pm]$

5.4 Multiple Fundamentals

Since, from (5.16), intrinsic bubbles correspond to the case of $k = \alpha$ $(> 0 > \underline{k})$, the cyclic bubbles in Section 5.3, which are obtained when $k < \underline{k}$, are not intrinsic bubbles. Does this imply that any intrinsic bubbles move non-cyclically? Using a multiple-fundamental model, we can derive a negative answer. The point is that, in contrast to the case of a single fundamental, in which cyclic movements of bubbles are generated by interactions between the fundamentals- and time-driven factors, interactions among multiple fundamentals-driven factors can cause cyclic dynamics in bubbles.

Let us introduce another fundamental, v, and replace (5.1) by

$$e(t) = m(t) + v(t) + (1/\alpha) E[\mathrm{d}e(t) \mid \Omega_t]/\mathrm{d}t, \tag{5.20}$$

where v follows a Brownian motion with constant drift $\nu\,(> 0)$ and constant diffusion coefficient $\delta\,(> 0)$:

$$dv(t) = \nu \mathrm{d}t + \delta \mathrm{d}w, \quad v(0) = v_0. \tag{5.21}$$

Here, w is a standard Wiener process. For the sake of simplicity, it is assumed that the two fundamentals, m and v, are independent of each other.

Set:

$$e(m, v, t) = f(m, v, t) + b(m, v, t).$$

As in Subsection 5.2.2, the fundamental exchange rate process is obtained as:

$$f(m, v, t) = m(t) + v(t) + (1/\alpha)(\mu + v).$$

From (5.2), (5.20), and (5.21), the bubble, $b(m, v, t)$, satisfies:

$$\begin{aligned} \alpha b(m, v, t) &= (1/2)\sigma^2 b_{mm}(m, v, t) + (1/2)\delta^2 b_{vv}(m, v, t) \\ &\quad + \mu b_m(m, v, t) + \nu b_v(m, v, t) + b_t(m, v, t). \end{aligned} \tag{5.22}$$

Again, the bubble solution can be obtained by the separation-of-variables method. First set

$$b(m, v, t) = x_1(m) \cdot x_2(v) \cdot y(t), \tag{5.23}$$

then (5.22) is decomposed into:

$$\begin{aligned} (1/2)\sigma^2 x_1''(m) + \mu x_1'(m) &= k x_1(m), \\ (1/2)\delta^2 x_2''(v) + \nu x_2'(v) &= (h - k) x_2(v), \\ y'(t) &= (\alpha - h) y(t), \end{aligned}$$

where k and h are the expected growth rate of the m-dependent factor, $x_1(m)$, and that of the product of the two fundamentals-dependent factor, $x_1(m) \cdot x_2(v)$, respectively. Each function is solved as:

$$\begin{aligned} x_1(m) &= A_1 \exp(\lambda_1 m) + A_2 \exp(\lambda_2 m), \\ x_2(v) &= B_1 \exp(\eta_1 v) + B_2 \exp(\eta_2 v), \\ y(t) &= C \cdot \exp\{(\alpha - h)t\}, \end{aligned}$$

where λ_i $(i = 1, 2)$ are given by (5.14) and (5.15), and η_i are given by:

$$\eta_1 = \frac{-\nu + \left[\nu^2 + 2\delta^2 (h - k)\right]^{1/2}}{\delta^2}, \tag{5.24}$$

$$\eta_2 = \frac{-\nu - \left[\nu^2 + 2\delta^2 (h-k)\right]^{1/2}}{\delta^2}. \tag{5.25}$$

Therefore, from (5.23), we obtain:

$$\begin{aligned} b(m, v, t) &= C \cdot \exp\{(\alpha - h) t\} \cdot [A_1 \exp(\lambda_1 m) + A_2 \exp(\lambda_2 m)] \\ &\quad \times [B_1 \exp(\eta_1 v) + B_2 \exp(\eta_2 v)]. \end{aligned} \tag{5.26}$$

From (5.26) the process b is an intrinsic bubble if and only if $\alpha = h$ so that it depends exclusively on the two fundamentals,[14] and is given by:

$$\begin{aligned} b(m, v) &= C_{11} \exp(\lambda_1 m + \eta_1 v) + C_{12} \exp(\lambda_1 m + \eta_2 v) \\ &\quad + C_{21} \exp(\lambda_2 m + \eta_1 v) + C_{22} \exp(\lambda_2 m + \eta_2 v), \end{aligned} \tag{5.27}$$

where $C_{ij} = A_i B_j$ $(i, j = 1, 2)$. (See Appendix A.4 for detailed discussions on the derivations of this equation and the other key relations in this section.)

This intrinsic bubble can display cyclical movements. Indeed, if $\bar{k}$ is defined as

$$\bar{k} = \alpha + \nu^2 / \left(2\delta^2\right),$$

then, from the definition of λ_i and η_i [(5.14), (5.15), (5.24), and (5.25)], the following relations are valid:

$$\begin{array}{lcl} k < \underline{k} & \Leftrightarrow & \left\{ \begin{array}{ll} \lambda_i & \text{are imaginary numbers,} \\ \eta_i & \text{are real numbers,} \end{array} \right. \\ \underline{k} \leqslant k \leqslant \bar{k} & \Leftrightarrow & \lambda_i \text{ and } \eta_i \text{ are real numbers,} \\ k > \bar{k} & \Leftrightarrow & \left\{ \begin{array}{ll} \lambda_i & \text{are real numbers,} \\ \eta_i & \text{are imaginary numbers.} \end{array} \right. \end{array}$$

Thus, a sufficiently large or small k implies cyclical intrinsic bubbles.

In order to examine the stochastic stability of the intrinsic bubble given by (5.27), consider first the case in which all roots are real ($\underline{k} \leqslant k \leqslant \bar{k}$). By similar reasoning to Section 5.3, the number of stable roots is determined by the signs of $\psi_{ij}(k)$ where $\psi_{ij}(k) = \lambda_i \mu + \eta_j \nu$ $(i, j = 1, 2)$. Although these functions may take various forms depending on the structural parameters, a typical example is illustrated in Fig. 4.[15] In the region $\underline{k} \leqslant k \leqslant \bar{k}$ it can be seen that as k increases from $\underline{k}$ to $\bar{k}$, the number of stable roots changes, in turn, from four to three, and then to two.

In the case where the λ_i are imaginary ($k \leq \underline{k}$), (5.27) is rewritten as:

$$b(m, v) = \sum_i^2 B_i \exp(-pm + \eta_i v) \cdot [a_{1i} \cos(qm) + a_{2i} \sin(qm)],$$

where $p = \mu/\sigma^2$ and $q = \left[|\ \mu^2 + 2\sigma^2 k\ |^{1/2}\right] / \sigma^2$. Thus, using the same reasoning as was used to derive proposition 2, the stochastic stability of b is determined

[14] A possible criticism of our solution is that the exchange rate should be a function of only one state variable, that is, the sum of two fundamentals, $m + v$, for $m + v$ itself follows a diffusion process. This property comes from the homogeneous fundamental processes, which is assumed only to simplify the analytical treatment. Instead, if one fundamental, m, is assumed to follow a mean reverting process: $dm = (\mu - \epsilon m)\,dt + \sigma dz$, for example, the two fundamentals, m and v, cannot be summed up to one aggregative diffusion process. In this case, the solution must depend on m and v independently (not on $m + v$).

[15] See Appendix A.5 for details concerning the curves in Fig. 4.

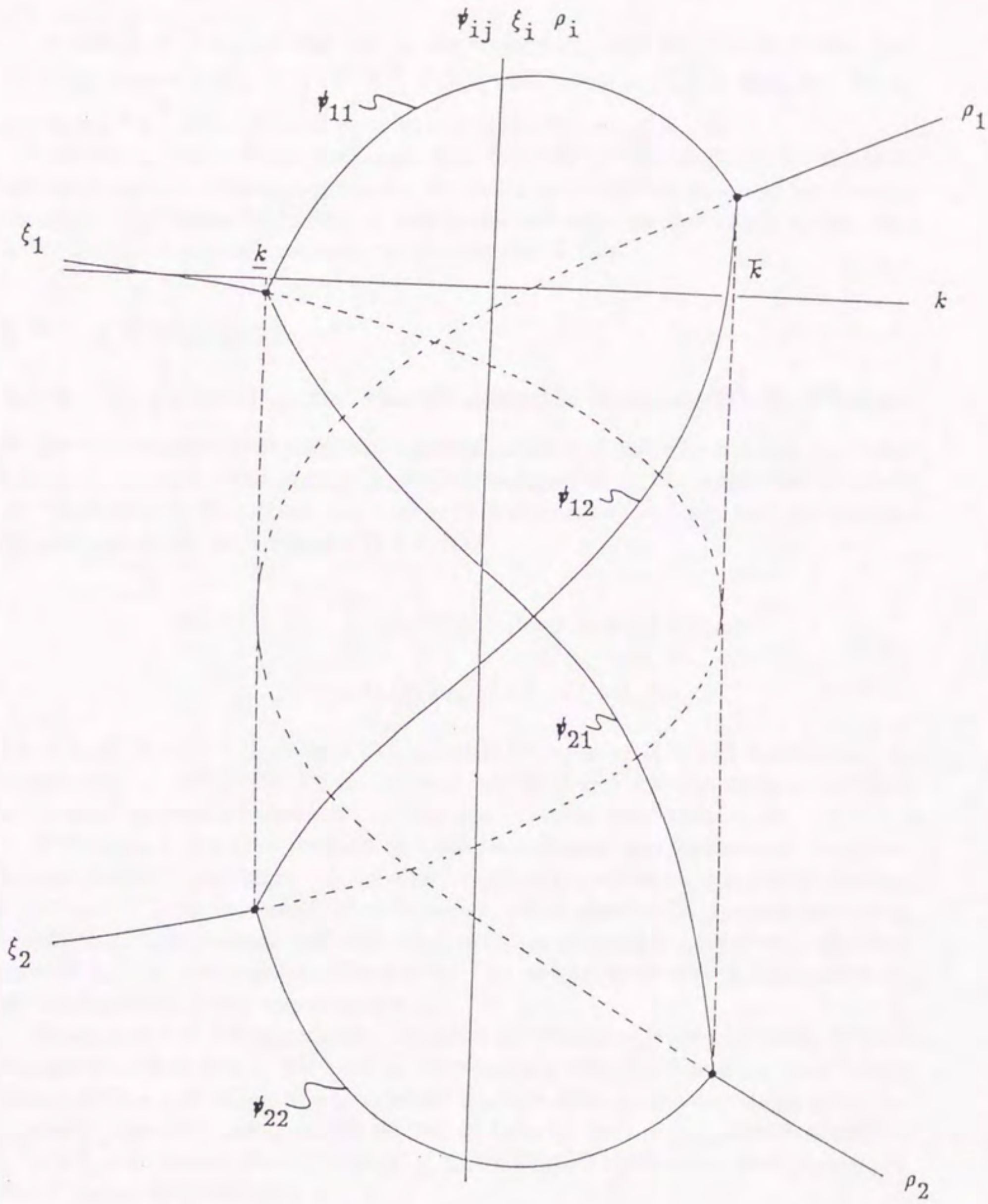


Fig. 4. Multiple fundamentals.

by the signs of $\xi_i(k)$ where $\xi_i(k) = -p\mu + \eta_i \nu$. Note that $\xi_i(\underline{k}) = \psi_{ii}(\underline{k})$. As k decreases from $\underline{k}$, the number of stable roots turns from two to one, as depicted in Fig.4.

Similarly if $k > \bar{k}$ so that the η_i are imaginary, the stability is determined by $\rho_i(k)$ where $\rho_i(k) = -(\nu^2/\delta^2) + \lambda_i \mu$, and hence $\rho_i(\bar{k}) = \psi_{ii}(\bar{k})$. These curves in Fig.4 show that there is only one stable root if $k > \bar{k}$.

Therefore, even if the exchange rate bubbles do not depend on extrinsic variables such as time and sunspots, the bubbles can either converge or diverge, showing either monotonic or non-monotonic movements (including cycles) due to interactions among fundamentals-dependent factors.

5.5 Discussion

5.5.1 Implications for the Stochastic Regime Shift Theory

By the principle of superposition, a general solution is given by a linear combination of solutions corresponding to possible values of k. In the single fundamental model presented in Section 5.2, a general solution to (5.8) can be obtained by integrating (5.16) with respect to k to give

$$\begin{aligned} b(m,t) &= \int_{-\infty}^{\infty} [A_1(k) \exp\{(\alpha - k)t + \lambda_1 m\} \\ &\quad + A_2(k) \exp\{(\alpha - k)t + \lambda_2 m\}] \, \mathrm{d}k, \end{aligned} \tag{5.28}$$

where each A_i $(i = 1, 2)$ is now defined as a function of k, which represents the distribution density. The exchange rate obtained can display various patterns of movements depending on the distribution density functions $A_i(k)$.

This general solution containing both time-driven and fundamentals-driven factors provides an insight for the analysis of macroeconomic stochastic process switching. When the effect of a time-dependent stochastic process switching of government policies is analyzed, the resulting rational expectations solutions depend on both time and fundamentals. Our solutions provide a foundation to investigate such policy experiments.

Assume $\mu = 0$ for simplicity. Consider an economy where *initially* the exchange rates float freely and then at time t_0 the government announces a future regime shift which affects the stochastic process of the future exchange rate. For example, suppose that it is announced at time t_0 that a regime characterized by $\phi(m)$ will be adopted at time $T \geq t_0$, so that the exchange rate process at time T is represented by:

$$e(T) = f(T) + \phi(m). \tag{5.29}$$

The $\phi(m)$ may be the solution to Krugman's (1991) target zone model of the solution to stochastic exchange rate pegging models of Flood and Garber (1983) and Froot and Obstfeld (1991a,b). This regime shift is supported by unrestricted intervention at time T. For example, if the exchange rate were outside a band at time T, the monetary authority would implement a discrete intervention to attain the exchange rate within the band.

Although the exchange rate process is time-independent after time T, it must be a function of t as well as m until time T, reflecting anticipations of the regime shift. In order to obtain this solution for time $t \in [t_0, T]$, consider the general solution, (5.28), within the range $k \leq \underline{k}$ ($= 0$ since $\mu = 0$). Defining $\tau = T - t$, it follows from (5.19) (where $\mu = 0$) that:

$$\begin{aligned} b(m,\tau) &= \int_{-\infty}^{0} \exp\{-(\alpha - k)\tau\}\{a_1(k)\cos(qm) + a_2(k)\sin(qm)\}\,\mathrm{d}k \\ &= \int_{0}^{\infty} \exp\{-(\alpha + \sigma^2 q^2/2)\tau\}\{c_1(q)\cos(qm) + c_2(q)\sin(qm)\}\,\mathrm{d}q, \end{aligned} \tag{5.30}$$

where $k = -\sigma^2 q^2/2$ and $c_i(q) = \sigma^2 q \cdot a_i[k(q)]$.

Since (5.30) must satisfy the boundary condition (5.29), $b(m,0) = \phi(m)$ or equivalently

$$\phi(m) = \int_{0}^{\infty} \{c_1(q)\cos(qm) + c_2(q)\sin(qm)\}\,\mathrm{d}q$$

must be valid. From the Fourier integral theorem, this equation is satisfied under some regularity condition if and only if the distribution functions, $c_i(q)$, are given by:[16]

$$c_1(q) = (1/\pi)\int_{-\infty}^{\infty} \phi(m)\cos(qm)\,\mathrm{d}m, \tag{5.31}$$

$$c_2(q) = (1/\pi)\int_{-\infty}^{\infty} \phi(m)\sin(qm)\,\mathrm{d}m. \tag{5.32}$$

By substituting (5.31) and (5.32) into (5.30), the solution for b can be obtained. Hence, the following result has been shown:[17]

Proposition 3: *Suppose that a future regime shift represented by* (5.29) *is announced at time* t_0. *Then, the equilibrium exchange rate at time* $t \in [t_0, T]$ *is given by:*

$$\begin{aligned} e(m,t) &= f(t) + \int_{0}^{\infty} \exp\{-(\alpha + \sigma^2 q^2/2)(T-t)\} \\ &\quad \cdot \{c_1(q)\cos(qm) + c_2(q)\sin(qm)\}\,\mathrm{d}q, \end{aligned}$$

where $c_i(q)\,(i = 1, 2)$ *are given by (5.31) and (5.32).*

Proof: See Appendix A.6.□

[16] In general the Fourier integral representation exists if $\phi(m)$ and $\phi'(m)$ are piecewise continuous and $\phi(m)$ is integrable over the interval $(-\infty, \infty)$. It is assumed that $\phi(m)$ satisfies these properties. Note that these conditions are satisfied in both Krugman's (1991) target zone model and Flood and Garber's (1983) model of stochastic exchange rate pegging.

[17] Ichikawa et al. (1990) obtain a closed form solution in the case of preannounced entry into the target zone of Krugman (1991). Miller and Sutherland (1992) discuss a time-dependent return to the gold standard.

In response to the announcement, the exchange rate instantaneously displays a discrete jump, and, as t goes to T, its distribution function defined over fundamentals continuously and uniformly approaches $f + \phi$, which represents the new regime. At time T the exchange rate follows the new regime with probability one. Otherwise, a discrete intervention would cause the exchange rate to jump at time T and to generate at expected infinitely large profit, which contradicts the rational expectations equilibrium.

5.5.2 Implications for Empirical Tests

In examining whether bubbles exist in asset prices, two tests are widely used; the variance bounds tests developed by Shiller (1981) and the cointegration test presented by Campbell and Shiller (1987). For example, using a monetary model Huang (1981) conducts variance bounds tests for the US dollar-mark, US dollar-sterling, and sterling-mark exchange rates and rejects the no-bubbles hypothesis. Implementing a similar test for the Australian dollar-US dollar exchange rate, Kearney and MacDonald (1990) find in favor of the no-bubbles hypothesis. MacDonald and Taylor (1991) and Gardeazabal and Regulez (1992) use cointegration tests to investigate the validity of the monetary model. If both the no-bubbles hypothesis and the monetary model are valid, fundamentals and exchange rates must be cointegrated. Using this fact MacDonald and Taylor (1991) find evidence which supports the joint hypothesis, whereas Gardeazabal and Regulez (1992) reject the monetary model. However, as is demonstrated below, these two tests, the variance bounds test and the cointegration test, may not be effective in testing no bubble hypotheses (in finite samples).[18] The results supporting the no-bubble hypothesis (Kearney and MacDonald, 1990; MacDonald and Taylor, 1991) should thus be carefully interpreted.

The Shiller test follows from the basic idea that asset price volatilities are enlarged in the presence of bubbles. However, this idea is valid only when bubbles are independent of the fundamentals. Asset prices can be less volatile even in the presence of bubbles if they are negatively correlated with the fundamentals. Indeed, by Ito's lemma we obtain from (5.4):

$$\mathrm{Var}\,(\mathrm{d}e) = (1 + b_m)^2\,\mathrm{Var}\,(\mathrm{d}m)\,,$$

implying that if $|1 + b_m| < 1$ the exchange rate is less volatile than the fundamentals. Fundamentals-dependent bubbles, therefore, may not be precluded by the variance bounds test using finite samples.

Cointegration tests suppose that, if bubbles exist, the order of non-stationarity of asset prices is higher than that of fundamentals. A cointegration relationship between asset prices and fundamentals, however, does not imply the absence of bubbles since almost every sample path of fundamentals-dependent bubbles could be convergent.[19] These stable bubbles do not appear to amplify the non-stationary of exchange rates (and other asset prices) in finite samples. Indeed,

[18] West (1987) develops another test for bubbles called a specification test. Our points do not dispute the power of the test, so that the results based on this test such a Meese (1986) would be valid even in the presence of fundamentals-dependent bubbles. See also Frankel (1985) who tests for exchange rate bubbles by computing crash risks.

[19] Even stochastically stable bubbles have necessarily explosive sample paths since they diverge in mean. However, these explosive sample paths occur extremely rarely. See Footnote 10.

Fukuta and Shibata (1993) verify this possibility through a Monte Carlo study. For specific parameter values they find that more than fifty percent of simulated asset prices with fundamentals-dependent bubbles are actually cointegrated with the fundamentals. Thus, cointegration tests might not be so powerful in testing for bubbles.[20] By introducing extrinsic crash risks, Evans (1991) also points out a potential difficulty in using cointegration tests to test for bubbles. Our results show that even without extrinsic factors fundamentals-dependency itself weakens the power of the cointegration method in testing for bubbles under fundamentals uncertainty.

As a final remark, it should be noted again that the absence of cointegration relationship between prices and fundamentals is sufficient for the presence of bubbles.

5.6 Conclusions

Using a monetary model of exchange rate determination, the dynamics of exchange rates which contain fundamentals-dependent bubbles have been analyzed. They can be either stochastically stable or unstable; either monotonic or non-monotonic, and in particular cyclic. In an extreme case, they converge with probability one, showing cyclic movements. Interactions between the fundamentals-driven factors and the time-driven factors or interactions among the fundamentals-driven factors themselves produce richer dynamics.

The general solution to the model provides a useful insight for the analysis of time-dependent stochastic process switching in currency and other asset markets. A general procedure has been presented to obtain forward-looking solutions to models containing an announcement of a future regime shift. In response to the announcement, the exchange rate instantaneously displays a discrete jump, and thereafter, as time goes to some promised date, its distribution function defined over fundamentals approaches continuously and uniformly the distribution function which represents the preannounced regime. This continuous transition is ensured by rational expectations investors, given potential discrete interventions.

A difficulty of the existing tests for bubbles in the foreign exchange market (and other asset markets) has been pointed out. Since the exchange rates with bubbles are likely to appear less volatile than the fundamentals and to appear stationary under fundamentals uncertainty in finite samples, both the variance bounds and cointegration tests which are widely used in the literature may have low power in testing no-bubble hypothesis.

[20] To be precise, the cointegration tests could detect fundamentals-dependent bubbles if infinitely many sample paths for given k could be observed. However, in reality we can observe only one sample path, and even in Monte Carlo experiments finite sample paths. In this practical context, the power of bubble tests mentioned above is weak.

5.7 Appendix for Chapter 5

5.7.1 Appendix A.1: Derivation of Eqs. in Section 5.2

We apply the separation-of-variables method to solve our fundamental partial differential equation (PDE),[21]

$$(1/2)\,\sigma^2 b_{mm}\,(m,t) + \mu b_m\,(m,t) + b_t\,(m,t) = \alpha b\,(m,t)\,. \tag{5.33}$$

Specify the function $b(m,t)$ in the separable form: $b(m,t) = x(m)\cdot y(t)$ as in eq. (5.9), where x and y are unknown functions which exclusively depend on m and t, respectively. Substitute this into eq. (5.33) to obtain:

$$(1/2)\,\sigma^2 x''\,(m)\,y(t) + \mu x'(m)y(t) + x(m)y'(t) = \alpha x(m)y(t).$$

Supposing $b \neq 0$ $(w.p.1)$, this can be rearranged by dividing the both sides by $x\cdot y(=b)$ as

$$(1/2)\,\sigma^2 x''\,(m)\,/x(m) + \mu x'(m)/x(m) = \alpha - y'(t)/y(t).$$

The resulting solution will turn out to satisfy indeed that $b \neq 0$ $(w.p.1)$. In the above equation the L.H.S. does not depend on t whereas the R.H.S. not on m. For this equality to be valid for all possible values for m and for t, therefore, the both sides must be independent of either m or t and take some constant value, say k:[22]

$$(1/2)\,\sigma^2 x''\,(m)\,/x(m) + \mu x'(m)/x(m) = k,$$
$$\alpha - y'(t)/y(t) = k,$$

which can be rearranged as:

$$(1/2)\,\sigma^2 x''(m) + \mu x'(m) = kx(m), \tag{5.34}$$

$$y'(t) = (\alpha - k)y(t). \tag{5.35}$$

These can be solved as, respectively,

$$x(m) = C_1\exp(\lambda_1 m) + C_2\exp(\lambda_2 m), \tag{5.36}$$

$$y(t) = C_3\exp\left[(\alpha - k)\,t\right], \tag{5.37}$$

where $C_i(i = 1,2,3)$ are constants and λ_1 and λ_2 are the eigenvalues of eq. (5.34), which are obtained from: $(1/2)\sigma^2\lambda^2 + \mu\lambda - k = 0$, namely

$$\lambda_1 = \frac{-\mu + (\mu^2 + 2\sigma^2 k)^{1/2}}{\sigma^2}, \tag{5.38}$$

$$\lambda_2 = \frac{-\mu - (\mu^2 + 2\sigma^2 k)^{1/2}}{\sigma^2}. \tag{5.39}$$

Combining (5.36) and (5.37), we obtain:

$$\begin{aligned} b(m,t) &= x(m)\cdot y(t) \\ &= A_1\exp\left[(\alpha - k)\,t + \lambda_1 m\right] + A_2\exp\left[(\alpha - k)\,t + \lambda_2 m\right], \end{aligned} \tag{5.40}$$

[21] For the separation-of-variables method, see Farlow (1983), Section 2.5.

[22] To be precise, parameter k can depend on some extraneous variables called sunspots. For simplicity, we ignore the possibility of sunspots as mentioned in footnote 7.

where $A_1(=C_1C_3)$ and $A_2(=C_2C_3)$ are constants.

As is aforementioned, it can be shown that solution (5.40) takes nonzero values with probability one: The solution can be rewritten as:

$$b(m,t)=\exp\{(\alpha-k)t\}\cdot[A_1\exp(\lambda_1 m)+A_2\exp(\lambda_2 m)],$$

implying that, for any finite k, $b(m,t)=0$ if and only if $A_1\exp(\lambda_1 m)=-A_2\exp(\lambda_2 m)$. When $A_1\neq 0$, this condition is equivalent to:

$$\exp\{(\lambda_1-\lambda_2)m\}=-(A_2/A_1).$$

For given A_1 and A_2, however, there exist at most a *countably* infinite number of values for m which satisfies this condition.[23] This implies that when $A_1\neq 0$, $b\neq 0\,(w.p.1)$ unless $b=0\,(w.p.1)$. When $A_1=0$, $A_2\exp(\lambda_2 m)$ and, hence, A_2 must equal zero for b to equal zero, implying that b trivially equals zero. It follows that $b\neq 0\,(w.p.1)$ unless $b=0\,(w.p.1)$.

Indeed, the function given by (5.40) is a solution to our PDE (5.33), as is shown now. By differentiating function (5.40) by m and t, we obtain:

$$b_m(m,t)=\lambda_1A_1\exp[(\alpha-k)t+\lambda_1 m]+\lambda_2A_2\exp[(\alpha-k)t+\lambda_2 m],$$

$$b_{mm}(m,t)=\lambda_1^2A_1\exp[(\alpha-k)t+\lambda_1 m]+\lambda_2^2A_2\exp[(\alpha-k)t+\lambda_2 m],$$

$$b_t(m,t)=(\alpha-k)b(m,t).$$

From these derivatives we obtain:

$$\begin{aligned}
&(1/2)\sigma^2 b_{mm}(m,t)+\mu b_m(m,t)+b_t(m,t)\\
=\;&(1/2)\sigma^2\lambda_1^2A_1\exp[(\alpha-k)t+\lambda_1 m]+(1/2)\sigma^2\lambda_2^2A_2\exp[(\alpha-k)t+\lambda_2 m]\\
&+\mu\lambda_1A_1\exp[(\alpha-k)t+\lambda_1 m]+\mu\lambda_2A_2\exp[(\alpha-k)t+\lambda_2 m]+(\alpha-k)b\\
=\;&\{(1/2)\sigma^2\lambda_1^2+\mu\lambda_1\}\cdot A_1\exp[(\alpha-k)t+\lambda_1 m]\\
&+\{(1/2)\sigma^2\lambda_2^2+\mu\lambda_2\}\cdot A_2\exp[(\alpha-k)t+\lambda_2 m]+(\alpha-k)b,\\
=\;&k\{A_1\exp[(\alpha-k)t+\lambda_1 m]+A_2\exp[(\alpha-k)t+\lambda_2 m]\}+(\alpha-k)b\\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(using the def. of }\lambda_i)\\
=\;&k\cdot b+(\alpha-k)b\\
=\;&\alpha\cdot b,
\end{aligned}$$

which proves that function (5.40) is indeed a solution to eq. (5.33).

5.7.2 Appendix A.2: Derivation of Proposition 1

We now prove proposition 1, which is concerned with the stochastic stability of b when λ_i are real numbers, i.e., $k\geq\underline{k}\;(=-\mu^2/(2\sigma^2)<0)$. It can be derived by applying straightforwardly the stability argument of Arnold (1973, pp.184-186).

Assume that $k\geq\underline{k}$. Let rewrite the solution (5.18) as:

$$b(m,t)=b_1(m,t)+b_2(m,t),$$

[23] When λ_i are distinct real roots, $\exp\{(\lambda_1-\lambda_2)m\}$ is strictly increasing in m since $\lambda_1-\lambda_2>0$, while $-(A_2/A_1)$ is a constant. Thus there exists at most unique value for m which satisfies $b=0$. When $\lambda_1=\lambda_2$, if $A_1=-A_2$ then $b=0$ $(w.p.1)$, and otherwise b cannot be zero. When λ_i are imaginary numbers, $\exp\{(\lambda_1-\lambda_2)m\}$ is a cycle defined over $m\in(-\infty,\infty)$, so that there exist at most countably infinite number of values for m which equates the value of this function to $-(A_2/A_1)$.

where $b_i = A_i{\cdot}\exp(\lambda_i m_0){\cdot}\exp\left[\{\alpha - \psi_i(k) + \lambda_i \sigma z(t)/t\}t\right]$ $(i = 1, 2)$, where $\psi_i(k) = k - \lambda_i\mu$ $(i = 1, 2)$. Note that by the strong law of large numbers [Arnold (1973, p.46) we have $\lim_{t\to\infty} [z(t)/t] = 0$ $(w.p.1)$ and, hence, $\lim_{t\to\infty} \{\alpha - \psi_i(k) + \lambda_i\sigma z(t)/t\} = \alpha - \psi_i(k)$ $(w.p.1)$. Thus the stochastic convergence of the i'th component b_i, converges almost surely to zero as $t \to \infty$ crucially depends on the relative magnitudes of $\psi_i(k)$ and α:

$$\lim_{t\to\infty} b_i = 0\ (w.p.1) \Leftrightarrow \psi_i(k) > \alpha,\ i = 1, 2,$$
$$b_i = \text{not converge } (w.p.1) \Leftrightarrow \psi_i(k) \leq \alpha,\ i = 1, 2.$$

Underlying intuition is as follows: b_i can be represented as:

$$b_i = D_i \exp[(\alpha - \psi_i)t] \cdot \exp[\lambda_i \sigma z(t)] \quad (i = 1, 2),$$

where $D_i = \exp(\lambda_i m_0)$ (a constant). That is, b_i evolves by two driving factors, $\exp[(\alpha - \psi_i)t]$ and $\exp[\lambda_i\sigma z(t)]$. Of these two factors the $\exp[\lambda_i\sigma z(t)]$ would diverge by the property of Wiener processed, while the factor $\exp[(\alpha - \psi_i)t]$ converges if $\psi_i > \alpha$ and diverges if $\psi_i < \alpha$. However, since $\lim_{t\to\infty} [z(t)/t] = 0$ $(w.p.1)$, the dynamics are dominantly governed by $\exp[(\alpha - \psi_i)t]$ as long as $\alpha \neq \psi_i$. This is intuitively because $\lambda_i\sigma z(t)$ diverges at the rate of, at most, the order of $t^{1/2}$ whereas $(\alpha - \psi_i)t$ evolves at the rate of the order of t whether it is positive or negative. It follows that the sign of $(\psi_i - \alpha)$ crucially determines the stochastic stability of b_i. [When $\alpha = \psi_i$, b_i diverges by the stochastic driving factor, $\exp[\lambda_i\sigma z(t)]$.]

Recalling that $\psi_1(k) \leq \psi_2(k)$, these relations imply:

1. $\alpha < \psi_1(k) \Leftrightarrow$both of b_1 and b_2 are convergent,
2. $\psi_1(k) \leq \alpha < \psi_2(k) \Leftrightarrow b_1$ is nonconvergent but b_2 is convergent,
3. $\psi_2(k) \leq \alpha \Leftrightarrow$both of b_1 and b_2 are nonconvergent,

which is equivalent to proposition 1.□

5.7.3 Appendix A.3: Derivation of Proposition 2

This subsection proves proposition 2, which is concerned with the stability of b in the case of imaginary roots ($k <\underline{k}$ (< 0)). In this case the roots given by (5.38) and (5.39) can be expressed as:

$$\lambda_1 = -p + qi \text{ and } \lambda_2 = -p - qi,$$

where $p = \mu/\sigma^2 (> 0)$ and $q = [|\ \mu^2 + 2\sigma^2 k\ |^{1/2}]/\sigma^2 (> 0)$. Then the solution (5.40) reduces to:

$$\begin{aligned}
b &= A_1 \exp[(\alpha - k)t + (-p + qi)m] + A_1 \exp[(\alpha - k)t + (-p - qi)m] \\
&= [A_1 \exp(qmi) + A_1 \exp(-qmi)] \cdot \exp[(\alpha - k)t - pm] \\
&= [a_1 \cos(qm) + a_2 \sin(qm)] \cdot \exp[(\alpha - k)t - pm] \text{ (by the Euler formula)} \\
&= [a_1 \cos(qm) + a_2 \sin(qm)] \cdot \exp(-pm_0) \cdot \exp[\{\alpha - (k + p\mu)\}t - p\sigma z(t)],
\end{aligned} \tag{5.41}$$

where the last equality comes from eq. (5.17).

Note in (5.41) that since $[a_1 \cos(qm)+a_2 \sin(qm)] = (a_1^2+a_2^2)^{1/2} \sin(qm+\beta)$, where β satisfies $\sin\beta = a_1/(a_1^2+a_2^2)^{1/2}$, it is valid that

$$| \, a_1 \cos(qm) + a_2 \sin(qm) \, | \leq (a_1^2 + a_2^2)^{1/2}.$$

The stability of bubbles is thus determined by the behavior of $\exp[\{\alpha - (k + p\mu)\}t - p\sigma z(t)]$. by the same argument as in Appendix A.2, therefore, the stability of b is decided by the relative magnitudes of $k + p\mu (= k - 2\underline{k})$ and α since $\lim_{t\to\infty} [z(t)/t] = 0$ $(w.p.1)$. This amounts to proposition 2. □

5.7.4 Appendix A.4: Derivations of Eqs. in Section 5.4

In this appendix we derive the bubble solution by solving the PDE under multiple fundamentals,

$$\begin{aligned} ab(m, v, t) &= (1/2)\sigma^2 b_{mm}(m, v, t) + (1/2)\delta^2 b_{vv}(m, v, t) \\ &\quad + \mu b_m(m, v, t) + \nu b_v(m, v, t) + b_t(m, v, t). \end{aligned} \tag{5.42}$$

Again the separation-of-variables method is applied. First set: $b(m, v, t) = x_1(m) \cdot x_2(v) \cdot y(t)$ and substitute it into eq. (5.42), obtaining

$$\begin{gathered} (1/2)\sigma^2 x_1''(m)x_2(v)y(t) + (1/2)\delta^2 x_2''(v)x_1(m)y(t) \\ +\mu x_1'(m)x_2(v)y(t) + \nu x_2'(v)x_1(m)y(t) + y'(t)x_1(m)x_2(v) \\ = \alpha x_1(m) \cdot x_2(v) \cdot y(t). \end{gathered}$$

Suppose $b \; (= x_1(m) \cdot x_2(v) \cdot y(t)) \neq 0$ $(w.p.1)$ and divide the both sides by b. (As in Appendix A.1, $b \neq 0$ will turn out to be satisfied by the resulting solution.) After some manipulation we obtain

$$\begin{gathered} (1/2)\sigma^2 x_1''(m)/x_1(m) + (1/2)\delta^2 x_2''(v)/x_2(v) \\ +\mu x_1'(m)/x_1(m) + \nu x_2'(v)/x_2(v) \\ = \alpha - y'(t)/y(t). \end{gathered}$$

In the above equation the L.H.S. does not depend on t whereas the R.H.S. not on either m or v. For this equality to be valid for all possible values for m, v and for t, therefore, the both sides must take some constant value, say h:

$$\begin{gathered} (1/2)\sigma^2 x_1''(m)/x_1(m) + (1/2)\delta^2 x_2''(v)/x_2(v) \\ +\mu x_1'(m)/x_1(m) + \nu x_2'(v)/x_2(v) = h, \end{gathered} \tag{5.43}$$

$$\alpha - y'(t)/y(t) = h. \tag{5.44}$$

Eq. (5.43) can be rewritten as

$$\begin{gathered} (1/2)\sigma^2 x_1''(m)/x_1(m) + \mu x_1'(m)/x_1(m) \\ = h - (1/2)\delta^2 x_2''(v)/x_2(v) - \nu x_2'(v)/x_2(v). \end{gathered}$$

Applying again the above argument to this equation of separated form, set the both sides equal to some constant, say k:

$$(1/2)\sigma^2 x_1''(m)/x_1(m) + \mu x_1'(m)/x_1(m) = k, \tag{5.45}$$

$$(1/2)\delta^2 x_2''(v)/x_2(v) + \nu x_2'(v)/x_2(v) = h - k. \tag{5.46}$$

From eqs. (5.44)-(5.46), we obtain the equations presented in the paper:

$$\begin{gathered}(1/2)\sigma^2 x_1''(m) + \mu x_1'(m) = k x_1(m),\\ (1/2)\delta^2 x_2''(v) + \nu x_2'(v) = (h-k)x_2(v),\\ y'(t) = (\alpha - h)y(t).\end{gathered}$$

By solving these ordinary equations and combining the results, the bubble solution b $(= x_1(m)x_2(v)y(t))$ is obtained as:

$$\begin{gathered}b(m,v,t) = C \cdot \exp\{(\alpha-h)t\}\cdot\\ [A_1\exp(\lambda_1 m) + A_2\exp(\lambda_2 m)][B_1\exp(\eta_1 v) + B_2\exp(\eta_2 v)],\end{gathered} \tag{5.47}$$

where λ_i are defined by eqs. (5.38) and (5.39); and η_i are the characteristic roots for (5.46) satisfying: $(1/2)\delta^2\eta^2 + \nu\eta - (h-k) = 0$, namely

$$\eta_1 = \frac{-\nu + [\nu^2 + 2\delta^2(h-k)]^{1/2}}{\delta^2}, \tag{5.48}$$

$$\eta_2 = \frac{-\nu - [\nu^2 + 2\delta^2(h-k)]^{1/2}}{\delta^2}. \tag{5.49}$$

In the same way that is used in Appendix A.1, it is easy to show that this solution takes nonzero values with probability one unless it denotes the trivial solution in that $b = 0$ $(w.p.1)$, and that (5.47) indeed satisfies PDE (5.42).

From (5.47), if $h = \alpha$, the bubble solution does not depend on t, implying that the resulting b is an intrinsic bubble, which depends exclusively on fundamentals m and v. Let us concentrate on proving that this intrinsic bubble can be cyclical as discussed in the last part of Section 5.4. When $h = \alpha$, solution (5.47) reduces to

$$\begin{aligned}b(m,v) &= C_{11}\exp(\lambda_1 m + \eta_1 v) + C_{12}\exp(\lambda_1 m + \eta_2 v) \quad (5.50)\\ &\quad + C_{21}\exp(\lambda_2 m + \eta_1 v) + C_{22}\exp(\lambda_2 m + \eta_2 v),\\ &\equiv b_{11}(m,v) + b_{12}(m,v) + b_{21}(m,v) + b_{22}(m,v), \quad (5.51)\end{aligned}$$

where $C_{ij} = A_iB_j (i,j = 1,2)$.

In the definition of η_i, (5.48) and (5.49), the term inside the squared root is positive when $k \le \bar{k}$ and strictly negative when $k > \bar{k}$, where $\bar{k} = \alpha + v^2/(2\delta^2)$. Recall on the other hand that the relative magnitudes of k and $\underline{k}$ determine whether λ_i take real numbers or imaginary ones. It follows that:

$$\begin{gathered}k < \underline{k} \Leftrightarrow \begin{cases}\lambda_i \text{ are imaginary numbers,}\\ \eta_i \text{ are real numbers,}\end{cases}\\ \underline{k} \le k \le \bar{k} \Leftrightarrow \lambda_i \text{ and } \eta_i \text{ are real numbers,}\\ k > \bar{k} \Leftrightarrow \begin{cases}\lambda_i \text{ are real numbers,}\\ \eta_i \text{ are imaginary numbers.}\end{cases}\end{gathered}$$

We check in order the stochastic stability in three possible cases: (i) $\underline{k} \le k \le \bar{k}$; (ii) $k < \underline{k}$; and (iii)$k > \bar{k}$.

Case (i): $\underline{k} \leq k \leq \bar{k}$ (λ_i and η_i are real numbers)

Consider component $b_{ij}(i = 1, 2)$ in (5.51). Since $m(t) = m_0 + \mu t + \sigma z(t)$ and $v(t) = v_0 + \nu t + \delta w(t)$, b_{ij} can be rewritten as

$$\begin{aligned} b_{ij} &= C_{ij} \exp[\lambda_i\{m_0 + \mu t + \sigma z(t)\} + \eta_j\{v_0 + \nu t + \delta w(t)\}] \\ &= C_{ij} \exp(\lambda_i m_0 + \eta_j v_0) \cdot \exp[\{(\lambda_i \mu + \eta_j v) + (\sigma z(t)/t) + (\delta w(t)/t)\}t] \\ &= D_{ij} \exp[\{\psi_{ij} + (\sigma z(t)/t) + (\delta w(t)/t)\}t], \end{aligned}$$

where $D_{ij} = \exp(\lambda_i m_0 + \eta_j\ v_0) \cdot C_{ij}$ and $\psi_{ij}(k) = \lambda_i \mu + \eta_j v$.

Applying the same argument as Section 5.3 to the above equation, the stochastic stability of component b_{ij} is crucially decided by the sign of $\psi_{ij}(k)$:

$$b_{ij} \text{ is stochastically stable.} \Leftrightarrow \psi_{ij}(k) < 0,$$

and, hence, the overall number of stable roots is determined by how the sign of $\psi_{ij}(k)$ depends on i, j. A typical example is illustrated in fig. 4. See Appendix A.5 for details of each curve in the figure.

Case (ii): $k < \underline{k}$ (λ_i are imaginary; and η_j are real)

In this case λ_i can be rewritten as: $\lambda_1 = -p + qi$ and $\lambda_2 = -p - qi$. Solution (5.50) then reduces to:

$$\begin{aligned} b(m, v) &= \exp(\eta_1 v)\{C_{11} \exp(\lambda_1 m) + C_{21} \exp(\lambda_2 m)\} \\ &\quad + \exp(\eta_2 v)\{C_{12} \exp(\lambda_1 m) + C_{22} \exp(\lambda_2 m)\} \\ &= \exp(\eta_1 v - pm)\{C_{11} \exp(qmi) + C_{21} \exp(-qmi)\} \\ &\quad + \exp(\eta_2 v - pm)\{C_{12} \exp(qmi) + C_{22} \exp(-qmi)\} \text{ (by def. of } \lambda_i) \\ &= \exp(\eta_1 v - pm)\{c_{11} \cos(qm) + c_{21} \sin(qm)\} \\ &\quad + \exp(\eta_2 v - pm)\{c_{12} \cos(qm) + c_{22} \sin(qm)\} \text{ (by the Euler formula)} \\ &= \exp(\xi_1 t + \eta_1 \delta w - p\sigma z)\{a_{11} \cos(qm) + a_{21} \sin(qm)\} \\ &\quad + \exp(\xi_2 t + \eta_2 \delta w - p\sigma z)\{a_{12} \cos(qm) + a_{22} \sin(qm)\}, \\ &\quad \text{(by def. of } m, v, \xi_i; \xi_i = -p\mu + \eta_i v) \\ &= \sum_i^2 \exp[\{\xi_i t - p\sigma z/t + \eta_i \delta w/t\}t] \cdot [a_{1i} \cos(qm) + a_{2i} \sin(qm)], \end{aligned}$$

which is equivalent to the equation presented in the paper.

Applying the same argument as was used in Appendix A.2 to the above expression, the stochastic stability of b is determined by the signs of $\xi_i(k)$ since $\lim_{t\to\infty} (z(t)/t) = 0$ ($w.p.1$) and $\lim_{t\to\infty} (w(t)/t) = 0$ ($w.p.1$). For example, the i'th component of the last expression is stochastically stable if and only if $\xi_i(k) < 0$.

Case (iii): $k > \bar{k}$ (η_i are imaginary; and λ_j are real)

In this case η_i can be rewritten as:

$$\eta_1 = -r + si \text{ and } \eta_2 = -r - si,$$

where $r = v/\delta^2$; $s = \mid v^2 + 2\delta^2(\alpha - k) \mid^{1/2} / \delta^2$. In the similar way to case (ii), solution (5.50) reduces to:

$$
\begin{aligned}
b(m,v) &= \exp(\lambda_1 m)\{C_{11}\exp(\eta_1 v) + C_{12}\exp(\eta_2 v)\} \\
&\quad + \exp(\lambda_2 m)\{C_{21}\exp(\eta_1 v) + C_{22}\exp(\eta_2 v)\} \\
&= \exp(\lambda_1 m - rv)\{C_{11}\exp(svi) + C_{12}\exp(-svi)\} \\
&\quad + \exp(\lambda_2 m - rv)\{C_{21}\exp(svi) + C_{22}\exp(-svi)\} \text{ (by def. of } \eta_i) \\
&= \exp(\lambda_1 m - rv)\{c_{11}\cos(sv) + c_{12}\sin(sv)\} \\
&\quad + \exp(\lambda_2 m - rv)\{c_{21}\cos(sv) + c_{22}\sin(sv)\} \text{ (by the Euler formula)} \\
&= \exp(\rho_1 t + \lambda_1\sigma z - r\delta w)\{a_{11}\cos(sv) + a_{12}\sin(sv)\} \\
&\quad + \exp(\rho_2 t + \lambda_2\sigma z - r\delta w)\{a_{21}\cos(sv) + a_{22}\sin(sv)\} \\
&\quad \text{(by def. of } m, v, \rho_i; \rho_i = -r\nu + \lambda_i\mu \; (= -v^2/\delta^2 + \lambda_i\mu)) \\
&= \sum_i^2 \exp[\{\rho_i t + \lambda_i\sigma z/t - r\delta w/t\}t] \cdot [a_{i1}\cos(sv) + a_{i2}\sin(sv)],
\end{aligned}
$$

which is equivalent to the equation presented in the paper.

Applying the same argument as was used in Appendix A.2 to the above expression, the stochastic stability of b is determined by the signs of $\rho_i(k)$ since $\lim_{t\to\infty}(z(t)/t) = 0$ $(w.p.1)$ and $\lim_{t\to\infty}(w(t)/t) = 0$ $(w.p.1)$.

5.7.5 Appendix A.5: The ψ_{ij}, ξ_i, and ρ_i Curves

This appendix examines the shapes of the ψ_{ij}, ξ_i, and ρ_i curves (i, j =1, 2) in Fig. 4.

By definition, the functions $\psi_{ij}(k)$ are given by:

$$
\begin{aligned}
\psi_{11}(k) &= \frac{-\mu^2+\mu\left(\mu^2+2\sigma^2 k\right)^{1/2}}{\sigma^2} + \frac{-\nu^2+\nu\left[\nu^2+2\delta^2(\alpha-k)\right]^{1/2}}{\delta^2}, \\
\psi_{12}(k) &= \frac{-\mu^2+\mu\left(\mu^2+2\sigma^2 k\right)^{1/2}}{\sigma^2} + \frac{-\nu^2-\nu\left[\nu^2+2\delta^2(\alpha-k)\right]^{1/2}}{\delta^2}, \\
\psi_{21}(k) &= \frac{-\mu^2-\mu\left(\mu^2+2\sigma^2 k\right)^{1/2}}{\sigma^2} + \frac{-\nu^2+\nu\left[\nu^2+2\delta^2(\alpha-k)\right]^{1/2}}{\delta^2}, \\
\psi_{22}(k) &= \frac{-\mu^2-\mu\left(\mu^2+2\sigma^2 k\right)^{1/2}}{\sigma^2} + \frac{-\nu^2-\nu\left[\nu^2+2\delta^2(\alpha-k)\right]^{1/2}}{\delta^2},
\end{aligned}
$$

which imply that $\psi_{11}(\bar{k}) = \psi_{12}(\bar{k})$, $\psi_{11}(\underline{k}) = \psi_{21}(\underline{k})$, $\psi_{12}(\underline{k}) = \psi_{22}(\underline{k})$ and $\psi_{21}(\bar{k}) = \psi_{22}(\bar{k})$. Thus,

$$
\begin{aligned}
\psi'_{11} &= -\psi'_{22} = \frac{\mu}{(\mu^2+2\sigma^2 k)^{1/2}} - \frac{\nu}{\left[\nu^2+2\delta^2(\alpha-k)\right]^{1/2}}, \\
\psi'_{12} &= -\psi'_{21} = \frac{\mu}{(\mu^2+2\sigma^2 k)^{1/2}} - \frac{\nu}{\left[\nu^2+2\delta^2(\alpha-k)\right]^{1/2}}.
\end{aligned}
$$

With $k < \underline{k} = -\mu^2/2\sigma^2$, the ξ_i-curves are represented as:

$$
\begin{aligned}
\xi_1(k) &= -\frac{\mu^2}{\sigma^2} + \frac{-\nu^2+\nu\left[\nu^2+2\delta^2(\alpha-k)\right]^{1/2}}{\delta^2}, \\
\xi_2(k) &= -\frac{\mu^2}{\sigma^2} + \frac{-\nu^2-\nu\left[\nu^2+2\delta^2(\alpha-k)\right]^{1/2}}{\delta^2},
\end{aligned}
$$

implying that $\xi_1(\underline{k}) = \psi_{11}(\underline{k}) \left[= \psi_{21}(\underline{k})\right]$ and $\xi_2(\underline{k}) = \psi_{12}(\underline{k}) \left[= \psi_{22}(\underline{k})\right]$, where

$$\xi_1'(k) = -\xi_2'(k) = -\frac{\nu}{\left[\nu^2 + 2\delta^2(\alpha - k)\right]^{1/2}} < 0.$$

Considering the case where $k \geqslant \bar{k}$, we obtain the ρ_i curves as:

$$\begin{aligned}
\rho_1(k) &= -\frac{\nu^2}{\delta^2} + \frac{-\mu^2 + \mu\left(\mu^2 + 2\sigma^2 k\right)^{1/2}}{\sigma^2}, \\
\rho_2(k) &= -\frac{\nu^2}{\delta^2} + \frac{-\mu^2 - \mu\left(\mu^2 + 2\sigma^2 k\right)^{1/2}}{\sigma^2}.
\end{aligned}$$

Thus, $\rho_1(\bar{k}) = \psi_{11}(\bar{k})\ [= \psi_{12}(\bar{k})]$ and $\rho_2(\bar{k}) = \psi_{22}(\bar{k})\ [= \psi_{21}(\bar{k})]$, where

$$\rho_1'(k) = -\rho_2'(k) = -\frac{\nu}{\left[\nu^2 + 2\delta^2(\alpha - k)\right]^{1/2}} > 0.$$

Let us finally show that values for the structural parameters can be chosen such that, as in Fig. 4, $\psi_{11}(\underline{k}) < 0$ and $\psi_{11}(\bar{k}) > 0$. By definition $\psi_{11}(\underline{k})$ is obtained as:

$$\psi_{11}(\underline{k}) = -\left[\left(\frac{\mu^2}{\sigma^2} + \frac{\nu^2}{\delta^2}\right) - \frac{\nu\left\{\nu^2 + 2\alpha\delta^2 + \mu^2\delta^2/\sigma^2\right\}^{1/2}}{\delta^2}\right].$$

After some manipulation, it can be shown that

$$\operatorname{sign}\psi_{11}(\underline{k}) = \operatorname{sign}\left[2\alpha\left(\frac{\nu^2}{\delta^2}\right) - \left(\frac{\mu^2}{\sigma^2}\right)^2 - \left(\frac{\mu^2}{\sigma^2}\right)\left(\frac{\nu^2}{\delta^2}\right)\right].$$

Therefore, $\psi_{11}(\underline{k})$ is negative if μ^2/σ^2 is sufficiently large and / or if α is sufficiently small.

Similarly, $\psi_{11}(\bar{k})$ is given as:

$$\psi_{11}(\bar{k}) = -\left[\left(\frac{\mu^2}{\sigma^2} + \frac{\nu^2}{\delta^2}\right) - \frac{\mu\left\{\mu^2 + 2\alpha\sigma^2 + \nu^2\sigma^2/\delta^2\right\}^{1/2}}{\delta^2}\right],$$

so that:

$$\operatorname{sign}\psi_{11}(\bar{k}) = \operatorname{sign}\left[2\alpha\left(\frac{\mu^2}{\sigma^2}\right) - \left(\frac{\nu^2}{\delta^2}\right)^2 - \left(\frac{\mu^2}{\sigma^2}\right)\left(\frac{\nu^2}{\delta^2}\right)\right].$$

Thus, a small (ν^2/δ^2) and / or a large α implies a positive $\psi_{11}(\bar{k})$. Finally, since

$$\begin{aligned}
\operatorname{sign}\psi_{11}(\underline{k}) \cdot \psi_{11}(\bar{k}) = \operatorname{sign}&\left[2\alpha\left(\tfrac{\nu^2}{\delta^2}\right) - \left(\tfrac{\mu^2}{\sigma^2}\right)^2 - \left(\tfrac{\mu^2}{\sigma^2}\right)\left(\tfrac{\nu^2}{\delta^2}\right)\right] \\
\times &\left[2\alpha\left(\tfrac{\mu^2}{\sigma^2}\right) - \left(\tfrac{\nu^2}{\delta^2}\right)^2 - \left(\tfrac{\mu^2}{\sigma^2}\right)\left(\tfrac{\nu^2}{\delta^2}\right)\right],
\end{aligned}$$

the inequalities $\psi_{11}(\underline{k}) < 0$ and $\psi_{11}(\bar{k}) > 0$ are valid if (μ^2/σ^2) is large enough while (ν^2/δ^2) is sufficiently small.

5.7.6 Appendix A.6: Derivation of Proposition 3

We first derive eq. (5.30). Assume that $\mu = 0$, so that, for $k < \underline{k}\ (= 0)$,

$$p = 0 \text{ and } q = (-2\sigma^2 k)^{1/2}/\sigma^2,$$

from the definition of p and q. The R.H.S. of solution (5.41) thus reduces to:

$$\begin{aligned}\text{The R.H.S. of (5.41)} &= [d_1\cos(qm) + d_2\sin(qm)]\cdot\exp[(\alpha - k)t] \\ &= [a_1\cos(qm) + a_2\sin(qm)]\cdot\exp[-(k-\alpha)\gamma],\end{aligned}$$

where $\gamma = T-t$; and $a_i = d_i\exp(-T(k-\alpha))$. Let us construct a general solution for $\gamma \geq 0$ by integrating the above solution over the range $k \in (-\infty, \underline{k}(=0))$:

$$b(m,\gamma) = \int_{-\infty}^{0}\exp\{-(\alpha - k)\gamma\}\{a_1(k)\cos(qm) + a_2(k)\sin(qm)\}dk,$$

where a_i are now density functions defined over k, $a_i = a_i(k)$. Note that $k = -\sigma^2 q^2/2$ by the definition of q. Thus $dk/dq = -\sigma^2 q < 0$ and, hence, k is monotonic in q (> 0). The above solution thus can be rewritten by changing the integrating valuable as

$$\begin{aligned}&b(m,\gamma)\\ =& \int_{\infty}^{0}\exp\{-(\alpha + \sigma^2 q^2/2)\gamma\}\\ &\times\{a_1(-\sigma^2 q^2/2)\cos(qm) + a_2(-\sigma^2 q^2/2)\sin(qm)\}(-\sigma^2 q^2/2)dq\\ =& \int_{0}^{\infty}\exp\{-(\alpha + \sigma^2 q^2/2)\gamma\}\{c_1(q)\cos(qm) + c_2(q)\sin(qm)\}dq,\end{aligned}$$

where $c_i(q) = \sigma^2 q\cdot a_i(-\sigma^2 q^2/2)$. This is equation (5.30).

The next step is to determine the density functions $c_i(q)$ so that (5.30) satisfies the boundary condition, (5.29): Since (5.30) must satisfy condition (5.29) at $\gamma = 0$, it must hold true that $b(m, 0) = \phi(m)$, or equivalently,

$$\phi(m) = \int_0^\infty\{c_1(q)\cos(qm) + c_2(q)\sin(qm)\}dq. \tag{5.52}$$

On the other hand, we can prove the following lemma:

Lemma A.1: *Under the regularity condition given by footnote 15, function $\phi(m)$ can be represented in the form of a version of the Fourier integral [see Farlow (1983, Section 2.11)]*

$$\phi(m) = \int_0^\infty\{f_1(q)\cos(qm) + f_2(q)\sin(qm)\}dq, \tag{5.53}$$

where $f_1(q)$ and $f_2(q)$ are Fourier sine- and cosine-transforms of $\phi(m)$, respectively:

$$
\begin{aligned}
f_1(q) &= (1/\pi)\int_{-\infty}^{\infty} \phi(m)\cos(qm)dm, \\
f_2(q) &= (1/\pi)\int_{-\infty}^{\infty} \phi(m)\sin(qm)dm.
\end{aligned} \tag{5.54}
$$

Proof: Applying the double-integral theorem of Fourier to $\phi(m)$, we obtain:

$$\phi(m) = (1/\pi)\int_0^{\infty} d\omega \int_{-\infty}^{\infty} \phi(\xi)\cdot\cos\omega(m-\xi)d\xi.$$

Here from the property of cosines, we have:

$$
\begin{aligned}
\cos\omega(m-\xi) &= \cos(\omega m - \omega\xi) \\
&= \cos(\omega m)\cdot\cos(\omega\xi) + \sin(\omega m)\cdot\sin(\omega\xi).
\end{aligned}
$$

Thus the above double-integral representation reduces to:

$$
\begin{aligned}
\phi(m) &= (1/\pi)\int_0^{\infty} d\omega \int_{-\infty}^{\infty} \phi(\xi)\cdot\cos(\omega m)\cos(\omega\xi)d\xi \\
&\quad +(1/\pi)\int_0^{\infty} d\omega \int_{-\infty}^{\infty} \phi(\xi)\cdot\sin(\omega m)\sin(\omega\xi)d\xi \\
&= \int_0^{\infty}[\int_{-\infty}^{\infty}(1/\pi)\phi(\xi)\cos(\omega\xi)d\xi]\cos(\omega m)d\omega \\
&\quad +\int_0^{\infty}[\int_{-\infty}^{\infty}(1/\pi)\phi(\xi)\sin(\omega\xi)d\xi]\sin(\omega m)d\omega \\
&= \int_0^{\infty} f_1(\omega)\cos(\omega m)d\omega + \int_0^{\infty} f_2(\omega)\sin(\omega m)d\omega,
\end{aligned}
$$

where f_i are the Fourier transforms given by (5.54). The last expression is the same as (5.53).□

Comparing (5.52) with (5.53), we can easily see that, if $c_i(q) = f_i(q)$, then the density functions necessarily satisfy (5.52) and, hence, the resulting function is a solution to the problem, We can also prove the converse to obtain the following lemma (For the proof, see the appendix of this supplement.):

Lemma A.2: $c_i(q)$ $(i = 1, 2)$ *satisfy the boundary condition (5.52) if and only if*

$$c_i(q) = f_i(q) \;\; (i = 1, 2),$$

where f_i are the Fourier transforms given by (5.54).

Proof: Let us prove the lemma provided in this supplement. From (5.52)and(5.53) we have:

$$\int_0^{\infty}\{[c_1(q) - f_1(q)]\cos(qm) + [c_2(q) - f_2(q)]\sin(qm)\}dq = 0 \;\; \forall m. \tag{5.55}$$

Differentiating (5.55) with respect to m gives

$$\int_0^\infty \{q[c_1(q) - f_1(q)]\sin(qm) - q[c_2(q) - f_2(q)]\cos(qm)\}dq = 0.$$

Differentiating this again with respect to m, we have

$$\int_0^\infty \{q^2[c_1(q) - f_1(q)]\cos(qm) + q^2[c_2(q) - f_2(q)]\sin(qm)\}dq = 0. \tag{5.56}$$

Similarly differentiating (5.55) $2n$ times yields

$$\int_0^\infty \{q^{2n}[c_1(q) - f_1(q)]\cos(qm) + q^{2n}[c_2(q) - f_2(q)]\sin(qm)\}dq = 0 \quad \forall m. \tag{5.57}$$

Evaluating (5.57) at $m = 0$, we obtain

$$\int_0^\infty q^{2n}[c_1(q) - f_1(q)]dq = 0.$$

Since this holds true for any integer $n(> 0)$, we have

$$c_1(q) - f_1(q) = 0 \quad \text{for any } q. \tag{5.58}$$

That is, functions c_1 and f_1 are exactly the same.

Substituting (5.58) into (5.57) yields

$$\int_0^\infty \{q^{2n}[c_2(q) - f_2(q)]\sin(qm)\}dq = 0 \quad \forall m.$$

Differentiating this with respect to m, we get

$$\int_0^\infty \{q^{2n+1}[c_2(q) - f_2(q)]\cos(qm)\}dq = 0 \quad \forall m.$$

Evaluating this at $m = 0$ gives

$$\int_0^\infty \{q^{2n+1}[c_2(q) - f_2(q)]\}dq = 0.$$

Since this holds valid for any n (> 0), we have

$$c_2(q) - f_2(q) = 0 \quad \text{for any } q,$$

which completes the proof.□

From this lemma, we obtain (5.31) and (5.32) in the paper. Substituting (5.31) and (5.32) into (5.30) yields proposition 3.

Indeed, solution (5.30) with a_i being given by (5.31) and (5.32) satisfies the PDE, (5.33), and the boundary condition, (5.29), as follows. First, from (5.30), the derivatives of b are given by:

$$
\begin{aligned}
b_t &= (\alpha - k)b, \\
b_m &= q \cdot \int_0^\infty \exp\{-(\alpha + \sigma^2 q^2/2)\gamma\}\{-c_1(q)\sin(qm) + c_2(q)\cos(qm)\}dq, \\
b_{mm} &= q^2 \cdot \int_0^\infty \exp\{-(\alpha + \sigma^2 q^2/2)\gamma\}\{-c_1(q)\cos(qm) - c_2(q)\sin(qm)\}dq \\
&= -q^2 b \\
&= 2kb/\sigma^2, \qquad \text{(using def. of } q\text{)}
\end{aligned}
$$

which implies:

$$
\begin{aligned}
(1/2)\sigma^2 b_{mm} + \mu b_m + b_t &= kb + (\alpha - k)b \qquad \text{(using } \mu = 0\text{)} \\
&= ab.
\end{aligned}
$$

This is equation (5.33). Second, given our solution, we have:

$$
\begin{aligned}
b(m, t = T) &= \int_0^\infty \{f_1(q)\cos(qm) + f_2(q)\sin(qm)\}dq \\
&= \phi(m),
\end{aligned}
$$

where f_i are given by (5.54); and the last equality follows from the Fourier integral theorem. This implies the boundary condition, (5.29). Therefore, the function given by (5.30), (5.31), and (5.32) is a solution to the PDE (5.33) subject to (5.29).

Bibliography

[1] Farlow, S., 1983, Partial Differential Equations (Wiley, New York, NY).

[2] Ikeda, S. and A. Shibata, 1992, "Fundamentals-Dependent Bubbles in Stock Prices," *Journal of Monetary Economics* 30, 143-168.

[3] Ikeda, S. and A. Shibata, 1995, "Fundamentals Uncertainty, Bubbles, and Exchange Rate Dynamics," forthcoming in *Journal of International Economics.*

[4] Arnold, L., 1973, Stochastic differential equations: Theory and applications (Wiley, New York, NY).

[5] Baillie, R. and P. McMahon, 1990, The foreign exchange market (Cambridge University Press, Cambridge).

[6] Bilson, J.F.O., 1978, Rational expectations and the exchange rate, in: J.A. Frenkel and H.G. Johnson, eds., The economies of exchange rates (Addison-Wesley, Reading, MA).

[7] Campbell, J. and R.J. Shiller, 1987, Cointegration and tests of present value models, Journal of Political Economy 95, 1062-1088.

[8] Diba, B.T. and H. Grossman, 1987, On the inception of rational bubbles, Quarterly Journal of Political Economics 102, 697-700.

[9] Dornbusch, R., 1976, Expectations and exchange rate dynamics, Journal of Political Economy 84, 1161-1176.

[10] Evans, G.W., 1991, Pitfalls in testing for explosive bubbles in asset prices, American Economics Review 81, 922-930.

[11] Farlow, S., 1983, Partial differential equations (Wiley, New York, NY).

[12] Flood, R.P. and P. Garber, 1980, Market fundamentals versus price level bubbles: The first tests, Journal of Political Economy 88, 745-770.

[13] Flood, R.P. and P. Garber, 1983, A model of stochastic process switching, Econometrica 51, 537-551.

[14] Flood, R.P. and P. Garber, 1992, The linkage between speculative attack and target zone models of exchange rates: Some extended results, in: P. Krugman and M. Miller, eds., Exchange rate targets and currency bands (Cambridge University Press, Cambridge).

[15] Frankel, J.A., 1985, The dazzling dollar, Brookings Papers on Economic Activity 1, 199-217.

[16] Frenkel, J.A., 1978, A monetary approach to the exchange rate, in: J.A. Frenkel and H.G. Johnson, eds., The economics of exchange rates (Addison Wesley, Reading, MA).

[17] Froot, K.A., and M. Obstfeld, 1991a, Stochastic process switching: Some simple solutions, Econometrica 59, 241-250.

[18] Froot, K.A., and M. Obstfeld, 1991b, Exchange-rate dynamics under stochastic regime shifts: A unified approach, Journal of International Economics 31, 203-229.

[19] Froot, K.A., and M. Obstfeld, 1991c, Intrinsic bubbles: The case of stock prices, American Economic Review 81, 1189-1214.

[20] Fukuta, Y. and A. Shibata, 1993, On the difficulty of testing for bubbles, mimeo.

[21] Gardeazabal, J. and M. Regulez, 1992, The monetary model of exchange rates and cointegration: Estimation, testing and prediction (Springer-Verlag, Berlin).

[22] Huang, R.D., 1981, The monetary approach to exchange rate in an efficient foreign exchange: Test based on volatility, Journal of Finance 36, 31-41.

[23] Ichikawa, M., M. Miller and A. Sutherland, 1990, Entering a preannounced currency band, Economics Letters 34, 363-368.

[24] Ikeda, S. and A. Shibata, 1992, Fundamentals-dependent bubbles in stock prices, Journal of Monetary Economics 30, 143-168.

[25] Kearney, C. and R. MacDonald, 1990, Exchange rate volatility, news and bubbles, Australian Economic Papers 70, 1-20.

[26] Klein, M.W., 1990, Playing with the band: Dynamic effects of target zones in an open economy, International Economic Review 31, 757-772.

[27] Kompas, T. and B. Spotton, 1989, A note on rational speculative bubbles, Economics Letters 30, 327-331.

[28] Krugman, P.R., 1991, Target zones and exchange rate dynamics, Quarterly Journal of Economics 106, 669-682.

[29] Levich, R.M., 1985, Empirical studies of exchange rates: Price behavior, rate determination and market efficiency, in: R.W. Jones and P.B. Kenen, eds., Handbook of international economics II (North-Holland, Amsterdam).

[30] MacDonald, R., 1988, Floating exchange rates: Theory and evidence (Unwin Hyman, London).

[31] MacDonald, R. and M.P. Taylor, 1991, The monetary approach to the exchange rate: Long-run relationships and coefficient restrictions, mimeo, International Monetary Fund.

[32] McCallum, B.T., 1983, On non-uniqueness in rational expectations models: An attempt at perspective, Journal of Monetary Economics 11, 139-168.

[33] Meese, R., 1986, Testing for bubbles in exchange markets: A case of sparkling rate? Journal of Political Economy 94, 345-373.

[34] Meese, R., 1990, Currency fluctuations in the post-Bretton Woods era, Journal of Economic Perspectives 2, 117-134.

[35] Miller, M. and A. Sutherland, 1992, Britain's returns to gold and entry into the EMS: Joining conditions and credibility, in: P. Krugman and M. Miller, eds., Exchange rate targets and currency bands (Cambridge University Press, Cambridge).

[36] Miller, M. and P. Weller, 1990, Currency bubbles which affect fundamentals: A qualitative treatment, Economic Journal 100 (supplement volume), 170-179.

[37] Miller, M., P. Weller, 1991, Exchange rate bands with price inertia, Economic Journal 101, 1380-1399.

[38] Miller, M., P. Weller and J. Williamson, 1989, The stabilizing properties of target zones, in: R.C. Bryant, D.A. Curie, J.A. Frenkel, P.R. Masson and R. Portes, eds., Macroeconomic policies in an interdependent world (Brookings Institution, Washington DC)

[39] Mussa, M. 1978, The exchange rate, the balance of payments, and monetary and fiscal policy under a regime of controlled floating, in: J.A. Frenkel and H.G. Johnson," eds., The economics of exchange rates (Addison-Wesley, Reading, MA).

[40] Sargent, T.J. and N. Wallace, 1973, The stability of models of money and growth with perfect foresight, Econometrica 41, 1043-1048.

[41] Shiller, R.J., 1981, Do stock prices move too much to be justified by subsequent changes in dividends? American Economic Review 71, 421-436.

[42] Svensson, L.O.E., 1991a, The term structure of interest rate differentials in a targets zone, Journal of Monetary Economics 28, 87-116.

[43] Svensson, L.O.E., 1991b, Target zones and interest rate variability, Journal of International Economics 31, 27-54.

[44] Svensson, L.O.E., 1992, The foreign exchange risk premium in a target zone with devaluation risk, Journal of International Economics 32, 21-40.

[45] West, K.D., 1987, A specification test for speculative bubbles, Quarterly Journal of Economics 102, 553-580.

Part III

Asset Pricing

Chapter 6

The Continuous-Time APT with Diffusion Factors and Rational Expectations: A Synthesis

Abstract: The APT is recast as a general theory of arbitrage asset valuation in a model with diffusion factors and rational expectations. Defining betas by factor elasticities of asset prices, the APT-type arbitrage-free condition is reformulated in terms of asset price function. The condition reduces to a partial differential equation with respect to the asset valuation function. The price function, as a solution of this equation, takes two alternative forms depending on how to design risk-adjustment. The resulting formulae consistently demonstrate the various existing ideas of arbitrage asset evaluation.

JEL Classification Number: G12, G13.

Keywords: Arbitrage, asset pricing, APT, options.

6.1 Introduction

Ross (1976, 1977) proposed an arbitrage theory of risk pricing in capital markets-the Arbitrage Pricing Theory (APT). The main proposition is that if asset returns conform to a factor generating model, then, by the law of one price, risk premia are determined as linear in sensitivity coefficients of returns to factors. With a k factor structure this linear relation asserts the existence of a set of k constants which evaluate factor risks. Regarded as an alternative to the Capital Asset Pricing Model (CAPM), the APT has so far been made rigorous and extended by many subsequent works.[1]

[1] For example, Huberman (1982) and Ingersoll (1984) gave proofs to the APT by rigorously reformulating the definition of arbitrage in infinite economies. Dybvig (1983) and Grinblatt and Titman (1983) discussed the mispricing upper bounds of the APT in settings with finitely many assets. See also Solnik (1983) and Ikeda (1991) for extensions of the APT to international asset pricing.

However, we can find some room to improve in the theory. First of all, the APT explored the risk-premium determination without saying anything about asset valuation. The risk premium of an asset, however, must be consistent with its price level in some sense. Secondly, in the APT, factor sensitivities of asset returns are assumed to be exogenous. More natural modelling would require these parameters to be endogenously determined reflecting some fundamental properties of the assets. Thirdly, with some minor exceptions[2] the discussion is limited within a two-period setting, so that they implicitly assumes away any possibilities of intertemporal changes in investment opportunities. Of course, these problems are interrelated. Given income gains from an asset, the factor sensitivities of return are essentially the same as those of the asset price. A multiperiod setting would be necessary to investigate valuation of the longer-term assets for which return-distributions change over time.

This paper develops a continuous-time model of asset pricing with k factors and rational expectations. The model is used for the purpose of recasting the APT as a general theory of the arbitrage asset valuation.[3] The key assumption is that underlying factors obey a joint diffusion process. This specification will enable us to apply the diffusion approach to arbitrage asset pricing. Here the diffusion approach is one which was originated by Black and Scholes (1973) for option pricing, applied to evaluating other derivative assets (*e.g.*, Vasicek (1977), Richard (1978), Brennan and Schwartz (1979), and Cox, Ingersoll and Ross (1981)), and made rigorous by Harrison and Kreps (1979) and Harrison and Pliska (1981). In pursuing the aim, we generalize the approach to evaluate assets with the function forms of their income streams and boundary values unspecified explicitly.

Our main messages are as follows. First of all, our dynamic asset pricing model of the Black and Scholes type is consistent with the factor return-generating model assumed in the APT. Given the result, the APT-type no-arbitrage conditions are recast in the dynamic setting with sensitivity coefficients defined by the partial elasticities of asset prices with respect to factors. The conditions assert the existence of a set of k predictable price processes which evaluate factor risk over time. Secondly, these conditions reduce to partial differential equations with respect to asset valuation functions. Thirdly, the price functions, as the solutions of these equations, can take two alternative forms depending on whether risk adjustment is made on: (a) a discount rate, or (b) expected factor dynamics. Finally, the resulting asset pricing formulae consistently demonstrate the various ideas of asset evaluation in the existing literature such as Modigliani and Miller (1959), Cox and Ross (1976), Rubinstein (1976), Harrison and Kreps (1979), and Harrison and Pliska (1981). In sum, this paper accomplishes synthesis in two different senses: it synthesizes the APT and the diffusion approach of the Black and Scholes type; and it integrate the existing ideas of asset evaluation from the viewpoint of arbitrage.

The remainder of this paper proceeds as follows. The model is presented

[2] Roll and Ross (1980) developed a simple intertemporal model of the APT where determinant factors follow Wiener processes with exogenous factor loadings. Solnik (1983) and Ikeda (1991) implicitly considered the same situation as in Roll and Ross (1980). However, neither discussed asset evaluation problems.

[3] A similar purpose has been pursued by Chang and Shanker (1987). They have applied the APT to option pricing, but the analysis there is limited within a static uni-factor framework and his interest is only in options markets.

in Section 6.2. In Section 6.3, we derive the APT-type no-arbitrage conditions on risk premia. Based on these conditions, Section 6.4 investigates arbitrage valuation of risky assets to obtain two alternative formulae for asset pricing in Subsections 6.4.1 and 6.4.2. Finally, Section 6.5 summarizes our analyses.

6.2 A Multi-Factor Asset Pricing Model

Ross (1976, 1977) based his discussion on a multi-factor *return-generating* process. In that setting, however, it would be difficult to obtain any explicit expression for asset valuation rules. Instead, we set up a multi-factor *pricing* system. The model is composed of the following set of assumptions, A1 through A6.

A1: *There are n kinds of risky assets which are infinitely divisible. They are indexed by $i = 1, 2, \ldots, n$. The set of these index numbers is denoted by A, i.e., $A = \{1, \ldots, n\}$. The markets are assumed to be: (i) perfect, with all investors taking asset prices as given; (ii) frictionless in that trading assets takes neither tax nor any other transaction costs; and (iii) open continuously in time.*

A2: *States of all the markets are described by k stochastic common factors. We denote the factors as $\phi_1(t), \ldots, \phi_k(t)$ where t represents time.*

Assumption A2 implicitly neglects the existence of asset-specific of idiosyncratic factors which the APT takes into account. However, this simplification will make little difference between the APT and our analysis since the APT assumes idiosyncratic factor risk to be unsystematic. The case with asset idiosyncrasy is briefly treated in Appendix A.2.

Contents of the common factors will depend on the set of assets under consideration. In the case of stocks, the factors may include some indices affecting real activities of firms. For derivative assets, factors may be certain variables conditioning underlying contracts. Anyway, their contents are left unspecified in this paper. What we need for our ends is specification of stochastic processes for factors.

A3: *The $(k \times 1)$ vector of factors, $\phi(t) = [\phi_j(t)]$, follows a diffusion process of the type:*[4]

$$d\phi(t) = I_\phi(t)\, \mu(\phi(t), t)\, dt + I_\phi(t) \sum (\phi(t), t)\, dz(t), \qquad (6.1)$$

where $I_\phi(t)$ denotes a $(k \times k)$ diagonal matrix whose $j-th$ component is $\phi_j(t)$. $\mu(\phi(t), t) = [\mu_j(\phi(t), t)]$ is a bounded $(k \times 1)$ vector of the expected rates of change, $\sum(\phi(t), t) = [\sigma_{jl}(\phi(t), t)]$, a bounded $(k \times k)$ matrix of diffusion coefficients of the unexpected rates of change, and $z(t) = [z_l(t)]$, a k-dimensional standard Wiener process.[5] *The covariance matrix of factors, which can be expressed as $I_\phi \sum \sum' I_\phi$ with the prime denoting transpose, is assumed to have*

[4] For detailed information on diffusion processes and stochastic differential equations, see, for example, Arnold (1973) and Friedman (1975). Applications of stochastic calculus to finance and economics are surveyed in Malliaris and Brock (1982) and Ingersoll (1987).

[5] By the definition of multi-dimensional Wiener processes, the components of z are independent of each other.

full rank.

Given assumptions A2 and A3, we specify information available to investors.

A4: *All investors are aware of the fact that asset markets are described as in assumptions A2 and A3, and have exact knowledge of the vector valued functions,* $\mu(\phi(t), t)$ *and* $\sum(\phi(t), t)$ *,as well as of the current values of factors and asset prices.*

Next, rationality of investors is assumed.

A5: *Investors behave rationally in the following two senses: (i)They prefer more wealth to less; and (ii) given the information set assumed in A4, expectations are formed consistent with the model.*

In A5, the former rationality, (i), excludes any profitable opportunities of riskless arbitrage. The latter assumption, (ii), is the rational-expectations hypothesis.

Assumptions A2 through A5 amount to the following. From assumption A3, the factors are jointly Markov, so that the probability distributions of their future values are independent of their past time-paths. Hence, by assumptions A2, A4 and A5 (ii), we can naturally set the time t price of the i-th asset, $P^i(t)\,(i \in A)$, equal to a function of the time t values of factors and the time, $P^i(\phi(t), t)$. Here, function $P^i(\phi(t), t)$ is not exogenously but endogenously determined. The determination mechanism is our main interest in Section 6.4.

Finally, instantaneous riskless bonds are assumed to exist.

A6: *Riskless instantaneous borrowing and lending are available at the interest rate,* r*. The rate is assumed to be an exogenous function of state variables,* $r(t) = r(\phi(t), t)$.

This is set only for the sake of simplicity. If risk-free bonds did not naturally exist, we could interpret $r(t)$ as the rate of return on the zero-beta portfolio without any change in our discussion below.

6.3 Arbitrage Determination of Risk Premia

From the model presented above, let us now derive the familiar APT-type linear equations for risk premia. We first show the consistency between our multi-factor asset-pricing model and the multi-factor return-generating processes assumed by the APT.

As pointed out previously, asset price $P^i(t)$ is represented as a function $P^i(\phi(t), t)$. This implies that the price follows a diffusion process since the factors are diffusions. Therefore, letting $\gamma_i(t) = \gamma_i(\phi(t), t)$ be some exogenous payout flows or income gains from the i-th asset, and $\alpha_i(t) = \alpha_i(\phi(t), t)$ be the total expected rate of return (the expected rate of price changes plus payout flows) on the asset, by Ito's Lemma, we can describe the dynamics of (ex-dividend) price $P^i(t)$ as

$$dP^i(t) = \left\{\alpha_i(t) P^i(t) - \gamma_i(t)\right\} dt + P^i_\phi(t) I_\phi(t) \sum(t)\, dz(t), \qquad (6.2)$$

where the expected instantaneous change of the price, $\alpha_i(t) P^i(t) - \gamma_i(t)$, is given by

$$\begin{aligned} \alpha_i(t) P^i(t) - \gamma_i(t) = P^i_\phi(t) I_\phi(t) \mu(t) + P^i_t(t) \\ + (1/2)\,\mathrm{trace}\left\{ P^i_{\phi\phi}(t) I_\phi(t) \Omega(t) I_\phi(t) \right\}. \end{aligned} \tag{6.3}$$

Here $\Omega(t) = \Omega(\phi(t), t)$ represents the covariance matrix of the unexpected rates of changes in factors, *i.e.*, $\Omega \equiv \sum \sum'$, and the subscripts on $P^i(t)$ denote partial derivatives. In particular, $P^i_\phi(t)$ stands for a $(1 \times k)$ vector whose j-th element is $\partial P^i(\phi(t), t) / \partial \phi_j(t)$, and $P^i_{\phi\phi}(t)$ is the Hessian of $P^i(\phi(t), t)$ without columns and rows of partial derivatives with respect to t.

The consistency between our setting and the multi-factor return-generating processes assumed by the APT is ascertained if we recall that term $\sum(t)\, dz(t)$ represents the unexpected rates of changes in factors, $I_\phi^{-1} d\phi(t) - \mu(t)\, dt$, and rewrite equation (6.2) as

$$\begin{aligned} \left\{ dP^i(t) + \gamma_i(t)\, dt \right\} / P^i(t) - \alpha_i(t)\, dt = \hat{P}^i_{\phi_1}(t) \left\{ d\phi_1(t) / \phi_1(t) - \mu_1(t)\, dt \right\} + \\ \cdots + \hat{P}^i_{\phi_k}(t) \left\{ d\phi_k(t) / \phi_k(t) - \mu_k(t)\, dt \right\}, \end{aligned} \tag{6.4}$$

where the hat (^) denotes the partial elasticity of an asset price with respect to each factor, *i.e.*,$\hat{P}^i_{\phi_j}(t) = \left(\phi_j(t) / P^i(t) \right) \left\{ \partial P^i(\phi(t), t) / \partial \phi_j(t) \right\}$. The left hand side of (6.4) stands for the unexpected rate of return on the i-th asset. Thus, this equation states that an innovation in the rate of return is generated as a linear combination of the unexpected changes of factors. This is the same situation as Ross assumed except for two points. First, in process (6.4) the sensitivity coefficients to factors' innovations are defined by the endogenous partial elasticities, $\hat{P}^i_{\phi_j}$, instead of exogenous betas. Secondly, no idiosyncratic factor appears in this equation.

Now that multi-factor return-generating processes are derived in equation (6.4), we discuss arbitrage determination of risk premia by extending the APT argument straightforwardly to the dynamic setting. Note that if one considers a self-financing sequence of portfolios, or a self-financing trading-strategy, $\{\theta\}_t = \left\{ (\theta_1, \ldots, \theta_n)' \right\}_t$, its local rate of return is given by,[6]

$$\theta(t)' \alpha(t)\, dt + \theta(t)' \hat{P}_\phi(t) \sum(t)\, dz(t), \tag{6.5}$$

where $\alpha(t)$ is the $(n \times 1)$ vector of the expected rates of return, and $\hat{P}_\phi(t)$ represents the $(n \times k)$ elasticity matrix whose (i, j)-element is $\hat{P}^i_{\phi_j}(t)$. Thus $\theta' \hat{P}_\phi$ determines the risk for the trading strategy. The absence of arbitrage requires: if $\theta' \hat{P}_\phi = 0$, then $\theta'(\alpha - r1_n) = 0$, where 1_n denotes the $(n \times 1)$ vector of ones. The algebraic consequence of this is that $(\alpha - r1_n)$ is linear in the k column vectors of $\hat{P}_\phi$, $\hat{P}_{\phi_1}, \ldots, \hat{P}_{\phi_k}$.

Lemma: *In the absence of arbitrage, the expected excess rates of return (over the risk-free interest rate) on risky assets are determined as linear in factor elasticities of their prices with weights which are the same across assets. Formally,*

[6] By the definition of self-financing strategies, expression (6.5) does not include any term of the continuous change in portfolio weighting. For this, see Ingersoll (1987, p.363) and Duffie (1988, p.157).

there exists a set of predictable processes, $\{\lambda\}_t = \{(\lambda_1, \ldots, \lambda_k)'\}_t$, *such that*

$$\alpha_i(t) - r(t) = \lambda_1(t)\hat{P}^i_{\phi_1}(t) + \cdots + \lambda_k(t)\hat{P}^i_{\phi_k}(t) \qquad \forall i \in A, \tag{6.6}$$

or, more integratingly in vector notation,

$$\alpha(t) - r(t)1_n = \hat{P}_\phi(t)\lambda(t). \tag{6.7}$$

The lemma recasts the original APT in the intertemporal setting in terms of endogenous partial elasticities of asset prices. Intuitively, the elasticities, $\hat{P}^i_{\phi_j}(i \in A, j = 1, \ldots, k)$, can be regarded as measures of local risk caused by factors, and predictable processes λ_j as risk prices. The statement then illustrates that the risk premium of any risky asset is determined as an inner product of the risk and risk-price vectors.

Coefficient λ_j is interpreted more clearly if we assume the existence of a sufficient number of non-redundant assets to introduce a *factor trading strategy* as Ross considered a factor portfolio in his static model.[7] Here the j-th factor trading strategy is defined as self-financing one whose instantaneous local rate of return always responds to the unexpected rate of change in the j-th factor with an elasticity of one and to the other factors with zero sensitivity. Explicitly, it is a trading strategy $\{\varepsilon^j\}_t = \left\{\left(\varepsilon^j_1, \ldots, \varepsilon^j_n\right)'\right\}_t$ satisfying

$$\hat{P}_\phi(t)'\varepsilon^j(t) = e^j, \tag{6.8}$$

where $\varepsilon^j(t)'1_n = 1$ and e^j stands for a $(k \times 1)$ vector with all zeros except for a unity in the j-th spot. Denoting the local expected rate of return on the j-th factor trading strategy by $\mu^j(t)$, its expected excess rate of return over the risk-free rate, or the j-th factor risk premium reduces to

$$\begin{aligned}
\mu^j(t) - r(t) &= \varepsilon^j(t)'(\alpha(t) - r(t)1_n) \\
&= \varepsilon^j(t)'\hat{P}_\phi(t)\lambda(t) && \text{(by (6.7))} \\
&= e^{j\prime}\lambda(t) && \text{(by (6.8))} \\
&= \lambda_j(t).
\end{aligned}$$

Therefore, λ_j is exactly equal to the j-th factor premium.

Remark: *As can easily be seen from the last discussion, the predictable vector process, λ, uniquely exists in complete markets where all of the k factor trading-strategies are available or marketed. When the markets are incomplete, on the other hand, the vector process which supports the arbitrage-free condition (6.6) is not unique. See Ingersoll (1987), pp.168-170.*

6.4 Arbitrage Asset Valuation

Let us next direct our attention to asset valuation functions. In general, conditions on risk premia of risky assets will, implicitly or explicitly, bind their prices. In fact, the APT-type equation (6.6) can be rewritten as a partial differential equation for the asset price function.

[7] See Ikeda (1988), pp.6-7, for the condition for the asset non-redundancy.

Corollary 1: *In the absence of arbitrage, the price function of any risky asset must satisfy the partial differential equation,*

$$(1/2)\,\mathrm{trace}\left\{P^i_{\phi\phi}(t)\, I_\phi(t)\, \Omega(t)\, I_\phi(t)\right\} + P^i_\phi(t)\, I_\phi(t)\,(\mu(t) - \lambda(t)) + P^i_t(t) \\ +\gamma_i(t) = r(t)\, P^i(t)\,,\ \forall i \in A. \tag{6.9}$$

Proof: Solve equation (6.3) for $\alpha_i(t)$ and substitute the result into no-arbitrage condition (6.6) to obtain equation (6.9). □

Differential equation (6.9) holds for any risky asset.[8] The forms of income gain functions and some boundary conditions will specify the characteristics of assets. Here we impose boundary conditions on asset prices at time $T\,(> t)$ in a general form:

$$P^i(\phi(T), T) = \Phi^i(\phi(T)) \qquad \forall i \in A, \tag{6.10}$$

where $\Phi^i(\phi(T))$ is an exogenous function evaluating the time-T value of the i-th asset.[9]

If asset i is a contingent claim with some maturity, the function is given by the maturity payoff. On the other hand, in the case of assets with eternal lives like stocks, it seems to be arbitrary to put boundary conditions like (6.10) on the prices. This, however, turns out not to be more restrictive than it might first appear if one notices the following two points. The first is that regarding $\Phi^i(\phi(T))$ as some investors' belief concerning the time T-value of asset i, the solution of (6.9) for $P^i(t)$, which will be expressed in terms of $\Phi^i(\phi(T))$, can be interpreted as a relative relationship between the present and expected future prices. Secondly, as will be shown later, investors discount the future value $\Phi^i(\phi(T))$ at some rate to evaluate the present value of asset i. Thus, we could assume away any effect of $\Phi^i(\phi(T))$ on $P^i(t)$ by taking a limit $T \to +\infty$ and neglecting any possibilities of price bubbles.

Asset valuation functions, if they exist, can be obtained by solving differential equations (6.9) subject to boundary conditions (6.10). In general, the risk-free asset would be priced as its expected present value discounted at the riskless interest rate. Also in the risk-neutral world, any assets would be evaluated in the same fashion. In our general setting, however, we must make some risk-adjustment on risk-neutral asset valuation. Assuming the existence and uniqueness of a positive solution of (6.9) and (6.10),[10] the following subsections show that these equations can be solved as two alternative pricing formulae depending on how to design risk-adjustment. The first is a formula with a discount rate reflecting factor risk. The second method for asset pricing adjusts expectations concerning state dynamics.

[8] Many articles which use the diffusion approach have presented the similar equations, but in some specific forms. See, for example, Black and Scholes (1973), Vasicek (1977) and Richard (1978). We can reduce equation (6.9) to theirs by specifying factors and the functional forms of μ, $\sum$, and γ_i arbitrarily.

[9] Conditions (6.10) implicitly assume that boundary dates are the same across all assets as in Cox, Ingersoll, and Ross (1985). This simplifies our analysis because if the dates differed from each other, hedging opportunities would change over time.

[10] A stochastic differential equation requires dome technical conditions to have a pathwise unique solution. The details are given in Arnold (1973, Chap. 6) and Friedman (1975, Chap. 5).

6.4.1 Solution (a)

The first approach defines some risk-adjusted discount rate in computing the expected present values of risky assets. Although the idea has so far been proposed by many studies as seen in some familiar textbooks such as Elton and Gruber (1987),[11] they have defined a discount rate as exogenously given. The problem here is how to find a discount rate which is consistent with the no-arbitrage condition, (6.6) or (6.9). To answer this, let us first consider a simple case where risky assets earn no income gain, *i.e.*, $\gamma_i(s) \equiv 0 \forall i \in A$ and $s \in [t, T]$. We denote asset prices in this case by small letters like $p^i(s) = p^i(\phi(s), s)$. Note that we are now concerned with the homogeneous equation of (6.9).

Given the setting, we find a self-financing trading-strategy whose relative value to each asset is a martingale in the absence of arbitrage. Formally, it is a self-financing strategy whose value process, say $D(s) = D(\phi(s), s)$, satisfies the following relationship:

$$p^i(t)/D(t) = E_t\left\{p^i(s)/D(s)\right\} \quad \forall i \in A, \quad s \in [t, T], \tag{6.11}$$

or equivalently,

$$E_t\left[d\left\{p^i(s)/D(s)\right\}\right] = 0 \quad \forall i \in A, \quad s \in [t, T], \tag{6.12}$$

where E_t denotes the mathematical expectation conditional on information at the current time, t. If one obtains this trading strategy, by definition (6.11), the price of asset i is expressed as an expected present value of its boundary payoff:

$$p^i(t) = E_t\left[p^i(T)\left\{D(t)/D(T)\right\}\right].$$

This exhibits that the appropriate discount factor in the case of no income-gain is given by the intertemporal relative price of the strategy, $D(t)/D(T)$. It is straightforward to prove that exactly the same discount factor is also applicable to the general case with income gains. This is the reason for our interest in the strategy defined by (6.11). One might be able to call it the *discounting trading strategy.*

The discounting trading strategy can be found in the following manner. Since its price conforms to a diffusion process by assumption, by Ito's Lemma, we can compute the expected local rate of change in the relative price, $p^i(s)/D(s)$, as follows:

$$\begin{aligned} E_s\left[\frac{d\left\{p^i(s)/D(s)\right\}}{\left\{p^i(s)/D(s)\right\}}\right] &= E_s[\left\{dp^i(s)/p^i(s)\right\} - \left\{dD(s)/D(s)\right\} + \left\{dD(s)/D(s)\right\}^2 \\ &\quad - \left\{dp^i(s)/p^i(s)\right\}\left\{dD(s)/D(s)\right\}]. \end{aligned} \tag{6.13}$$

Here the expected rates of return, $E_s\left\{dp^i(s)/p^i(s)\right\}$ and $E_s\left\{dD(s)/D(s)\right\}$, must satisfy the arbitrage-free relationship like (6.6), so that they aregiven as

$$E_s\left\{dp^i(s)/p^i(s)\right\} = \left\{r(s) + \hat{p}^i_\phi(s)\lambda(s)\right\}ds, \tag{6.14}$$

$$E_s\left\{dD(s)/D(s)\right\} = \left\{r(s) + \hat{D}_\phi(s)\lambda(s)\right\}ds, \tag{6.15}$$

[11] The idea was originally proposed by Modigliani and Miller (1959).

respectively, where $\hat{p}^i_\phi(s)$ and $\hat{D}_\phi(s)$ represent the $(1 \times k)$ vectors of factor elasticities. On the other hand, the realized rates of return, $dp^i(s)/p^i(s)$ and $dD(s)/D(s)$, can be calculated by application of Ito's Lemma as

$$dp^i(s)/p^i(s) = E_s\left\{dp^i(s)/p^i(s)\right\} + \hat{p}^i_\phi(s)\sum(s)\,dz(s)\,, \tag{6.16}$$

$$dD(s)/D(s) = E_s\{dD(s)/D(s)\} + \hat{D}_\phi(s)\sum(s)\,dz(s)\,. \tag{6.17}$$

These two equations imply that the second order moments in (6.13) are given by

$$E_s\{dD(s)/D(s)\}^2 = \hat{D}_\phi(s)\,\Omega(s)\,\hat{D}_\phi(s)'\,ds, \tag{6.18}$$

$$E_s\left\{dp^i(s)/p^i(s)\right\}\{dD(s)/D(s)\} = \hat{p}^i_\phi(s)\,\Omega(s)\,\hat{D}_\phi(s)'\,ds. \tag{6.19}$$

Finally, by substituting equations (6.14), (6.15), (6.18) and (6.19) into (6.13) and rearranging the result, the expression of the expected rate of change in $p^i(s)/D(s)$ reduces to

$$E_s\left[\frac{d\left\{p^i(s)/D(s)\right\}}{\left\{p^i(s)/D(s)\right\}}\right] = \left\{\hat{p}^i_\phi(s) - \hat{D}_\phi(s)\right\}\left\{\lambda(s) - \Omega(s)\,\hat{D}_\phi(s)'\right\}ds. \tag{6.20}$$

Thus, if we choose and reshuffle the portfolio continuously such that $\forall s \in [t, T]$,

$$\hat{D}_\phi(s) = \lambda(s)'\,\Omega(s)^{-1}\,, \tag{6.21}$$

then the resulting sequence is the discounting trading strategy because portfolio-formation rule (6.21) with expression (6.20) implies equation (6.12).

An explicit price dynamics of the discounting trading strategy can now be obtained by successive substitution of (6.15) and (6.21) into (6.17) as:

$$dD(s)/D(s) = \left[r(s) + \lambda(s)'\,\Omega(s)^{-1}\lambda(s)\right]ds + \lambda(s)'\,\Omega(s)^{-1}\sum(s)\,dz(s)\,. \tag{6.22}$$

After logarithmic transformation of $D(s)$,[12] we integrate this differential equation by Ito's rule to get an expression of the intertemporal relative price:

$$D(t)/D(T) = \exp\{-\chi(T)\}\,, \tag{6.23}$$

where $\chi(\tau)$ $(\tau \in [t, T])$ is defined as

$$\begin{aligned}\chi(\tau) \;&=\; \textstyle\int_t^\tau \left[r(\phi(s), s) + (1/2)\,\lambda(\phi(s), s)'\,\Omega(\phi(s), s)^{-1}\,\lambda(\phi(s), s)\right]ds \\ &\quad + \textstyle\int_t^\tau \lambda(\phi(s), s)'\,\Omega(\phi(s), s)^{-1}\sum(\phi(s), s)\,dz(s)\,.\end{aligned} \tag{6.24}$$

[12] Through Ito's Lemma we have:

$$d\ln D = \frac{\partial \ln D}{\partial D}dD + \frac{1}{2}\frac{\partial^2 \ln D}{\partial D^2}(dD)^2\,.$$

Using equation (6.22), one can eliminate dD and $(dD)^2$ from the above expression to obtain

$$d\ln D = \left(r + (1/2)\,\lambda'\Omega^{-1}\lambda\right)ds + \lambda'\Omega^{-1}\sum dz.$$

In the case of no income-gain, as stated earlier, the prices of assets are determined as their expected present values and the arbitrage-free discounting factor is now given by equations (6.23) and (6.24).

Not surprisingly, this discussion can be extended straightforwardly to the general case with income gains, *i.e.*,$\gamma_i(s) \neq 0$. We obtain the first formula for asset pricing.

Proposition 1: *In the absence of arbitrage, the price of any risky asset is determined as an expected present value of its future income stream and boundary price. Here the discounting factor is given by the intertemporal relative price of the discounting trading strategy. That is,*$\forall i \in A$,

$$P^i(\phi(t),t) = E_t\left[\Phi^i(\phi(T))\exp\{-\chi(T)\} + \int_t^T \gamma_i(\phi(\tau),\tau)\exp\{-\chi(\tau)\}\,d\tau\right], \tag{6.25}$$

where $\chi(\tau)$ *is defined by* (6.24).

Proof: See Appendix A.1.□

The above formula accomplishes an *exponential* risk-adjustment in that the discount rate is adjusted by risk. However, the same equation can be rearranged as an *additive* risk-adjustment formula provided by Rubinstein (1976). He considered securities within a discrete-time framework to present a pricing formula with risk-adjustment done in an additive manner.

To show the equivalence, let us introduce a random variable $Y(\tau)$ ($\tau \in [t,T]$) as

$$\begin{aligned} Y(\tau) \;=\;& \exp\{-\textstyle\int_t^\tau (1/2)\,\lambda(\phi(s),s)'\,\Omega(\phi(s),s)^{-1}\,\lambda(\phi(s),s)\,ds \\ & -\textstyle\int_t^\tau \lambda(\phi(s),s)'\,\Omega(\phi(s),s)^{-1}\sum(\phi(s),s)\,dz(s)\}, \end{aligned}$$

and suppose that the expectation operator and the integral with respect to time are interchangeable in order. Then, the following corollary results from the definition of covariance:

Corollary 2: *Formula (6.25) can be rewritten as,* $\forall i \in A$,

$$\begin{aligned} P^i(\phi(t),t) \;=\;& E_t\Big[\Phi^i(\phi(T))\exp\Big\{-\textstyle\int_t^T r(\phi(\tau),\tau)\,d\tau\Big\} \\ & +\textstyle\int_t^T \gamma_i(\phi(\tau),\tau)\exp\Big\{-\int_t^\tau r(\phi(s),s)\,ds\Big\}\,d\tau\Big] + \Psi_i(t), \end{aligned} \tag{6.26}$$

where

$$\begin{aligned} \Psi_i(t) \;=\;& \mathrm{Cov}_t\Big[\Phi^i(\phi(T))\exp\Big(-\textstyle\int_t^T r(\tau)\,d\tau\Big), Y(T)\Big] \\ & +\textstyle\int_t^T \mathrm{Cov}_t\left\{\gamma_i(\tau)\exp\left(-\int_t^\tau r(s)\,ds\right), Y(\tau)\right\}d\tau. \end{aligned} \tag{6.27}$$

Equation (6.26) states that the price of any risky asset is given as its expected present value discounted at the risk-free rate *plus* some risk-adjusting term,Ψ_i. The latter, by equation (6.27), is determined by normalized covariance of a mixed factor risk Y and the discounted income stream on the asset. This scenario is fairly consistent with that of Rubinstein (1976). In effect, equations (6.26) and (6.27) reduce to the same formula as Rubinstein's Theorem 1 if one makes another assumption that the interest rate is deterministic and takes a limit $T \to +\infty$ while neglecting any possibilities of price bubbles.

Furthermore, formula (6.25) can be rewritten in the form of martingale measure representation proposed by Harrison and Kreps (1979), Harrison and Pliska (1981), Huang (1987), and Chamberlin (1988). Regarding $Y(\tau)$ as a Radon-Nikodym derivative,[13] let us define a new equivalent probability measure $\bar{M}(\tau)$ for $\phi(\tau)$ from the original one, $M(\tau)$, as

$$\bar{M}(\tau) = \int Y(\tau)\, dM(\tau).$$

Then, the immediate result is the following.

Corollary 3: *Formula (6.25) can be rewritten as, $\forall i \in A$,*

$$\begin{aligned} P^i(\phi(t), t) &= \bar{E}_t[\Phi^i(\phi(T)) \exp\left\{-\int_t^T r(\phi(\tau), \tau)\, d\tau\right\} \\ &\quad + \int_t^T \gamma_i(\phi(\tau), \tau) \exp\left\{-\int_t^\tau r(\phi(s), s)\, ds\right\} d\tau], \end{aligned}$$

where $\bar{E}$ denotes the expectation with respect to the equivalent martingale measure $\bar{M}$.

That is, an asset price is determined such that the price process discounted by the riskless rate becomes a martingale with respect to some artificial probability measure.

6.4.2 Solution (b)

The second approach considers an alternative economy which supports the same relationship as (6.9). Let us note that condition (6.9) can be also obtained in a hypothetical economy defined by the following two assumptions: (i) the expected rates of return on any risky assets are determined as the risk-free rate, *i.e.,*

$$\alpha_i(t) = r(t) \quad i \in A, \tag{6.28}$$

and (ii) capital markets are completely described by new state variables ϕ^*, rather than ϕ, which jointly follow a risk-adjusted process,

$$d\phi^*(t) = I_{\phi^*}(t)\{\mu(\phi^*(t), t) - \lambda(t)\}\, dt + I_{\phi^*}(t) \sum (\phi^*(t), t)\, dz(t), \tag{6.29}$$

[13] By Radon-Nikodym's theorem, $E_t[Y(\tau)] = 1$ must be valid for $Y(\tau)$ to be a Radon-Nikodym derivative. The equality can be proven as follows. First, define $q^i(\tau)$ as the i-th asset's price discounted with r, i.e.,

$$q^i(\tau) = P^i(\tau) \exp\left(-\int_t^\tau r(s)\, ds\right).$$

Then, we can verify in the same way as in the text that:

$$q^i(t) = E_t\left[q^i(\tau) Y(\tau)\right] \qquad \forall i \in A.$$

Next, think of asset i as an asset or a self-financing trading-strategy which yields $\exp\left(\int_t^\tau r(s)\, ds\right)$ at time τ ($\tau \in [t, T]$). the price of the asset must satisfy $q^i(t) = q^i(\tau) = 1$. Substituting this into the above expression, we obtain $E_t[Y(\tau)] = 1$ as desired.

where the parametric vector $\lambda(t)$ is the same as given by (6.7) in the *original* economy.[14] This equivalence implies that, given all other things equal, asset prices in the hypothetical economy are equal to ones in the original setting. On the other hand, one can easily guess that the former prices are determined as expected present values of the assets with discounting done at the risk-free rate. As a consequence, we get the second formula for asset pricing in our original economy:

Proposition 2: *The arbitrage-free price of any risky asset is also expressed as its expected present value with discounting done at the risk-free rate if expectations are formed along the risk-adjusted factor dynamics. Formally,* $\forall i \in A$,

$$\begin{aligned} P^i(t) \;=\; & E_t^*\left[\Phi^i(\phi^*(T))\exp\left\{-\int_t^T r(\phi^*(\tau),\tau)\,d\tau\right\}\right. \\ & \left.+\int_t^T \gamma_i(\phi^*(\tau),\tau)\exp\left\{-\int_t^\tau r(\phi^*(s),s)\,ds\right\}d\tau\right], \end{aligned} \tag{6.30}$$

where E_t^* *denotes the mathematical expectation along dynamics (6.29).*

Proof: Applying Theorem 5.3 of Friedman (1975) to equations (6.9) and (6.10) results in solution (6.30). □

The above solution technique is similar to one presented by Cox and Ross (1976) for option pricing. They found an option price function of the underlying stock price by considering an expectational economy with risk-neutral investors. At first sight, however, the two seem to be a little different from each other because risk-neutrality implies that factor risk premia, $\lambda_i(t)$, must be equal to zero in the absence of arbitrage. The difference comes from the fact that factors are not specified as market prices here. If they were, the expected local rates of return on factor trading strategies must be equal to those on factors themselves, *i.e.*, $\mu^i(t) = \mu_i(t)$. As a result, the risk-adjusted expected rates of return on factors would be reduced as

$$\mu^i(t) - \lambda^i(t) = \mu^i(t) - \left\{\mu^i(t) - r(t)\right\} = r(t).$$

In this case, the interest-rate-determined return hypothesis (6.28) *cum* risk-adjusted factor dynamics (6.29) is equivalent to the risk-neutrality assumption. Thus, the approach by Cox and Ross can be regarded as a special case of ours.

Remark: *Our analysis has been limited to the case of no idiosyncratic factor. But Appendix A.2 shows that all the above discussions can be extended straightforwardly to the general case with idiosyncratic factors if, as in Ross (1976, 1977), the number of assets, n, is assumed to be large enough for investors to diversify away the idiosyncratic risk. In particular, it proves that the same risk pricing equation as (6.6) and asset pricing formulae as (6.25) and (6.30) must (approximately) hold in that setting. The only difference is that in that case the i-th idiosyncratic factor enters into functions* α_i, γ_i, *and* Φ^i. *Needless to say, this does not mean that idiosyncratic factors have no influence on asset pricing. The exact implication is that the idiosyncratic risk does not exert any effect on the price levels. The expected trend of the idiosyncratic factor dynamics does affect the asset price through changing the expected income stream and boundary payoff.*

[14] It can easily be verified by using Ito's Lemma that the left hand side of (6.9) with ϕ replaced by ϕ^* is equal to the expected return generated by hypothetical dynamics (6.29).

6.5 Conclusion

In this paper we have developed an arbitrage asset pricing model with diffusion factors and rational expectations. Using the diffusion approach of the Black and Scholes (1973) type in a general way, the APT of Ross (1976, 1977) is extended to a general theory of arbitrage asset valuation.

Our main discussion is summarized as follows. First of all, defining sensitivity coefficients by partial elasticities of asset prices with respect to factors, one can reformulate the APT-type conditions in terms of asset price functions. The conditions assert that there must be a set of predictable factor price processes in the absence of free lunches.

Secondly, these conditions reduce to partial differential equations with respect to asset price functions.

Finally, asset price functions, which are solutions of the differential equations, take two alternative forms depending on how risk-adjustment is designed. The first method of risk-adjustment is to define a discount rate as reflecting factors' risk. We have shown that, in the absence of arbitrage, the discount rate is given by the intertemporal relative price of a trading strategy whose relative price to each asset is a martingale, and that the resulting asset valuation formula consistently demonstrates the additive risk-adjusting proposed by Rubinstein (1976) and the martingale measure representation given by Harrison and Kreps (1979) and others. Even with discounting done at the risk-free interest rate, risk-adjusting can also be carried out by taking expectations along some risk-adjusted factor dynamics. This is the second method of risk-adjusting. It has been pointed out that the risk-neutrality approach by Cox and Ross (1976) is a special case of this method, and that if we additionally suppose that factors are market prices, then two approaches are equivalent.

The resulting formulae of asset pricing have general applicability since the function forms of factor dynamics, payout flows, and terminal payoffs have been left unspecified in the model.

6.6 Appendix for Chapter 6

6.6.1 Appendix A.1: Proof of Proposition 1

Here we prove that price function (6.25) is the solution of differential equation (6.9) (given the assumption that the solution uniquely exists). Let us define a random variable $x(\tau), \tau \in [t, T]$, as

$$x(\tau) = P^i(\tau) \exp\{-\chi(\tau)\} + \int_t^\tau \gamma_i(s) \exp\{-\chi(s)\} ds,$$

where $\chi(s)$ is given by (6.24). Then, if it is shown that under condition (6.9) $x(\tau)$ is a martingale, and so $x(t) = E_t\{x(T)\}$, it completes the proof because $x(t) = P^i(t)$ and $E_t\{x(T)\}$ is equal to the right hand side of (6.25) by construction. In the following we shall show $E_t\{dx(\tau)\} = 0$.

By Ito's Lemma, one obtains:

$$
\begin{aligned}
dx(\tau) &= \exp(-\chi)\, dP^i - P^i \exp(-\chi)\, d\chi + \gamma_i \exp(-\chi)\, d\tau \\
&\quad + (1/2) P^i \exp(-\chi)\, (d\chi)^2 - \exp(-\chi)\, d\chi dP^i \\
&= \exp(-\chi) \left\{ dP^i - P^i d\chi + \gamma_i d\tau + (1/2)\, P^i (d\chi)^2 - d\chi dP^i \right\}.
\end{aligned} \tag{6.31}
$$

Here $d\chi$ can be computed from definition (6.24) as

$$
d\chi(\tau) = \left(r + (1/2)\, \lambda' \Omega^{-1} \lambda\right) d\tau + \lambda' \Omega^{-1} \sum dz. \tag{6.32}
$$

Combining expressions (6.2), (6.3) and (6.32) yields

$$
\begin{aligned}
(d\chi)^2 &= \lambda' \Omega^{-1} \lambda d\tau, \\
d\chi dP^i &= P^i_\phi I_\phi \lambda d\tau.
\end{aligned} \tag{6.33}
$$

Substituting equations (6.2), (6.3), (6.32) and (6.33) into (6.31), and taking expectations at the initial time, we obtain, $\forall \tau \in [t, T]$,

$$
\begin{aligned}
E_t \{dx(\tau)\} &= E_t \Big[\exp(-\chi) \Big[(1/2)\, \mathrm{trace} \left(P^i_{\phi\phi} I_\phi \Omega I_\phi \right) \\
&\quad + P^i_\phi I_\phi (\mu - \lambda) + P^i_\tau - r P^i + \gamma_i \Big] d\tau \Big].
\end{aligned}
$$

It follows that if equilibrium condition (6.9) is satisfied, $E_t \{dx(\tau)\} = 0$ holds. □

6.6.2 Appendix A.2: The Case with Idiosyncratic Factors

This appendix briefly treats the case with asset idiosyncrasy. We first introduce asset idiosyncrasy by replacing A2 in the text with the following two assumptions.

A7: *Each asset market is affected by some idiosyncratic factor as well as the common ones. We denote the factor specific to the i-th asset ($i \in A$) as $\varepsilon_i(t)$.*

A8: *The idiosyncratic factors follow diffusion processes of the form:*

$$
d\varepsilon_i(t) = \varepsilon_i(t)\, \eta_i(\varepsilon_i(t), t)\, dt + \varepsilon_i(t)\, \delta_i(\varepsilon_i(t), t)\, dw_i(t), \tag{6.34}
$$

where η_i is the expected rate of change in ε_i, δ_i the diffusion coefficient, and w_i a standard Wiener process. Functions $\eta_i(\varepsilon_i(t), t)$ and $\delta_i(\varepsilon_i(t), t)$ are supposed bounded. These factors are assumed to be unrelated to the common ones, i.e., $Cov_t(d\varepsilon_i(t), d\phi_j(t)) = 0\ \forall i \in A$ and $j = 1, \ldots, k$.

Given the change in assumptions, we naturally redefine functions $P^i(t)$, $P^i(T)$, $\alpha_i(t)$, and $\gamma_i(t)$ as:

$$
\begin{aligned}
&P^i(t) = P^i(\phi(t), \varepsilon_i(t), t), \quad P^i(T) = \Phi^i(\phi(T), \varepsilon_i(T)), \\
&\alpha_i(t) = \alpha_i(\phi(t), \varepsilon_i(t), t), \quad \gamma_i(t) = \gamma_i(\phi(t), \varepsilon_i(t), t).
\end{aligned}
$$

In the same way as in Section 6.2, it can easily be seen that the dynamics of price $P^i(t)$ are described as

$$
\begin{aligned}
dP^i(t) &= \{\alpha_i(t) P^i(t) - \gamma_i(t)\}\, dt + P^i_\phi(t)\, I_\phi(t) \sum(t)\, dz(t) \\
&\quad + P^i_{\varepsilon_i}(t)\, \varepsilon_i(t)\, \delta_i(t)\, dw_i(t),
\end{aligned} \tag{6.35}
$$

where the expected change of the price is given by

$$\begin{aligned}&\alpha_i(t)P^i(t)-\gamma_i(t)=P^i_\phi(t)I_\phi(t)\mu(t)+P^i_{\varepsilon i}(t)\varepsilon_i(t)\eta_i(t)+P^i_t(t)\\&+(1/2)\operatorname{trace}\left\{P^i_{\phi\phi}(t)I_\phi(t)\Omega(t)I_\phi(t)\right\}+(1/2)P^i_{\varepsilon_i\varepsilon_i}(t)\varepsilon_i(t)^2\delta_i(t)^2,\end{aligned}\tag{6.36}$$

so that the return-generating process is written as

$$\begin{aligned}&\{dP^i(t)+\gamma_i(t)dt\}/P^i(t)-\alpha_i(t)dt=\hat{P}^i_{\phi_1}(t)\{d\phi_1(t)/\phi_1(t)-\mu_1(t)dt\}+\\&\cdots+\hat{P}^i_{\phi_k}(t)\{d\phi_k(t)/\phi_k(t)-\mu_k(t)dt\}+\hat{P}^i_{\varepsilon_i}(t)\{d\varepsilon_i(t)/\varepsilon_i(t)-\eta_i(t)dt\}.\end{aligned}\tag{6.37}$$

As shown by the last term on the right hand side of (6.37), the process now includes innovations in the idiosyncratic factor.

We finally set a well-diversification assumption according to the APT's spirit.

A9: *The number of assets, n, is large enough for investors to asymptotically diversify away the idiosyncratic risk.*

This permits us to apply the APT discussion to return-generating process (6.37). The result is: in the absence of asymptotic arbitrage, the linear equation (6.6) holds with a cross-sectional mean square error equal to zero, and it prices most of the assets correctly (see Huberman (1982) and Ingersoll (1984, 1987).

Corresponding to corollary 1, the partial differential equation can be derived by solving (6.36) for α_i and substituting the result into APT condition (6.6):

$$\begin{aligned}&(1/2)\operatorname{trace}\left\{P^i_{\phi\phi}(t)I_\phi(t)\Omega(t)I_\phi(t)\right\}+(1/2)P^i_{\varepsilon_i\varepsilon_i}(t)\varepsilon_i(t)^2\delta_i(t)^2\\&+P^i_\phi(t)I_\phi(t)(\mu(t)-\lambda(t))+P^i_{\varepsilon_i}(t)\varepsilon_i(t)\eta_i(t)+P^i_t(t)+\gamma_i(t)\\&=r(t)P^i(t)\quad\forall i\in A.\end{aligned}\tag{6.38}$$

The effects of asset idiosyncrasy are shown by the second and fourth terms on the left hand side of (6.38).

Equation (6.38) is solved in the form of formula (6.25) in just the same manner as in Subsection 6.4.1 in the text if by assumption A9 one neglects noise effects on the discounting trading strategy.

Finally, the same formula as equation (6.30) is obtained by applying Theorem 5.3 of Friedman (1975) again on equation (6.38) and the terminal condition. However, in this case expectation operator E^* must be defined as one with respect to the *non-adjusted* dynamics of the idiosyncratic factor (6.34) and the risk-adjusted processes of the common factors (6.29).

Bibliography

[1] Arnold, Ludwig, 1973, Stochastic differential equations: Theory and applications (John Wiley & Sons, Inc., New York).

[2] Black, Fischer and Myron Scholes, 1973, The pricing of options and corporate liabilities, Journal of Political Economy 81, 145-166.

[3] Brennan, Michael J. and Eduardo S. Schwartz, 1979, A continuous time approach to the pricing of bonds, Journal of Banking and Finance 3, 133-155

[4] Chamberlain, Gray, 1988, Asset pricing in multiperiod securities markets, Econometrica 56, 1283-1300

[5] Chang, Jack S. K. and Latha Shanker, 1987, Option pricing and the arbitrage pricing theory, Journal of Financial Research 10, 1-16.

[6] Cox, John C. and Stephen A. Ross, 1976, The valuation of options for alternative stochastic processes, Journal of Financial Economics 3, 145-166.

[7] Cox, John C., Jonathan E. Ingersoll, Jr. and Stephen A. Ross, 1981, A re-examination of traditional hypotheses about the term structure of interest rates, Journal of Finance 36, 769-799.

[8] Cox, John C., Jonathan E. Ingersoll, Jr. and Stephen A. Ross, 1985, An intertemporal general equilibrium model of asset prices, Econometrica 53, 363-384.

[9] Duffie, Darrell, 1988, Security markets: Stochastic models (Academic Press, New York).

[10] Dybvig, Philip H., 1983, An explicit bound on individual assets' deviations from APT pricing in a finite economy, Journal of Financial Economics 12, 483-496.

[11] Elton, Edwin J. and Martin J. Gruber, 1987, Modern portfolio theory and investment analysis 3rd ed. (John Wiley & Sons, Inc.,New York).

[12] Friedman, Avner, 1975, Stochastic differential equations and applications 1 (Academic Press, New York).

[13] Grinblatt, Mark and Sheridan Titman, 1983, Factor pricing in a finite economy, Journal of Financial Economics 12, 497-507.

[14] Harrison, J. Michael and David M. Kreps, 1979, Martingales arbitrage in multiperiod securities markets, Journal of Economic Theory 20, 381-408.

[15] Harrison, J. Michael and Stanley R. Pliska, 1981, Martingales and stochastic integrals in the theory of continuous trading, Stochastic Processes and Their Application 11, 215-260.

[16] Huang, Chi-fu, 1987, An intertemporal general equilibrium asset pricing model: The case of diffusion information, Econometrica 55, 117-142.

[17] Huberman, Gur, 1982, A simple approach to arbitrage pricing theory, Journal of Economic Theory 28, 183-191.

[18] Ikeda, Shinsuke, 1988, The continuous-time APT with diffusion factors and rational expectations: A synthesis, ISER Discussion Paper, Osaka University, No.175.

[19] Ikeda, Shinsuke, 1991, Arbitrage asset pricing under exchange risk, Journal of Finance 46, 447-455.

[20] Ingersoll, Jonathan E., Jr., 1984, Some results in the theory of arbitrage pricing, Journal of Finance 39, 1021-1039.

[21] Ingersoll, Jonathan E., Jr., 1987, Theory of financial decision making (Rowman and Littlefield, Totowa, New Jersey).

[22] Malliaris, A.G. and W.A. Brock, 1982, Stochastic methods in economics and finance (North-Holland, Amsterdam).

[23] Modigliani, Franco and Merton H. Miller, 1959, The cost of capital, corporation finance, and the theory of investment: reply, American Economic Review 49, 655-669.

[24] Richard, Scott F., 1978, An arbitrage model of the term structure of interest rates, Journal of Financial Economics 6, 33-57.

[25] Roll, Richard and Stephen A. Ross, 1980, An empirical investigation of the arbitrage pricing theory, Journal of Finance 35, 1073-1103.

[26] Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, Journal of Economic Theory 13, 341-360.

[27] Ross, Stephen A., 1977, Return, risk, and arbitrage, in I. Friend and J. L. Bicksler, eds., Risk and return in finance (Ballinger Publishing Co., Cambridge, Mass), 189-218.

[28] Rubinstein, Mark, 1976, The valuation of uncertain income streams and the pricing of options, Bell Journal of Economics 7, 407-425.

[29] Solnik, Bruno, 1983, International arbitrage pricing theory, Journal of Finance 38, 449-457.

[30] Vasicek, Oldrich, 1977, An equilibrium characterization of the term structure, Journal of Financial Economics 5, 177-188.

Chapter 7

Arbitrage Asset Pricing under Exchange Risk

Abstract: This paper extends the APT to an international setting. Specifying a linear factor return-generating model in local currency terms, we show that the usual risk-diversification rule in the APT does not yield a riskless portfolio unless currency fluctuations obey the same factor model as asset returns. We then consider an arbitrage portfolio whose exchange risk is hedged by foreign riskless bonds. Under the resulting no-arbitrage conditions, the expected returns are not on the same hyperplane, unlike the closed-economy APT, unless they are adjusted by the cost of exchange risk hedging.

JEL Classification Numbers: G12, G11.

Keywords: APT, exchange risk.

7.1 Introduction

In analyses of international economic phenomena, exchange risk is one of the most important elements to be considered. This paper studies arbitrage asset pricing in an international setting and shows how the introduction of exchange risk changes the Arbitrage Pricing Theory (APT) formulated in closed-economy models by Ross (1976, 1977), Huberman (1982), Ingersoll (1984), and others.

The extension of the APT to an international framework (IAPT) is successfully undertaken by Solnik (1983). He proves by straightforward application of the APT that, if the asset returns measured in an arbitrarily given numeraire. currency jointly follow a linear factor model, then, in the absence of arbitrage, the vector of expected returns in a given currency is spanned by the vector of ones and factor loading vectors.[1] But this direct applicability of the APT to an international setting is somewhat counterintuitive since it is natural to conjecture that exchange risk will introduce a new element into arbitrage activities. The key assumption in this puzzle is that the return-generating process is

[1] The same conclusion is reached by Ross and Walsh (1982). Levine (1989) generalizes Solnik's result in an inflationary model with purchasing power parity deviations. See also Kleidon and Pfleiderer (1983) for a discussion of the Solnik IAPT.

specified in a numeraire. currency. It can easily be seen that under this specification the exchange risk of asset returns is automatically diversified when one constructs an arbitrage portfolio according to the rule used in the APT. This enables Solnik to apply the closed-economy APT straightforwardly to international asset pricing.

Instead, we set up a linear factor return-generating process in local currency terms in Section 7.2 to emphasize the effect of exchange risk on international arbitrage asset pricing. We first show that the usual risk-diversification rule in the APT does not yield a risk-free portfolio in this setting if exchange rate fluctuations do not have the same factor structure as asset returns. Our main analysis presents a way of constructing an arbitrage portfolio hedged against exchange risk and verifies that expected returns are not on the same hyperplane unless they are adjusted by the cost of exchange risk hedging. Given this result, we also clarify the arbitrage determination of covariances between asset returns and exchange rate fluctuations. The results obtained are compared with Solnik's IAPT in Section 7.3.

7.2 Arbitrage Asset Pricing with Exchange Risk Hedging

7.2.1 A Linear Factor Model with Exchange Risk

We consider a world with N countries indexed 1 to N. In each country, there exist one risky asset and one locally riskless national bond which are freely traded in perfect international capital markets. The risky assets, as well as the national bonds, are assumed to be denominated in their respective local currencies. Given this assumption, we specify the generating process of the risky asset returns as the familiar linear K-factor model in local currency terms:

$$\tilde{r}_i^i = \bar{r}_i^i + b_{i1}\tilde{f}_1 + \cdots + b_{iK}\tilde{f}_K + \tilde{\varepsilon}_i \ (i = 1, \cdots, N). \tag{7.1}$$

In the equation above, $\tilde{r}_i^i$ is the random return on the i-th country risky asset in terms of the local currency, $\bar{r}_i^i$ denotes its expected value, the $\tilde{f}_k$ values ($k = 1$, $\cdots$, K) represent international common factors with zero means, b_{ik} denotes the sensitivity of return $\tilde{r}_i^i$ to fluctuations in factor k, and $\tilde{\varepsilon}_i$ is a nonsystematic risk component with zero mean, bounded variance, and $E[\tilde{\varepsilon}_i|\tilde{f}_k] = 0$ for all i and k.[2] As usual in the APT, the number of risky assets N is assumed to be large enough to permit the law of large numbers to hold.[3]

We assume flexible foreign exchange rates and denote their random processes by the general form:

$$\tilde{s}_i^j = \bar{s}_i^j + \tilde{\delta}_i^j \ \ (i, j = 1, \cdots, N), \tag{7.2}$$

where $\tilde{s}_i^j$ represents the random rate of appreciation of country i's currency in terms of country j's. Values $\bar{s}_i^j$ and $\tilde{\delta}_i^j$ denote the expected and random parts

[2] The $\tilde{\varepsilon}_i$ values can be correlated with each other.

[3] Since by assumption there is only one risky asset in each country, this means that the number of countries is large. If there are many risky assets in each country, we can easily consider an alternative world in which there are only a few countries without any change in our main results (below). This point will be discussed at the end of Section 7.3.

of currency variation $\tilde{s}_i^j$, respectively. By definition, $E[\tilde{\delta}_i^j] = 0$ for all i and j. Trivially, when $i = j$, $\tilde{s}_i^j$, $\bar{s}_i^j$, and $\tilde{\delta}_i^j$ are all equal to zero.

Here, in contrast to Solnik (1983), we do not assume that currency fluctuations have the same factor structure as equation (7.1). In order to explicitly demonstrate the implication of not making this assumption, consider risky asset returns from the viewpoint of a given currency, say currency n. If we define P_i^j and S_i^j as the currency j prices of asset i (the risky asset of country i) and currency i, respectively, then the price of asset i measured in terms of currency n is given by the law of one price as $P_i^n = P_i^i S_i^n$. Thus, by application of Ito's lemma, one can compute the currency n return on the asset as:[4]

$$\tilde{r}_i^n = \tilde{r}_i^i + \tilde{s}_i^n + \mathrm{Cov}(\tilde{r}_i^i, \tilde{s}_i^n) \quad (i = 1, \cdots, N), \tag{7.3}$$

where $\mathrm{Cov}(\cdot, \cdot)$ denotes covariance. Substitution of equations (7.1) and (7.2) (where $j = n$) into (7.3) yields:

$$\tilde{r}_i^n = \bar{r}_i^n + b_{i1}\tilde{f}_1 + \cdots + b_{iK}\tilde{f}_K + \tilde{\varepsilon}_i + \tilde{\delta}_i^n \quad (i = 1, \cdots, N), \tag{7.4}$$

where the expected return $\bar{r}_i^n$ is:

$$\bar{r}_i^n = \bar{r}_i^i + \bar{s}_i^n + \mathrm{Cov}(\tilde{r}_i^i, \tilde{s}_i^n). \tag{7.5}$$

The last term on the right hand side of equation (7.4), $\tilde{\delta}_i^n$, represents the exchange risk for country n's investors. If this term were assumed to be characterized by the same K-factor model as equation (7.1), then by substituting it for $\tilde{\delta}_i^n$ in (7.4), the asset returns measured in the numeraire currency n could be rewritten as being generated by the K-factor model as well.[5] Thus, the usual APT (in a closed economy setting) could be applied to that case as in the Solnik model. Empirically, however, it might be difficult to extract international common factors simultaneously, demonstrating both the price variations of risky assets (e.g., stocks) and currency fluctuations while keeping residual or idiosyncratic risk small. This matters because, as shown by Huberman (1982) and Ingersoll (1984), an increase in idiosyncratic risk worsens the fit of the APT equations.[6] Our formulation avoids this difficulty.

In our setting, however, it is impossible to construct riskless portfolio in the same manner as the usual APT because the exchange risk in equation (7.4), $\tilde{\delta}_i^n$, depends on the asset index i and, at the same time, is undiversifiable, unlike residual risk $\tilde{\varepsilon}_i$. To see this, construct a portfolio from N risky assets denoted by $\omega = (\omega_1, \cdots, \omega_N)'$ with ω_i being the investment proportions, according to the usual rule:

$$\left.\begin{array}{rcl} \omega' b_1 & = & 0, \\ & \vdots & \\ \omega' b_K & = & 0, \end{array}\right\} \tag{7.6}$$

$$\omega'\tilde{\varepsilon} \simeq 0, \tag{7.7}$$

[4] Following Solnik (1983), we define asset returns over a short interval of time and assume that the conditions required for Ito's calculus are met. See Roll and Ross (1980) and Ikeda (1991) for the continuous time APT.

[5] This case will be discussed explicitly in Section 7.3.

[6] See Cho, Eun, and Senbet (1986) for empirical research of the Solnik IAPT.

where $b_k = (b_{1k}, \cdots, b_{Nk})'$ and $\tilde{\varepsilon} = (\tilde{\varepsilon}_1, \cdots, \tilde{\varepsilon}_N)'$. Then, the currency n return on the portfolio can be computed from equations (7.4), (7.6), and (7.7) as:

$$\begin{aligned} \omega'\tilde{r}^n &= \omega'\bar{r}^n + \omega' b_1 \tilde{f}_1 + \cdots + \omega' b_K \tilde{f}_K + \omega'\tilde{\varepsilon} + \omega'\tilde{\delta}^n \qquad (7.8) \\ &= \omega'\bar{r}^n + \omega'\tilde{\delta}^n, \end{aligned}$$

where $\tilde{r}^n = (\tilde{r}_1^n, \cdots, \tilde{r}_N^n)'$, $\bar{r}^n = (\bar{r}_1^n, \cdots, \bar{r}_N^n)'$, and $\tilde{\delta}^n = (\tilde{\delta}_1^n, \cdots, \tilde{\delta}_N^n)'$. Equation (7.8) reveals that the simple hedging rule given by (7.6) and (7.7) does not yield a risk-free portfolio since the exchange risk $\omega'\tilde{\delta}^n$ would be left undiversified by this rule. The next subsection confronts this problem by considering an arbitrage portfolio which hedges against exchange risk by means of foreign lending and borrowing at locally riskless rates.

7.2.2 Arbitrage Pricing of Hedged Assets

Consider a portfolio of $2N$ assets, $\theta = (\theta_1, \cdots, \theta_{2N})'$, proportion θ_i $(i = 1, \cdots, N)$ of which is invested in risky asset i and proportion θ_{N+i} of which is invested in the national bond of country i. Based on the portfolio defined by equations (7.6) and (7.7), we now specify portfolio θ as:

$$\theta_i = \omega_i \quad (i = 1, \cdots, N), \qquad (7.9)$$

$$\theta_{N+i} = -\omega_i \quad (i = 1, \cdots, N). \qquad (7.10)$$

The resulting bundle is an arbitrage portfolio since equations (7.9) and (7.10) imply that investments in the risky assets are fully financed by holding the opposite positions of national bonds.

At the same time the following two facts imply that portfolio θ is riskless. First, as can be seen from equations (7.6), (7.7), and (7.9), the risk from common factors $\tilde{f}_k$ and residual factors $\tilde{\varepsilon}_i$ is diversified, and second, the investment in country i's risky asset is protected against exchange risk by opposite trading in the national bond of the same country. Explicitly, if $\tilde{\rho}_i^j$ denotes the currency j return on the national bond of country i (so that all of the ρ_i^i values are deterministic), return $\tilde{\rho}_i^n$ can be computed in the same way used to derive (7.3):

$$\tilde{\rho}_i^n = \rho_i^i + \bar{s}_i^n + \tilde{\delta}_i^n \quad (i = 1, \cdots, N). \qquad (7.11)$$

Thus, setting $\tilde{\rho}^n = (\tilde{\rho}_1^n, \cdots, \tilde{\rho}_N^n)'$, $\rho = (\rho_1^1, \cdots, \rho_N^N)'$, and $\bar{s}^n = (\bar{s}_1^n, \cdots \bar{s}_N^n)'$, the currency n return on portfolio θ, $\tilde{r}_\theta^n$, reduces to:

$$\begin{aligned} \tilde{r}_\theta^n &= \sum_{i=1}^{N} \theta_i \tilde{r}_i^n + \sum_{i=1}^{N} \theta_{N+i} \tilde{\rho}_i^n \qquad (7.12) \\ &= \omega'\tilde{r}^n - \omega'\tilde{\rho}^n \quad \text{(by (7.9) and (7.10))} \\ &= \omega'(\bar{r}^n - \rho - \bar{s}^n). \quad \text{(by (7.8) and (7.11))} \end{aligned}$$

This verifies that the return measured in currency n is free from any risk.[7]

[7] Exchange risk hedging by holding locally riskless bonds is also proposed in utility-based International Asset Pricing Models (IAPM). See, for example, Solnik (1973, 1974), Stulz (1981), and Adler and Dumas (1983).

Now that a riskless arbitrage portfolio has been obtained from the viewpoint of currency n, we can use the same algebraic argument as the APT to determine risk premia for country n's investors. For free lunches to be absent, the currency n return on portfolio θ must be equal to zero. Explicitly, from equation (7.12), the following equation must be valid:

$$\omega'(\bar{r}^n - \rho - \bar{s}^n) = 0. \tag{7.13}$$

Given the portfolio formation rule of equation (7.6), the condition of equation (7.13) requires that the expected net return vector $(\bar{r}^n - \rho - \bar{s}^n)$ be spanned by factor loading vectors b_k $(k = 1, \cdots, K)$ for some parameters $\lambda_1^n, \cdots, \lambda_K^n$:

$$\bar{r}^n - \rho - \bar{s}^n = \lambda_1^n b_1 + \cdots + \lambda_K^n b_K.$$

Note here that portfolio θ retains the property of a riskless arbitrage portfolio even if its return is evaluated in any currency other than currency n. It follows that the above discussion is applicable to excluding arbitrage opportunities in terms of any currency. This leads to our main result:

Proposition: *Suppose that risky asset returns and currency fluctuations are generated by processes* (7.1) *and* (7.2), *respectively. Then, if no arbitrage opportunities are left unexploited in terms of any currency, there exist N sets of K scalars, $\lambda_1^j, \cdots, \lambda_K^j$, such that* :

$$\bar{r}^j - \rho - \bar{s}^j = \lambda_1^j b_1 + \cdots + \lambda_K^j b_K \qquad (j = 1, \cdots, N). \tag{7.14}$$

The left hand side of (7.14) denotes the expected net returns in terms of currency j on risky assets whose exchange risk is hedged by foreign borrowing and lending.[8] On the other hand, given the local currency specification of the return-generating process (7.1), it is natural to regard b_{ik} as a measure of local factor risk, *i.e.*, factor risk in local currency terms. Thus, the theorem asserts that, in each country or currency, there are K local factor risk prices, λ_k^j, which determine the expected net returns on hedged assets. Put otherwise, the expected returns themselves are not on the same hyperplane, unlike the closed-economy APT, unless they are adjusted by the cost of exchange risk hedging represented by $\rho + \bar{s}^j$.

Formally, if variables $\bar{s}_i^j$, ρ_i^i, and λ_k^j are given, the no-arbitrage conditions of equation (7.14) yield N^2 expected returns in terms of different currencies, $\bar{r}_i^j$. To see the mechanism from another point of view, let us substitute equation (7.5) (setting $n = j$) for $\bar{r}_i^j$ in the i-th row of (7.14). We obtain:

$$\bar{r}_i^i + \mathrm{Cov}(\tilde{r}_i^i, \tilde{s}_i^j) - \rho_i^i = \lambda_1^j b_{i1} + \cdots + \lambda_K^j b_{iK}. \tag{7.15}$$

This provides an alternative expression showing that N^2 no-arbitrage conditions yield N expected values of the risky assets' returns in terms of respective

[8] To define the concept of exchange risk hedging used here more precisely, it is one in which the resulting excess return on an asset is affected by currency fluctuations only insofar as they affect the local currency return on the asset. Thus, these excess returns do not correspond to returns hedged against currency risk in a minimum-variance sense because the covariance of excess returns with currency fluctuations is not zero.

local currencies and $N(N-1)$ covariances between the risky assets' returns and currency fluctuations.

The arbitrage determination of asset-currency covariances can be shown as follows. First, take the difference between the no-arbitrage conditions given by equation (7.15) for two different currencies, say currencies j and ℓ. Next, note that $\mathrm{Cov}(\tilde{r}_i^i, \tilde{s}_i^j) - \mathrm{Cov}(\tilde{r}_i^i, \tilde{s}_i^\ell) = \mathrm{Cov}(\tilde{r}_i^i, \tilde{s}_\ell^j)$. Then, we obtain our second proposition:

Corollary: *Suppose that the conditions in the theorem hold. Then, in the absence of arbitrage, the covariances between local currency returns on risky assets and exchange rate fluctuations are given by:*

$$\mathrm{Cov}(\tilde{r}, \tilde{s}_\ell^j) = \lambda_1^{j\ell} b_1 + \cdots + \lambda_K^{j\ell} b_K \qquad (j, \ell = 1, \cdots, N), \tag{7.16}$$

where $\mathrm{Cov}(\tilde{r}, \tilde{s}_\ell^j) = (\mathrm{Cov}(\tilde{r}_1^1, \tilde{s}_\ell^j), \cdots, \mathrm{Cov}(\tilde{r}_N^N, \tilde{s}_\ell^j))'$ *and* $\lambda_k^{j\ell} = \lambda_k^j - \lambda_k^\ell$.

This shows that covariance vector $\mathrm{Cov}(\tilde{r}, \tilde{s}_\ell^j)$ is also spanned by factor loadings b_k. But here the weights applied are risk price differences $\lambda_k^{j\ell}$. A close look at the derivation reveals the reason for this. By definition, covariance $\mathrm{Cov}(\tilde{r}_i^i, \tilde{s}_\ell^j)$ equals the difference between the expected net return on a hedged investment in risky asset i in terms of currency j, $\bar{r}_i^i +$ $\mathrm{Cov}(\tilde{r}_i^i, \tilde{s}_i^j) - \rho_i^i$, and that in terms of currency ℓ, $\bar{r}_i^i + \mathrm{Cov}(\tilde{r}_i^i, \tilde{s}_i^\ell) - \rho_i^i$. From equation (7.14), on the other hand, it is international discrepancies in factor risk prices that cause cross-currency differences in expected net returns on hedged investments. It follows that the asset-currency covariance $\mathrm{Cov}(\tilde{r}_i^i, \tilde{s}_\ell^j)$ must be determined so as to precisely match the effect of cross-currency differences in factor risk prices.

7.3 A Comparison with Solnik's IAPT

To clarify the implications of our results, we finally compare our discussion with that of Solnik (1983). He specifies the return-generating process in a numeraire. currency which is arbitrarily chosen. As pointed out in Subsection 7.2.1, his analysis can be replicated in our setting by additionally assuming that currency fluctuations have the same factor structure as the asset returns given by equation (7.1). For example, let currency variation $\tilde{\delta}_i^n$ be generated by:

$$\tilde{\delta}_i^n = c_{i1}^n \tilde{f}_1 + \cdots + c_{iK}^n \tilde{f}_K + \tilde{\mu}_i^n \qquad (i = 1, \cdots, N,\ i \neq n), \tag{7.17}$$

where c_{ik}^n denotes the factor loading and $\tilde{\mu}_i^n$ represents an idiosyncratic risk component with zero mean, bounded variance, and $E[\tilde{\mu}_i^n | \tilde{f}_k] = 0$ $(k = 1, \cdots, K)$. Then, equation (7.4) reduces to:

$$\tilde{r}_i^n = \bar{r}_i^n + b_{i1}^n \tilde{f}_1 + \cdots + b_{iK}^n \tilde{f}_K + \tilde{\varepsilon}_i^n, \tag{7.18}$$

where $b_{ik}^n = b_{ik} + c_{ik}^n$ and $\tilde{\varepsilon}_i^n = \tilde{\varepsilon}_i + \tilde{\mu}_i^n$.

This means that the returns in terms of numeraire. currency n follow a linear K-factor model, which is the situation Solnik supposes. As he proves, this factor structure is invariant to the currency chosen if the same factor model as equation

(7.17) holds for all currencies. Applying the usual arbitrage argument to the model, we then obtain Solnik's result:

$$\bar{r}^j - \rho_j^j 1 = \pi_1^j b_1^j + \cdots + \pi_K^j b_K^j \qquad (j = 1, \cdots, N), \tag{7.19}$$

where 1 is an N-dimensional vector of ones, the π_k^j values represent factor risk prices, and the b_k^j values are factor loading vectors. Equation (7.19) shows that if expected returns are measured in the same currency, all of them are on the same hyperplane as in the case of the closed-economy APT despite the existence of exchange risk.

Equation (7.18) (and so (7.17)) plays a key role in producing this result. In this setting, the factors $\tilde{f}_k$ also generate currency fluctuations. Naturally, the factor loadings b_{ik}^j $(= b_{ik} + c_{ik}^j)$ can be regarded as measures of total or composite factor risk, i.e., factor risk in local currency terms plus exchange risk and parameter π_k^j in (7.19) as the prices of total factor risk including the price of exchange risk. In this sense, Solnik demonstrates international arbitrage pricing in terms of *total* factor risk prices.

In contrast, our model (in the previous section) considers a case in which one cannot successfully extract common factors simultaneously generating asset returns and currency fluctuations. Our theorem avoids this difficulty by explaining arbitrage pricing of hedged securities in terms of *local* factor risk prices.

We can also derive Solnik's result concerning asset-currency covariances from equation (7.19) in a way similar to that used to obtain (7.16):[9]

$$\mathrm{Cov}(\tilde{r}^j, \tilde{s}_\ell^j) = \pi_1^{j\ell} b_1^j + \cdots + \pi_K^{j\ell} b_K^j \qquad (j, \ell = 1, \cdots, N), \tag{7.20}$$

where $\mathrm{Cov}(\tilde{r}^j, \tilde{s}_\ell^j)$=$(\mathrm{Cov}(\tilde{r}_1^j, \tilde{s}_\ell^j), \cdots, \mathrm{Cov}(\tilde{r}_N^j, \tilde{s}_\ell^j))'$ and $\pi_k^{j\ell} = \pi_k^j - \pi_k^\ell$.

This shows that the covariances between asset returns measured in a given currency, rather than respective local currencies, and currency fluctuations are on the same hyperplane. Again, the result depends crucially on the specification of the return-generating process in terms of a numeraire. currency. In this case, factor loadings are measures of *total* factor risk, and therefore the set of K factor loadings measures the risk of return variations (correlated with currency fluctuations) in *numeraire.* currency terms. In contrast, according to the result in our corollary, if the return-generating process is specified in local currency terms, arbitrage ensures a similar linear relation between the covariance of the *local* currency return on a risky asset with currency variations and factor loadings b_{ik}, which are measures of *local* factor risk.

Our model can easily be extended to the case in which there is more than one risky asset denominated in each currency. Such an extension incurs the marginal cost of notational complexity (so that we do not treat this case explicitly) but yields the marginal benefit of making it possible, in the absence of a locally riskless bond, to conduct the analysis using a zero beta portfolio constructed from risky assets of the same nationality. Although the results are essentially the

---

[9] As Solnik (1983) shows, returns on locally riskless bonds are also spanned in the same way as (7.19):

$$\rho_j^j + \bar{s}_j^l = \rho_l^l + \pi_1^l c_{j1}^l + \cdots + \pi_K^l c_{jK}^l.$$

This relation is used to derive equation (7.20).

same as presented here, one additional result is that, in this case, the returns on risky assets of the same nationality are, if they are large in number, spanned in the same way as in the usual APT. This result corresponds to the aforementioned result of Solnik. This is not surprising because the return process of assets of the same nationality satisfies his assumption concerning asset returns.

7.4 Conclusions

This paper has pursued an international extension of the APT. In order to focus on the effect of exchange risk on arbitrage asset pricing, we have specified a linear factor model in local currency terms (rather than numeraire. currency terms). In this setting, the usual risk-diversification rule in the APT does not yield a risk-free portfolio because the exchange risk of an asset is undiversifiable and at the same time varies depending on the nationality of the asset. Confronted with this problem, we have considered an arbitrage portfolio which hedges against exchange risk by means of foreign lending and borrowing at locally riskless rates. The resulting no-arbitrage conditions require that the expected net returns on risky assets whose exchange risk is covered by foreign borrowing and lending be determined as linear combinations of factor loadings. This means that the expected returns themselves are not on the same hyperplane unless they are adjusted by the cost of exchange risk hedging.

Bibliography

[1] Adler, Michael and Bernard Dumas, 1983, International portfolio choice and corporation finance: A synthesis, The Journal of Finance 38, 925-984.

[2] Cho, D. Chinhyung, Cheol S. Eun, and Lemma W. Senbet, 1986, International arbitrage pricing theory: An empirical investigation, The Journal of Finance 41, 313-329.

[3] Huberman, Gur, 1982, A simple approach to arbitrage pricing theory, Journal of Economic Theory 28, 183-191.

[4] Ikeda, Shinsuke, 1991, The continuous-time APT with diffusion factors and rational expectations: A synthesis, The Economic Studies Quarterly 42, 124-138.

[5] Ingersoll, Jr., Jonathan E., 1984, Some results in the theory of arbitrage pricing, The Journal of Finance 39, 1021-1039.

[6] Kleidon, Allan W. and Paul Pfleiderer, 1983, Discussion, The Journal of Finance 38, 470-472.

[7] Levine, Ross, 1989, An international arbitrage pricing model with PPP deviations, Economic Inquiry 27, 587-599.

[8] Roll, Richard and Stephen A. Ross, 1980, An empirical investigation of the arbitrage pricing theory, The Journal of Finance 35, 1073-1103.

[9] Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, Journal of Economic Theory 13, 341-360.

[10] Ross, Stephen A., 1977, Return, risk, and arbitrage, in I. Friend and J. Bicksler, eds.: Risk and Return in Finance (Ballinger, Cambridge, MA).

[11] Ross, Stephen A. and Michael M. Walsh, 1982, A simple approach to the pricing of risky assets with uncertain exchange rates, Research in International Business and Finance 3, 39-54.

[12] Solnik, Bruno H., 1973, European Capital Markets (D.C. Heath, Lexington).

[13] Solnik, Bruno H., 1974, An equilibrium model of the international capital market, Journal of Economic Theory 8, 500-524.

[14] Solnik, Bruno H., 1983, International arbitrage pricing theory, The Journal of Finance 38, 449-457.

[15] Stulz, René M., 1981, A model of international asset pricing, Journal of Financial Economics 9, 383-406.

Chapter 8

An Intertemporal Capital Asset Pricing Model with Stochastic Differential Utility

Abstract: Intertemporal capital asset pricing and the stochastic properties of optimal portfolios and consumption are examined in a continuous-time recursive utility (stochastic differential utility) model with multiple state variables dynamically affecting investment opportunities. Although Merton-Richard's multi-beta intertemporal CAPM (ICAPM) relationships are valid, they do not collapse to the consumption-based single beta CAPM (CCAPM). Instead, several multi-beta versions of CCAPM are proposed in both heterogeneous and homogeneous agents settings. The multiple correlation coefficient of the individual agent's optimal consumption with the aggregate consumption and multiple state variables is unity.

JEL Classification Numbers: G12, G11.

Keywords: Asset pricing, risk premium, consumption beta, stochastic differential utility, multi-fund separation.

8.1 Introduction

This paper examines intertemporal capital asset pricing and the stochastic properties of optimal portfolio-consumption decisions in a continuous-time exchange economy in which the lifetime utility of households is represented by time-nonadditive recursive preferences. Extending the traditional static capital asset pricing model (CAPM) [e.g., Sharpe (1964) and Lintner (1965)] to a dynamic model with stochastic investment opportunities, Merton (1973) and Richard (1979) constructed an intertemporal capital asset pricing model (ICAPM). They proved: (i) the multi-beta linear structure in risk premium determination; and (ii) the multi-fund separation theorem. Although they consider asset pricing from the viewpoint of rational consumption choice, the direct and close relationship between the consumption rate and asset pricing was not clearly understood until Breeden's (1979) work appeared. Defining a risk measure by the 'consumption beta,' which is given as the covariance between the asset return and changes in aggregate consumption divided by the variance of aggregate consumption changes, he shows that the multi-beta linear structure of equilibrium risk premia can be represented in terms of a single beta, i.e., the consumption beta.

These intertemporal versions of CAPM, however, have been criticized for both theoretical and empirical reasons. First, these analyses are based on the restrictive assumption that investors' preferences have time-additive von Neumann-Morgenstern representations. As is well known [e.g., Kreps and Porteus (1978) and Epstein and Zin (1989)], this utility function cannot disentangle attitudes towards risk and towards time. Second, the ICAPMs presented above seem to be inconsistent with empirical data, as pointed out by Mehra and Prescott (1985), Mankiw and Shapiro (1986), Breeden, Gibbons, and Litzenberger (1989), and Hansen and Jagannathan (1991).

Based on the theory of stochastic differential utility (SDU), which is developed by Duffie and Epstein (1992a, 1992b) and Duffie and Lions (1992) as a continuous-time version of non-separable recursive utility,[1] this paper reexamines equilibrium intertemporal capital asset pricing and the stochastic properties of optimal portfolio-consumption decisions.[2] Specifically, I do that by applying SDU to the Breeden (1979)-Richard (1979) model with X-dimensional Brownian information continuously affecting investment opportunities. To generalize ICAPM by using recursive utility, the SDU is useful and tractable in two points: First, dynamic optimality conditions can be easily derived using a modified Bellman equation developed by Duffie and Epstein (1992a, 1992b); and secondly, the SDU formulation retains the local linear structure of the continuous time model.

The main results obtained in this exercise can be summarized as follows. First, under any preference structures defined by SDU, the ICAPM relationships derived by Merton (1973) and Richard (1979), i.e., the $X+2$ fund separation theorem and the multi-beta structure, are still valid. Secondly, however, the multi-beta structure of risk premia cannot be reduced to the single-beta representation in terms of aggregate consumption. That is, Breeden's CCAPM does not hold in time-nonadditive recursive utility models. Thirdly, if individuals' preferences are heterogeneous, risk premia cannot be generally given as a linear combination of the market and consumption betas, in contrast to the two-beta CAPM developed by Epstein and Zin (1989) and Giovannini and Weil

[1] An underlying formulation of recursive utility is developed by Epstein and Zin (1989) within a discrete time framework. See also Svensson (1989) and Ma (1993). Svensson develops a continuous-time extension of time-nonadditive utility, limiting his attention to the case of constant elasticity of intertemporal substitution and constant relative risk aversion. Ma examines the equilibrium property of a recursive-utility model with heterogeneous agents.

[2] I can also refer to studies in which dynamic asset pricing is examined under an alternative generalized class of preferences: habit persistent preferences. See, for example, Sundaresan (1989), Constantinides (1990), Detemple and Zapatero (1991), and Ingersoll (1992). In these articles past consumption directly affects current utility, while in recursive utility models, with which I am concerned here, the consumption history affects current utility only by changing current wealth. Moreover, in contrast to recursive preferences, habit formation models assume von Neumann-Morgenstern type utility, which cannot distinguish attitudes towards risk from attitudes towards time.

(1989). Finally, the multiple correlation coefficient of the individual's optimal consumption with the aggregate consumption and X state variables is unity.

The paper is structured as follows. In Section 8.2 a model of perfect capital markets which is characterized by stochastic investment opportunities and by the SDU is presented. It is shown that the standard Merton-Richard ICAPM is valid even under the time-nonadditive recursive utility represented by the SDU. Section 8.3 examines equilibrium risk premium determination from the viewpoint of aggregate consumption and the market portfolio, developing a multifactor CCAPM. Section 8.4 explores the stochastic properties of optimal consumption. Section 8.5 contains conclusions.

8.2 The Model

Let us extend the Breeden (1979)-Richard (1979) model of intertemporal capital asset pricing by specifying investors' attitudes towards risk and time by means of stochastic differential utility (SDU): Consider an exchange economy populated with K certain-lived rational investors who maximize lifetime utility from consumption streams. Underlying the model is a complete probability space (Ω, F, P) where Ω is the set of states of nature, F is the σ-field of events, and P is a probability measure on (Ω, F). An A-dimensional standard Wiener process, z_a, and an X-dimensional standard Wiener process, z_x, are defined on (Ω, F, P). The flow of information is described by the filtration $\{F_t\}$ generated by these multi-dimensional standard Wiener processes.

There is a single perishable consumption good, which serves as the numeraire. Consumers can trade $(A+1)$ assets: A risky capital assets and one instantaneously riskless asset. Individual k has wealth $W^k(t)$ at time t, continuously allocating it to the consumption good and the assets. These assets are traded in perfect competitive markets that are frictionless. Trading is possible only at equilibrium prices. The capital assets yield stochastic rates of return (or simply, stochastic returns), whose probability distributions depend on the state of the economy. The state of the economy is described by an X-dimensional Markovian vector process.

Given this basic structure, I characterize investment opportunities with the following stochastic differential equation for risky-asset prices:[3]

$$dP = \{I_a \mu_a(x,t) - \delta_a(x,t)\}\, dt + I_a \sigma_a(x,t)\, dz_a, \tag{8.1}$$

where P represents the $(A \times 1)$ vector of the prices of risky assets; I_a is an $(A \times A)$ diagonal matrix whose i-th diagonal element is the price of the i-th risky asset; μ_a is the $(A \times 1)$ vector of expected rates of total return on risky assets; δ_a is the $(A \times 1)$ vector of payout flows; and σ_a is the $(A \times A)$ matrix of diffusion coefficients of return processes.

Concerning (8.1), note first that the stochastic vector process of total returns, which are the sum of capital gains dP plus income gains $\delta_a dt$, is given by $I_a \mu_a dt + I_a \sigma_a dz_a$. Secondly, for a solution to (8.1) to exist, it is sufficient that functions μ_a and σ_a satisfy the regularity conditions known as the Lipschitz

[3] For detailed information on stochastic differential equations, see, for example, Arnold (1973).

condition and the growth condition.[4] This condition is assumed to be satisfied here. I also make the same assumption in (8.2), introduced just below.

The investment opportunities specified above are conditioned by an X-dimensional Markov state process x, which affects at each instant the mean return vector μ_a and the diffusion coefficient matrix σ_a. The stochastic process of x is given by:

$$dx = \eta_x(x,t)\,dt + \sigma_x(x,t)\,dz_x, \tag{8.2}$$

where η_x is the $(X \times 1)$ vector of means; and σ_x the $(X \times X)$ matrix of diffusion coefficients.

The instantaneously (or locally) riskless asset yields the risk-free rate, r. This interest rate also depends on state variable vector x, and hence fluctuates over time randomly.

Given the investment opportunities, an investor, say the k-th investor, chooses his consumption process, c^k, and portfolio allocation so as to maximize his lifetime utility, $V^k(t)$. The lifetime utility is specified by the SDU. It is defined as a solution to the following recursive equation:

$$V^k(t) = E_t\left[\int_t^{T^k} f^k\left(c_s^k, V_s^k\right) ds\right], \tag{8.3}$$

where $T^k > 0$ is investor k's time of death,[5] and E_t is the expectation operator conditional on the state of the economy at time t. $f^k\left(c_s^k, V_s^k\right)$ denotes an aggregator function which evaluates the time-s consumption rate, depending on lifetime utility at that time, $V^k(s)$.[6] It is assumed that this function is concave and strictly increasing in consumption. It is also assumed that the aggregator function f is continuous, Lipschitz in utility, and satisfies a growth condition in consumption. These regularity conditions ensure the existence of an SDU process satisfying (8.3).[7]

Note here that the SDU presented above contains as special cases familiar classes of preference structures. In particular, if the aggregator function is given by:

$$f^k = u^k\left(c^k\right) - \beta^k V^k \text{ and hence } f_{cV}^k = 0,$$

where β is a strictly positive constant, and the subscripts denote partial derivatives, then, the recursive utility function given by (8.3) reduces to the time-

[4] See Arnold (1973, p. 105).

[5] For simplicity, time horizon T^k is assumed finite here, as in Breeden (1979), Merton (1973), and Duffie and Epstein (1992b). By assuming some additional technical condition, the utility representation given by (8.3) obtains also in the case of infinite time horizon $T^k \to \infty$. See Duffie and Epstein (1992a, Appendix C).

[6] In Duffie and Epstein (1992a, b), the SDU is primitively defined by a pair of aggregator functions, one of which determines the degree of intertemporal substitution of consumption, and the other of which denotes a local measure of risk aversion. As proven there, this primitive utility can be represented in terms of one normalized aggregator after a simplifying change of variables. In (3), function f denotes this normalized aggregator.

[7] See theorem 1 of Duffie and Epstein (1992a, p. 366).

additive von Neumann-Morgenstern utility function,[8]

$$V^k(t) = E_t\left[\int_t^{T^k} u^k\left(c_s^k\right)\exp\left\{-\beta^k(s-t)\right\}ds\right].$$

Letting w^k denote the $(A \times 1)$ vector of portfolio proportions invested in risky assets and, hence, $1 - \sum_{a=1}^{A} w_a^k$ denote the investment proportion of the riskless asset, the budget constraint for each investor is given by:

$$dW^k = \left\{w^{k'}(\mu_a - r\iota)W^k + rW^k + y^k - c^k\right\}dt + w^{k'}\sigma_a W^k dz_a, \qquad (8.4)$$

where ι is the vector of ones and y^k denotes labor income, the process of which is exogenously given in this paper.

Let $J^k\left(W^k, x, t\right)$ be the maximum lifetime utility that is attainable under given W^k, x, t, and let Ω_{ij} $(i, j = a, x)$ be instantaneous variance-covariance matrices. From Duffie and Epstein (1992a)'s proposition 9, controls $c^k\left(W^k, x, t\right)$ and $w^k\left(W^k, x, t\right)$ give the optimal consumption- and portfolio-decision if they solve the following modified Bellman equation subject to constraint (8.4):

$$\sup_{(c^k, w^k)} DJ^k\left(W^k, x, t\right) + f^k\left\{c^k, J^k\left(W^k, x, t\right)\right\} = 0,$$

where:

$$\begin{aligned} DJ^k\left(W^k, x, t\right) = {} & J_t^k + J_W^k\left\{w^{k'}(\mu_a - r\iota)W^k + rW^k + y^k - c^k\right\} + J_x^k\eta_x \\ & + \tfrac{1}{2}w^{k'}\Omega_{aa}w^k\left(W^k\right)^2 J_{WW}^k + w^{k'}W^k\Omega_{ax}J_{xW}^k + \tfrac{1}{2}\Omega_{xx}J_{xx}^k. \end{aligned}$$

Assume the interior optimum. Then the first-order conditions are given by:

$$f_c^k\left(c^k, J^k\right) = J_W^k\left(W^k, x, t\right), \qquad (8.5)$$

$$(\mu_a - r\iota)J_W^k + \Omega_{aa}w^k W^k J_{WW}^k + \Omega_{ax}J_{xW}^k = 0, \qquad (8.6)$$

the latter of which can be solved for the asset demand:

$$w^k W^k = \Omega_{aa}^{-1}(\mu_a - r\iota)\left(-J_W^k/J_{WW}^k\right) - \Omega_{aa}^{-1}\Omega_{ax}\left(J_{xW}^k/J_{WW}^k\right). \qquad (8.7)$$

Note that the aggregator function $f^k(c^k, V^k)$, which specifies the intertemporal recursive structure of utility, does not directly appear in (8.6) or (8.7).

[8] For another special case, let β^k be a function of consumption,

$$f^k = u^k\left(c^k\right) - \beta^k\left(c^k\right)V^k.$$

The resulting utility function is the one of the Uzawa type. The Kreps-Porteus (1978) utility function is obtained by setting:

$$f(c, V) = \frac{\beta}{\rho}\cdot\frac{(c)\rho - (\alpha V)^{\rho/\alpha}}{V^{(\rho/\alpha - 1)}}, \quad 0 \neq \rho \leq 1, \ \beta \geq 0, \ \alpha \leq 1,$$

where superscript k is suppressed; ρ represents the elasticity of intertemporal substitution in consumption; and α denotes a parameter of risk aversion. This Kreps-Porteus utility does not satisfy the Lipschitz condition which is a part of a sufficient condition for the existence of an SDU process. Given the Brownian information structure, however, the SDU can be shown to exist uniquely. See Duffie and Epstein (1992a, p. 367).

Indeed, it could be seen that, given the value function, the conditions for optimal portfolios, (8.6) or (8.7), are essentially identical to those which are derived in the Merton (1973)-Richard (1979) ICAPM with the standard time-additive von Neumann-Morgenstern utility. As might be supposed, their ICAPM relationships are valid also in the present recursive utility model.

Proposition 1 *(validity of the Merton-Richard ICAPM): For any preference structures defined by SDU in (8.3), the Merton-Richard ICAPM remains valid in the following two senses: (i) In equilibrium each investor can choose their optimal portfolios from $(X+2)$ funds, i.e., the instantaneously riskless asset, the market portfolio (aggregate wealth), and the X hedging portfolios;* [9] *and (ii) equilibrium risk premia of any risky assets have the linear multi-beta structure,*

$$\mu_a - r\iota = \beta_{aM}\left(\mu_M - r\right) + \beta_{ax}\left(\mu_x - r\iota\right), \tag{8.8}$$

where μ_M denotes the expected return on the market portfolio; μ_x is the $(X \times 1)$ vector of the expected returns on the hedging portfolios;[10] *and the elements of $(A \times 1)$ vector β_{aM} and $(A \times X)$ matrix β_{ax} are multiple-regression betas of A risky assets on the market and the hedging portfolios respectively.*

Note that this proposition does not imply the irrelevance of the recursive structure of utility to risk premium determination. Preference structures affect the maximum value process J^k, which determines the risk premia of the $(X+1)$ basis securities which appear on the right hand side of (8.8).

8.3 Aggregate Consumption, the Market Portfolio, and Capital Asset Pricing

8.3.1 Aggregate Consumption and Capital Asset Pricing

Let us now reconsider capital asset pricing presented above from the viewpoint of the aggregate consumption rate. As proven by Breeden (1979), given the von Neumann-Morgenstern utility function, the multi-beta structure of risk premia as in (8.8) reduces to a single-beta representation where the beta coefficient is defined in terms of the covariance with changes in the aggregate consumption rate, i.e., the consumption beta. The same proposition is, however, not valid in the present model with non-separable recursive utility, as I shall now show.

From (8.6), risk premia must satisfy:

$$\mu_a - r\iota = \Omega_{aW_k}\left(-J^k_{WW}/J^k_W\right) + \Omega_{ax}\left(-J^k_{xW}/J^k_W\right), \tag{8.9}$$

[9] To be precise, proposition (i) in proposition 1 is composed of two statements: The first one is the multi-fund separation theorem. It states that there exist $X+2$ mutual funds each composition of which does not depend on investors' preferences such that any optimal portfolios can be chosen from these funds. This proposition is concerned with subjective optimality, but not with market equilibrium. The second statement is that equilibrium returns on the market portfolio are linearly independent of returns on X hedging portfolios so that the market portfolio, together with the X hedging portfolios and the riskless asset, can comprise the $X+2$ mutual funds.

[10] Return vector μ_x is not identical with the drift vector of x, η_x, unless the state variables are prices of marketed assets.

where Ω_{aW_k} denotes the $(A \times 1)$ vector of covariances between asset returns and changes in individual k's wealth, dW^k, i.e., $\Omega_{aW_k} = \Omega_{aa} w^k W^k$. In order to rewrite this equation in terms of the covariances of returns with changes in consumption, note from (8.5) that agent k's optimal consumption can be taken as a function of W^k, x, and t; $c^k = c^k\left(W^k, x, t\right)$. Applying Ito's Lemma to this function, I obtain the vector of covariances between asset returns and changes in agent k's consumption rate, Ω_{ac_k}, as:

$$\Omega_{ac_k} = \Omega_{aW_k} c_W^k + \Omega_{ax} c_x^k.$$

Solve this equation for Ω_{aW_k}, and substitute the result into (8.9). Then, after some manipulation, I can obtain

$$H^k\left(\mu_a - r\iota\right) = \Omega_{ac_k} + \Omega_{ax} L_x^k, \tag{8.10}$$

where $H^k$and L_x^k are given by:

$$H^k = -c_W^k J_W^k / J_{WW}^k, \tag{8.11}$$

$$L_x^k = c_W^k J_{xW}^k / J_{WW}^k - c_x^k. \tag{8.12}$$

Equation (8.10) can be aggregated over k. Letting C denote the aggregate consumption rate: $C = \sum_k c^k$, I obtain

$$\mu_a - r\iota = \Omega_{aC}\left(1/H\right) + \Omega_{ax}\left(L_x/H\right), \tag{8.13}$$

where $H = \sum_k H^k$ and $L_x = \sum_k L_x^k$.

In equation (8.13), equilibrium risk premia are expressed not only in terms of covariances between asset returns and changes in the aggregate consumption rate but also in terms of covariances between asset returns and changes in state variables. As can be seen from (8.13), the Breeden-type single-beta CCAPM holds valid only if L_x equals zero. I can prove that the validity of this condition crucially depends on whether or not f_{cV}^k equals zero; that is, whether or not the utility function is time-additive.

Proposition 2 *(consumption betas and risk premia):*

(i) The multi-beta ICAPM given by (8.8) reduces to the single-beta CCAPM if and only if $f_{cV}^k = 0$ so that the utility function is of the time-additive type.[11]

(ii) Suppose that $f_{cV}^k \neq 0$. Then, if the variance-covariance matrix of the aggregate consumption growth rate and changes in state variables, $\Omega_{\ln C, x}$, has full rank, and if capital markets are complete, equilibrium risk premia are determined by the following multi-beta version of the CCAPM:

$$\mu_a - r\iota = \theta_{aC}\left(\mu_C - r\right) + \theta_{ax}\left(\mu_x - r\iota\right), \tag{8.14}$$

where μ_C denotes the expected return on a portfolio which is perfectly correlated with changes in the aggregate consumption. θ_{aC} and θ_{ax} are multiple-regression betas on returns on the consumption portfolio and the X hedging portfolios.

[11] The invalidity of the single-beta CCAPM has been verified in a more specific context by Bergman (1985).

Proof: (i) From (8.13), the single consumption-beta CCAPM remains valid if and only if $L_x \ (= \sum_k L_x^k) = 0$. Recalling that $c^k = c^k \left(W^k, x, t\right)$, differentiate both sides of (8.5) with respect to x to obtain:

$$J_{xW}^k = f_{cc}^k c_x^k + f_{cV}^k J_x^k.$$

Substitution of this into (8.12) yields:

$$L_x^k = f_{cV}^k \left(c_W^k J_x^k - c_x^k J_W^k\right) / J_{WW}^k.$$

This equation implies that L_x^k equals zero if and only if f_{cV}^k equals zero since $c_W^k J_x^k - c_x^k J_W^k$ is not identically zero. This completes the proof of (i) of proposition 1.

(ii) Suppose that $f_{cV}^k \neq 0$. If capital markets are complete, one can construct the X hedging portfolios and the consumption portfolio returns of which are perfectly and positively correlated with the rate of aggregate consumption growth. From (8.13), these mimicking portfolios must satisfy:

$$\begin{bmatrix} \mu_C - r \\ \mu_x - r\iota \end{bmatrix} = \Omega_{\ln C, x} \begin{bmatrix} C/H \\ L_x/H \end{bmatrix}.$$

If the variance-covariance matrix $\Omega_{\ln C,x}$ has full rank, this equation can be solved for $(C/H, L_x/H)$. Substitution of the result into (8.13) yields (8.14). □

This result can be intuitively understood by recalling the marginal utility pricing rule: As well known,[12] the equilibrium risk premium of a risky asset must be determined as equal to (minus) the covariance between its return and the next instant marginal utility of each agent. Indeed, a close look at (8.9) reveals that:

$$\mu_a - r\iota = -Cov_t \left(dJ_W^k / J_W^k, I_a^{-1} dP\right), \tag{8.15}$$

where $Cov_t(\cdot,\cdot)$ denotes the $(A \times 1)$ vector of the covariance per unit of time conditional on time-t information F_t. In the model of stochastic investment opportunities, the indirect marginal utility J_W^k depends on the wealth and state variables. Given the local linear structure of the model, this produces the Merton-Richard type multi-beta ICAPM. On the other hand, from the first-order condition for optimal consumption, the J_W^k must equal the direct marginal utility, which depends only on the consumption rate under the time-additive von Neumann-Morgenstern utility function. This produces Breeden's single consumption-beta representation of the CCAPM. In contrast, in the present SDU model, the direct marginal utility defined by f_c^k depends on the indirect utility $J^k \left(W^k, x, t\right)$ and the current consumption rate c^k, where W^k can be eliminated from this relation by using first order condition (8.5). Consequently, the direct marginal utility is a function of consumption and state variables. This results in the multi-beta CCAPM of (8.14).[13,14]

[12] See, for example, Lucas (1978) and Cox, Ingersoll, and Ross (1985).

[13] From these discussions it is easy to see that the multi-beta CCAPM given by (8.14) can be obtained even by assuming time-additive von Neumann-Morgenstern utility if instantaneous utility is specified as state-dependent: $u^k = u^k \left(c^k, x\right)$. However, this utility function cannot disentangle preferences toward risk and time.

[14] Given equation (8.15), it is easy to show that in a representative-agent economy equilibrium price P_i must also satisfy the same partial differential equation as derived by Cox, Ingersoll, and Ross (1985) from a separable von Neumann-Morgenstern utility model .

8.3.2 Homothetic Preferences and a Multi-Beta CCAPM in a Representative-Agent Economy

As is shown above, the multi-beta ICAPM does not reduce to the single beta CCAPM under nonseparable recursive utility, reflecting the fact that aggregate consumption does not play a key role as a sufficient statistic in asset pricing without the von Neumann-Morgenstern expected utility. By assuming homogeneous preferences and hence a representative agent economy, however, the multi-beta structure given by (8.14) can be rewritten in a simpler and more tractable form.

Following Duffie and Epstein (1992b), suppose that the preferences are homothetic and the value function J is homogeneous in that it is expressed as

$$J(W, x, t) = j(x, t) \cdot W^{\gamma}, \ \gamma > 0, \tag{8.16}$$

where suffix k is suppressed because of the homogeneous-agent assumption. Then I have $J = WJ_W/\gamma$, implying that the value function solely depends on wealth and the marginal utility. From this relationship and the first order condition (8.5), the marginal utility, J_W, is a function of c and W alone: $J_W = \psi(c, W)$. It follows from the marginal utility pricing rule, (8.15), that any risk premia are determined by covariances of returns with consumption and those with wealth (i.e., with the market portfolio in the present setting). It is straightforward to derive from this observation Duffie and Epstein's (1992b) two-beta CCAPM:[15]

$$\mu_a - r\iota = \lambda_{aC}(\mu_C - r) + \lambda_{aM}(\mu_M - r), \tag{8.17}$$

where the elements of the vectors λ_{aC} and λ_{aM} denote multiple-regression betas on returns on the consumption portfolio and returns on the market portfolio.[16]

The above two-beta structure is obtained under homothetic preference (8.16). In order to weaken this assumption, consider instead a non-homothetic preference given by:

$$J(W, x, t) = j\left(x^1, x^2, t\right) \cdot W^{\gamma} + g\left(x^1, t\right), \ \gamma > 0, \tag{8.18}$$

where $x^i\,(i = 1, 2)$ are $\left(X^i \times 1\right)$ vectors, where $X^1 + X^2 = X$, satisfying $x' = \left(x^{1\prime}, x^{2\prime}\right)$. Then I have $J = W\left(J_W - g\right)/\gamma + g$. Substituting this into the first order condition (8.5), it is seen that indirect marginal utility J_W is a function of c, W, and x^1: $J_W = \phi\left(c, W, x^1\right)$. Finally, substitute this function into the marginal utility pricing rule (8.15) to obtain the following:

Proposition 3: *Consider the representative agent economy with the value function given by (8.18). Then equilibrium risk premia are given by the $\left(X^1 + 2\right)$-beta CCAPM:*

$$\mu_a - r\iota = \chi_{aC}(\mu_C - r) + \chi_{aM}(\mu_M - r) + \chi_{ax^1}(\mu_{x^1} - r\iota),$$

where the elements of vectors χ_{aC}, χ_{aM}, and χ_{ax^1} denote multiple-regression betas on returns on the consumption portfolio, the market portfolio, and the X^1 hedging portfolios.

---

[15] The two-beta CCAPM was first proposed in Mankiw and Shapiro (1986)'s empirical work without any theoretical reasoning. Its theoretical formulation is first given by Epstein and Zin (1989) and Giovannini and Weil (1989) in discrete time models with homogeneous agents.

[16] I can show that the two-beta CCAPM, given by (8.17), entails the absence of preference differentials among investors.

8.4 Stochastic Properties of Optimal Consumption

Let us examine the relationship between the aggregate consumption rate and each individual's consumption rate. Concerning this point, Breeden (1979) proves a positive perfect correlation between the two consumption rates. In the SDU model, however, the same relationship is not valid and what is proven is just the following.

By the multi-fund separation theorem, given in proposition 1, the optimal portfolio can be taken as being constructed from independent $(X+2)$ portfolios, say the consumption portfolio (whose returns duplicate changes in the aggregate consumption C), the X hedging portfolios, and the riskless asset. Thus random changes in W^k are perfectly multiple-correlated with changes in C and x:

$$\mathrm{corr}_t\left\{dW_t^k, (dC_t, dx_t)\right\} = 1. \tag{8.19}$$

On the other hand, from (10), c^k is a function of W^k, x and t: $c^k = c^k\left(W^k, x, t\right)$, so that local changes in c^k are perfectly multiple correlated with $\left(W^k, x\right)$:

$$\mathrm{corr}_t\left\{dc_t^k, \left(dW_t^k, dx_t\right)\right\} = 1. \tag{8.20}$$

Combining (8.19) and (8.20) yields the following:

Proposition 4 *(individual agent's consumption and aggregate consumption): Suppose that capital markets are complete. Then changes in each individual's optimal consumption rate, c^k, are perfectly multiple-correlated with changes in the aggregate consumption rate and returns on the hedging portfolios:*

$$\mathrm{corr}_t\left\{dc^k, (dC, dx)\right\} = 1.$$

As proven by Breeden and Litzenberger (1978), one can construct the optimal sharing rule in complete markets with time-additive von Neumann-Morgenstern utility. Each individual's consumption rate is represented by an increasing function of aggregate consumption alone. This results in perfect correlation between dC and dc^k. Proposition 2 implies that with nonseparable recursive utility, I can no longer construct such a simple sharing rule. Hence, c^k is a function of x as well as of C, so that individual agents' consumption rates are related to the aggregate consumption rate, depending on the state of nature.

8.5 Conclusions

In a model with SDU, I have reexamined intertemporal capital asset pricing and the stochastic properties of optimal consumption-portfolio decisions under a generalized class of preference structures. As in Breeden (1979)'s and Richard (1979)'s articles, the model used is characterized by Brownian information which continuously affects investment opportunities. The results obtained are as follows. First, under any preference structures given by SDU, the ICAPM relationships which Merton (1973) and Richard (1979) have derived, i.e., the equilibrium $X + 2$ fund separation theorem and the multi-beta structure, are still valid.

Secondly, however, the multi-factor structure of risk premia cannot be represented in terms of the aggregate consumption rate alone. That is, Breeden's CCAPM, which is based on the time-additive expected utility representation of preference, does not hold here. Instead, several multi-beta versions of the CCAPM are proposed. If preferences are homogeneous across agents and homothetic, the two-beta CAPM developed by Epstein and Zin (1989) and Giovannini and Weil (1989) obtains: Risk premia are determined by the consumption beta as well as the market beta. Under more generous assumptions betas with respect to the state variables as well as the consumption beta are required in risk premium determination.

Finally, the multiple correlation coefficient of the individual's optimal consumption with the equilibrium aggregate consumption and X state variables is unity.

Bibliography

[1] Arnold, L., 1973, Stochastic differential equations: Theory and applications (John Wiley & Sons, New York).

[2] Bergman, Y.Z., 1985, Time preference and capital asset pricing models, Journal of Financial Economics 14, 145-159.

[3] Breeden, D., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, Journal of Financial Economics 7, 265-296.

[4] Breeden, D., M.R. Gibbons, and R.H. Litzenberger, 1989, Empirical tests of the consumption-oriented CAPM, Journal of Finance 44, 231-262.

[5] Breeden, D. and R.H. Litzenberger, 1978, Prices of state-contingent claims implicit in option prices, Journal of Business 51, 621-651.

[6] Constantinides, G.M., 1990, Habit formation: A resolution of the equity premium puzzle, Journal of Political Economy 98, 519-543.

[7] Cox, J., J. Ingersoll and S. Ross, 1985, An intertemporal general equilibrium model of asset prices, Econometrica 53, 363-384.

[8] Detemple, J.B. and F. Zapatero, 1991, Asset prices in an exchange economy with habit formation, Econometrica 59, 1633-1657.

[9] Duffie, D. and L.G. Epstein, 1992a, Stochastic differential utility, Econometrica 60, 353-394.

[10] Duffie, D. and L.G. Epstein, 1992b, Asset pricing and stochastic differential utility, Review of Financial Studies 5, 411-436.

[11] Duffie, D. and P.L. Lions, 1992, PDE solutions of stochastic differential utility, Journal of Mathematical Economics 21, 577-606.

[12] Duffie, D. and W. Zame, 1989, The consumption-based capital asset pricing model, Econometrica 57, 1279-1297.

[13] Epstein, L.G. and S. Zin, 1989, Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework, Econometrica 57, 937-969.

[14] Epstein, L.G. and S. Zin, 1991, Substitution, risk aversion and the temporal behavior of consumption and asset returns: An empirical analysis, Journal of Political Economy 99, 263-286.

[15] Giovannini, A. and P. Weil, 1989, Risk aversion and intertemporal substitution in capital asset pricing model, NBER working paper no. 2824.

[16] Hansen, L.P. and R. Jagannathan, 1991, Implications of security market data for models of dynamic economies, Journal of Political Economy 99, 225-262.

[17] Ingersoll, J.E., 1992, Optimal consumption and portfolio rules with intertemporally dependent utility of consumption, Journal of Economic Dynamics and Control 16, 681-712.

[18] Kreps, D. and E. Porteus, 1978, Temporal resolution of uncertainty and dynamic choice theory, Econometrica 46, 185-200.

[19] Lintner, J., 1965, The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets, Review of Economics and Statistics 47, 13-37.

[20] Lucas, R.E., 1978, Asset prices in an exchange economy, Econometrica 46, 1429-1445.

[21] Ma, C., 1993, Market equilibrium with heterogeneous recursive-utility-maximizing agents, Economic Theory 3, 243-266.

[22] Mankiw, G. and M. Shapiro, 1986, Risk and return: consumption beta versus market beta, Review of Economics and Statistics 68, 452-459.

[23] Merton, R., 1973, An intertemporal capital asset pricing model, Econometrica 41, 867-880.

[24] Mehra, R. and E. Prescott, 1985, The equity premium: A puzzle, Journal of Monetary Economics 15, 145-161.

[25] Richard, S.F., 1979, A generalized capital asset pricing model,in: E.J. Elton and M.J. Gruber, eds., Portfolio theory, 25 years after: Essays in honor of Harry Markowitz (North-Holland, Amsterdam), 215-232.

[26] Sharpe, W., 1964, Capital asset prices: A theory of market equilibrium under condition of risk, Journal of Finance 19, 425-442.

[27] Sundaresan, 1989, Intertemporally dependent preferences and the volatility of consumption and wealth, Review of Financial Economics 2, 73-89..

[28] Svensson, L.O., 1989, Portfolio choice with non-expected utility in continuous time, Economics Letters 30, 313-317.

Chapter 9

Optimal Consumption and Asset Pricing under Market Incompleteness: A Simple Approach

Abstract: To analyze consumption/portfolio choices and asset pricing under market incompleteness, the consumers' problem is transformed into a reduced space induced by the security markets. The prices of securities with orthonormal payoffs in the reduced space, say pseudo state prices, plays precisely the same key role as state prices in complete markets: the absence of arbitrage is equivalent to the unique existence of a martingale measure constructed from the pseudo state prices; the martingale measure is used to, following the martingale approach, form a single lifetime budget constraint for a static problem equivalent to the original dynamic problem. Implications for the representative-agent asset pricing formula and the APT are discussed. The theory is extended straightforwardly to a multi-period model with short-sale constraints.

JEL Classification Numbers: D52, G10, G11, G12.

Keywords: Incomplete markets, Arrow-Debreu securities, asset pricing, martingale, representative agent, APT, short sale constraints.

9.1 Introduction

The purpose of this paper is to propose a simple approach to the analysis of the optimal consumption/portfolio choice and asset pricing under incomplete security markets. The key proposal is to transform the consumption space into a reduced space induced by marketed-security payoffs, thereby representing consumers' choice problems in the same form as in the case of complete markets. In the resulting canonical representation, basis securities with orthonormal payoffs are defined as pseudo Arrow-Debreu securities and the prices of the securities as pseudo state prices. Pseudo Arrow-Debreu securities can be regarded as a normalized consumption lottery, and the reduced consumption space a composite consumption or lottery space. Arbitrage pricing is characterized in precisely the same manner as in the complete market case. Applying the martingale approach [e.g., Pliska (1986), Cox and Huang (1989, 1991)], this unique pseudo martingale measure is used to form the lifetime budget constraint in a static utility maximization problem equivalent to the original dynamic problem.

In the case of incomplete security markets, the infinitely many sets of the state prices or equivalent martingale measures are consistent with the absence of arbitrage. When, following the martingale approach, the dynamic utility maximization problem is reduced to a static one, infinitely many lifetime budget constraints must be considered corresponding to arbitrage-free equivalent martingale measures. To cope with this difficulty, some "completizing" of models is generally needed. He and Pearson (1991a, b) proposed choosing uniquely a martingale measure, called a minimax martingale measure, by solving a dual problem, thereby extending the feasible consumption space.

However, their sophisticated and well-established theory is faced with some difficulties, which motivate this research. First, to obtain the minimax martingale measure, it is necessary to solve the indirect-utility minimization problem with respect to the arbitrage-free state price vector. Although, once the dual problem is solved, the primary dynamic problem can be easily solved by solving the resultant static problem, the dual problem itself is not always easier to solve compared with the primary problem. In contrast, my approach proposed here reduces the consumption space into a composite consumption space which reflects market incompleteness. Given this transformation, only the arbitrage argument is needed to form uniquely the lifetime budget constraint, as in the case of complete markets.

Secondly, the minimax martingale measure is agent-specific because the underlying implicit state prices are not equalized under market incompleteness. Due to this property, it is prohibitively difficult to consider heterogeneous-agent economies using the duality approach.[1] In the present approach, contrastingly, the pseudo martingale measure is equalized among agents though arbitrage. This facilitates to analyze heterogeneous-agent economies. Specifically, it will be shown that a hypothetical single-agent economy can be constructed such that it produces a no-trade equilibrium supported by the same security price vector as in the equilibrium of the original heterogeneous-agent economy. This is an incomplete-market extension of the representative-agent asset pricing formula developed by Duffie (1996).

Implications for the arbitrage asset pricing (APT) are also discussed. Under market incompleteness, there exist some unavoidable factor risks. Avoiding this problem, an incomplete-market version of the APT will be derived using the pseudo state price vector without the zero-beta and factor portfolios.

The rest of the paper proceeds as follows. In the next section, a simple one-period model of exchange economy is presented. In Section 9.3, I develop a space-reducing approach and reconstruct the arbitrage and equilibrium arguments. Section 9.4 considers implications and applications of the approach. In Section 9.5 are conclusions.

[1] He and Pearson (1991a, b) avoid this difficulty by assuming a single-agent economy.

9.2 The Model

Consider a one-period exchange economy with uncertainty described by a finite set: $\Omega \triangleq \{\omega_1, \cdots, \omega_S\}$ of states, one of which will realize at the end of period. The probability measure is given by Π, under which each state can occur with a nonzero probability. There are N state-contingent securities. Each asset is characterized by its price at the beginning of period, p_n, $n = 1, \cdots, N$, and payoff vector giving its payoffs in each of the S state of nature, $x_n \triangleq (x_n(\omega_s)) \in \mathbf{R}^S$. The entire security markets are thus described by the payoff matrix, $x \in \mathbf{R}^{S \times N}$, and the security price vector, $p \in \mathbf{R}^N$.

Whether security markets (x, p) are complete or incomplete is determined by the relative magnitudes of the rank of payoff matrix x and the number of state, S: They are complete if and only if $\text{rank}(x) = S$, otherwise $[\text{rank}(x) < S]$ they are incomplete. Putting otherwise, letting M denote the marketed (or attainable) payoff space, defined by the span of payoff vectors $x_1, \cdots, x_N$, the security markets are complete if and only if $M = \mathbf{R}^S$ and incomplete otherwise $(M \subset \mathbf{R}^S)$. In order to consider incomplete markets, assume that $\text{rank}(x)(\equiv L) < S$. Setting $K \equiv S - L$ (and hence $1 \leq K \leq S-1$), I call K the deficiency in the security markets.

The economy are populated with I households, ith of whom is characterized by endowment vector $e^i \triangleq (e_0^i, e_1^i) \in \mathbf{R}_+^{S+1}$, where $e_1^i \triangleq (e_1^i(\omega_s)) \in \mathbf{R}_+^S$, and utility functional $U^i = U^i(c^i)$, where $c^i \triangleq (c_0^i, c_1^i) \in \mathbf{R}_{++}^{S+1}$ denotes the consumption vector with $c_1^i \triangleq (c_1^i(\omega_s)) \in \mathbf{R}_{++}^S$. $U^i(c^i)$ is strictly increasing and concave. Letting $\theta^i \in \mathbf{R}^N$ denote the security holding vector, agent i's consumption choice problem is represented as problem (P) below:

$$(\text{P}) \quad \max_{c^i \in C(e^i)} \; U^i(c^i),$$

where $C(e^i)$ represents the budgt feasible consumption set:

$$C(e^i) \triangleq \{c^i \in \mathbf{R}_+^{S+1} \mid \exists \theta^i; e_0^i = p^\top \theta^i + c_0^i \text{ and } e_1^i + x\theta^i = c_1^i\}. \tag{9.1}$$

Given the consumers' choices, the market-clearing conditions for security markets are represented as:

$$\sum_{i \in I} \theta^i = 0. \tag{9.2}$$

From the budget equations in (9.1), the security market equilibrium (9.2) implies $\sum_{i \in I} c_0^i = \sum_{i \in I} e_0^i$ and $\sum_{i \in I} c_1^i = \sum_{i \in I} e_1^i$.

9.3 Transformation

In order to analyze this incomplete market model, consider a linear transformation induced by the following $S \times S$ matrix $A \triangleq (a_{ij}) \triangleq \begin{pmatrix} A_L \\ A_{\setminus L} \end{pmatrix}$, where $A_L \in \mathbf{R}^{L \times S}$ and $A_{\setminus L} \in \mathbf{R}^{K \times S}$:

1. A is orthogonal matrix, i.e., $A^\top A = E$ and hence $A^\top = A^{-1}$;

2. $A_L > 0$;

3. $A_{\setminus L}x = 0$.

Setting $\tilde{x} \triangleq A_L x \in \mathbf{R}^{L\times N}$, the payoff matrix is then transformed by A into:

$$Ax = \begin{pmatrix} \tilde{x} \\ 0 \end{pmatrix}.$$

That is, transformation A reduces the commodity space defined by $\mathbf{R}^S$ into an L-dimensional subspace (Note: rank($\tilde{x}$)$= L$). In what follows, we represent the L dimensional commodity vectors in the reduced space by tildes, e.g., $\tilde{x}_n$.

The idea of the present research is to simplify the incomplete market model using this reduced space representation. To do so, I define below a set of securities which would be Arrow-Debreu securities under market completeness:

Definition 1: A security with payoff vector $\tilde{d}_l \in \mathbf{R}^L$ whose elements are all zero except unity on the i'th spot,

$$\tilde{d}_l \triangleq \left(\overbrace{0, \cdots, 0, \underset{l\text{'th}}{1}, 0, \cdots, 0}^{L} \right)^{\top}, \tag{9.3}$$

is called the l'th *pseudo Arrow-Debreu security*, and its price the l'th *pseudo state price.*

The set of L pseudo Arrow Debreu securities is an orthonormal basis of the reduced space. As I shall show below, under market incompleteness the pseudo state prices play precisely the same key role as state prices in complete markets.

9.3.1 Arbitrage Pricing

To begin with I rewrite the definition of arbitrage in the reduced space.

Lemma: *A portfolio $\theta \in \mathbf{R}^N$ is an arbitrage if and only if*

$$\begin{pmatrix} -p^{\top}\theta \\ \tilde{x}\theta \end{pmatrix} \geq 0 \ \ but \ \neq 0. \tag{9.4}$$

Proof: Recall that arbitrage is defined as a portfolio $\theta \in \mathbf{R}^N$ such that

$$\begin{pmatrix} -p^{\top}\theta \\ x\theta \end{pmatrix} \geq 0 \text{ but } \neq 0.$$

Since it is valid from the definition of A that $x\theta \geq 0 \Leftrightarrow \tilde{x}\theta \geq 0$ and $x\theta = 0 \Leftrightarrow \tilde{x}\theta = 0$, condition (9.4) follows from this definition of arbitrage.□

Therefore, the usual no-arbitrage restriction on asset prices can be duplicated in the reduced space.

Proposition 1: *The security price system is arbitrage-free if and only if there uniquely exists a pseudo state price vector $\phi \in \mathbf{R}^L_{++}$, such that :*

$$p = \tilde{x}^\top \phi. \tag{9.5}$$

Proof: Applying the separating hyperplane theorem to (9.4) yields (9.5). □

To characterize the above arbitrage pricing using a martingale concept, consider a marketed security, indexed by 0, with payoffs of unity in the reduced consumption space:

Definition 2: Security $\tilde{x}_0 \triangleq (1, \cdots, 1) \in \mathbf{R}^L_{++}$ is called the *pseudo riskless security.*

Since the payoff vector of the pseudo riskless security can be duplicated by holding one unit of each pseudo Arrow-Debreu security, the price of the security, p_0, is given in the absence of arbitrage as

$$p_0 = \sum_{k=1}^{L} \phi_k. \tag{9.6}$$

Next, let me define a new probability measure:

Definition 3: $\rho = (\rho_l) \in \mathbf{R}^L_{++}$, where $\rho_l \triangleq \phi_l / \left(\sum_{k=1}^{L} \phi_k \right) \in (0,1)$ (and hence $\sum_{l=1}^{L} \rho_l = 1$), is called a *pseudo martingale measure.*

Then, I obtain the following corollary from equations (9.5) and (9.6):

Corollary 1: *The security price system is arbitrage-free if and only if there uniquely exists a pseudo martingale measure ρ, such that*

$$p = p_0 \mathcal{E}_\rho \left[\tilde{x} \right], \tag{9.7}$$

where $E\rho$ represents expectations with respect to ρ, e.g., $\mathcal{E}_\rho \left[\tilde{x}_n \right] \triangleq \tilde{x}_n^\top \rho$.

Contrary to the standard state pricing, in the pseudo state pricing proposed above, the arbitrage-free price functional or (pseudo) martingale measure is unique and observable even under market incompleteness since the pseudo Arrow Debreu securities are necessarily marketed. This would facilitate applying the martingale approach to incomplete market models.

9.3.2 Consumers' Choice Problems

Next, I transform consumers' choice problem (P) into the reduced space. Letting $\tilde{c}_1^i$ and $\hat{c}_1^i$ denote: $\tilde{c}_1^i \triangleq A_L c_1^i \in \mathbf{R}^L_{++}$ and $\hat{c}_1^i \triangleq A_{\setminus L} c_1^i \in \mathbf{R}^K$ (and hence $A c_1^i \triangleq \begin{pmatrix} \tilde{c}_1^i \\ \hat{c}_1^i \end{pmatrix}$), respectively, the utility functional is rewritten as

$$U^i \left(c_0^i, c_1^i \right) = U^i \left(c_0^i, A^{-1} A c_1^i \right) = U^i \left(c_0^i, A_L^\top \tilde{c}_1^i + A_{\setminus L}^\top \hat{c}_1^i \right). \tag{9.8}$$

As for the budget constraint, the second equation in (9.1) is rewritten by multiplying A from the left as

$$\begin{pmatrix} \tilde{c}_1^i \\ \hat{c}_1^i \end{pmatrix} = \begin{pmatrix} \tilde{e}_1^i \\ \hat{e}_1^i \end{pmatrix} + \begin{pmatrix} \tilde{x}\theta^i \\ 0 \end{pmatrix}, \tag{9.9}$$

where $\tilde{e}_1^i = A_L e_1^i$ and $\hat{e}^i = A_{\setminus L} e_1^i$.

Eliminate $\hat{c}_1^i$ from (9.8) by substituting the second equation of (9.9). Then, given $\hat{e}^i$, the resultant utility functional depends only on $\left(c_0^i, \tilde{c}_1^i\right) \in \mathbf{R}_{++}^{L+1}$. Denote this reduced utility functional by $\tilde{U}^i\left(c_0^i, \tilde{c}_1^i\right)$, i.e.,[2]

$$\tilde{U}^i\left(c_0^i, \tilde{c}_1^i\right) \triangleq U^i\left(c_0^i, A_L^\top \tilde{c}_1^i + A_{\setminus L}^\top \hat{e}_1^i\right),$$

where, if U^i is differentiable, the derivatives of $\tilde{U}^i\left(c_0^i, \tilde{c}_1^i\right)$ are given by

$$\tilde{U}_0^i = U_0^i \text{ and } \tilde{U}_1^i = A_L U_1^i, \tag{9.10}$$

with $\tilde{U}_0^i = \partial\tilde{U}^i/\partial c_0^i$ and $\tilde{U}_1^i = \partial\tilde{U}^i/\partial\tilde{c}_1^i$, etc.

Given this transformation, consider the following problem, (P$'$):

$$(\mathrm{P}') \quad \max_{(c_0, \tilde{c}_1) \in \tilde{C}(\tilde{e}^i)} \tilde{U}^i\left(c_0^i, \tilde{c}_1^i\right),$$

where $\tilde{C}\left(\tilde{e}^i\right)$ represents the set of budget-feasible consumption in the reduced space:

$$\tilde{C}\left(\tilde{e}^i\right) \triangleq \left\{\tilde{c}^i \in \mathbf{R}_+^{L+1} \mid \exists\theta^i; e_0^i = p^\top\theta^i + c_0^i \text{ and } \tilde{e}_1^i + \tilde{x}\theta^i = \tilde{c}_1^i\right\}. \tag{9.11}$$

Then, it is easy to prove the following proposition.

Proposition 2: *Consumers' choice problem (P) is equivalent to problem (P$'$) in the following sense: (i) If* $(c_0^i, c_1^i, \theta^i) = (c_0^{i*}, c_1^{i*}, \theta^{i*})$ *is a solution to (P), then* $(c_0^i, \tilde{c}_1^i, \theta^i) = (c_0^{i*}, A_L c_1^{i*}, \theta^{i*})$ *is a solution to (P$'$), with* $\tilde{U}^i(c_0^{i*}, A_L c_1^{i*}; \hat{e}_1^i) = U^i(c_0^{i*}, c_1^{i*})$*; Conversely, (ii) if* $(c_0^i, \tilde{c}_1^i, \theta^i) = (c_0^{i*}, \tilde{c}_1^{i*}, \theta^{i*})$ *is a solution to (P$'$), then* $(c_0^i, c_1^i, \theta^i) = (c_0^{i*}, A_L^\top \tilde{c}_1^{i*} + A_{\setminus L}^\top \hat{e}_1, \theta^{i*})$ *is a solution to (P).*

Proof: (i) is self-evident. To prove (ii), suppose that $(c_0^i, \tilde{c}_1^i, \theta^i) = (c_0^{i*}, \tilde{c}_1^{i*}, \theta^{i*})$ *is a solution to (P$'$)*. By the definitions of $\tilde{c}_1^i$ and A, the corresponding optimal consumption c_1^{i*} is given by: $c_1^{i*} = A^{-1}\begin{pmatrix} \tilde{c}_1^{i*} \\ \hat{e}_1^i \end{pmatrix} = A^\top \begin{pmatrix} \tilde{c}_1^{i*} \\ \hat{e}_1^i \end{pmatrix}$, which implies (ii). □

Note that problem (P$'$) has the same structure as in complete market settings. Indeed, as in Pliska (1986) and Cox and Huang (1989, 1991), dynamic problem (P$'$) can be transformed into a static problem faced with a *single* lifetime budget constraint.

Corollary 2: *Consumer problem (P$'$) and hence (P) have the following martingale representation, (P$''$):*

$$(P'') \quad \max_{(c_0^i, \tilde{c}_1^i,) \in \tilde{B}(\tilde{e}^i)} \tilde{U}^i\left(c_0^i, \tilde{c}_1^i\right),$$

[2] Although $\tilde{U}^i\left(c_0^i, \tilde{c}_1^i\right)$ depends on $\hat{e}^i$, it is safely abbreviated from the notation because I will not consider any endowment shocks throughout this paper.

where $\tilde{B}(\tilde{e}^i)$ represents the lifetime-budget-feasible set,

$$\tilde{B}(\tilde{e}^i) \triangleq \left\{\tilde{c}^i \in \mathbf{R}_+^{L+1} \mid c_0^i - e_0^i + p_0 \mathcal{E}_\rho\left[\tilde{c}_1^i - \tilde{e}_1^i\right] = 0\right\}. \tag{9.12}$$

Proof: Noting that $\tilde{x}$ has full rank L, assume without loss of generality that $\tilde{x}$ represents the payoff matrix for the set of L pseudo Arrow Debreu securities: $\tilde{x} = E_L$, the L-dimensional unit matrix. Solving the first inequaliy in (9.11) for θ^i and substituting the result into the second inequality yields

$$c_0^i - e_0^i + \phi^\top\left(\tilde{c}_1^i - \tilde{e}_1^i\right) = 0,$$

which implies $\tilde{B}(\tilde{e}^i) = \tilde{C}(\tilde{e}^i)$ in equation (9.11). □

In general, analyzing incomplete market equilibrium entails some 'completization' of the model. For this purpose, He and Pearson (1991a, b) choose uniquely the state price vector by solving a dual problem, thereby *extending* the feasible consumption space. In applying this approach, a dual indirect-utility minimization problem, which is as difficult as the primary one, must be solved to determine the direction in which to extend the consumption space. In contrast, the present approach *reduces* the consumption space into a composite consumption space which reflects the deficiency of some security markets. In this way of 'completizing' the problem, only required procedure is coordinate transformation and there is no need to solve any associate problem. This facilitates the analysis of incomplete market models, as shown in the next section.[3]

9.4 Discussions

9.4.1 Comparison with Other Martingale Measures

Based on propositions 1 and 2, the familiar martingale measure concepts can be related to pseudo state prices. Letting $Ker(A_L)$ represent

$$Ker(A_L) \triangleq \left\{\nu \in \mathbf{R}^S; A_L\nu = 0\right\},$$

and 1_S denote $(1, \cdots, 1)^\top \in \mathbf{R}^S$, equivalent martingale measures and minimax martingale measures can be characterized in this context as follows:

Proposition 3: *Suppose that the equilibrium pseudo state price vector of a incomplete security market economy,* $(\{e^i\}_{i\in I}, \{U^i\}_{i\in I}, x)$ *is given by* ϕ *and an associate optimal solution to (P′) by* $\{c_0^*, \tilde{c}_0^{i*}\}_{i\in I}$*. Then,*

1. *equivalent martingale measure q is given by*

$$q = \frac{\psi}{\psi^\top 1_S},$$

where $\psi = A_L^\top\phi + \nu$, $\nu \in Ker(A_L)$; and

[3] Milne (1981, 1988) proposed an approach in which the utility function over consumption is transformed as one defined over portfolios. Our method is a variant of his induced preference approach in that consumers' preference is transformed as reflecting the security market.

2. *minimax martingale measure* q^i_{minimax} *for agent* i *is given by*

$$q^i_{\text{minimax}} = \frac{\psi^{i*}}{\psi^{i*\top} 1_S},$$

where $\psi^{i*} = U^i_1\left(c^{i*}_0, A_L^\top \tilde{c}^{i*}_0 + A_{\setminus L}^\top \tilde{e}^i_1\right) / U^i_0\left(c^{i*}_0, A_L^\top \tilde{c}^{i*}_0 + A_{\setminus L}^\top \tilde{e}^i_1\right)$.

Proof: See Appendix A.1.□

Note from (9.10) that $\tilde{U}^i_1/\tilde{U}^i_0 = A_L U^i_1/U^i_0$ where $\text{rank}A_L$ is L. Thus, while $\tilde{U}^i_1/\tilde{U}^i_0$ are equalized among agents by pseudo state prices ϕ_l, U^i_1/U^i_0 and hence equilibrium state price vector ψ^{i*} are generically agent-specific. This makes it difficult to apply the minimax martingale method to the case of heterogeneous agent economies. The next subsection shows that the space-reducing approach proposed here is useful to avoid this problem.

9.4.2 Constructing a Representative-Agent Economy

From the fundamental theorem of welfare economics, it is well known that in a complete market economy with heterogeneous agents a hypothetical single agent economy can be constructed such that the resulting no-trade equilibrium produces the same competitive equilibrium prices as in the original economy [see Duffie (1996)]. Using the approach presented above, I can extend this proposition to the present incomplete market setting. The key point is that the competitive equilibrium of the security market economy considered here is constrained-Pareto optimum [see, e.g., Milne (1981, 1988)].

Given endowment vectors $\left\{e^i\right\}_{i\in I}$, let $\lambda \in \mathbf{R}^I_+$ denote an aggregation weight vector to define a (transformed) "aggregate" utility, $\tilde{U}^\lambda\left(c^A_0, \tilde{c}^A_1\right)$, as:

$$\tilde{U}^\lambda\left(c^A_0, \tilde{c}^A_1\right) \triangleq \sup_{\left\{c^i_0, \tilde{c}^i_1\right\}_{i\in I}} \sum_{i\in I} \lambda_i \tilde{U}^i\left(c^i_0, \tilde{c}^i_1\right) \text{ subject to } c^A_0 \geq \sum_{i\in I} c^i_0 \text{ and } \tilde{c}^A \geq \sum_{i\in I} \tilde{c}^i_1. \tag{9.13}$$

Then, I can prove the incomplete-market version of the above proposition as follows:

Proposition 4: *Suppose that* $\left(\left\{c^{i*}\right\}_{i\in I}, \left\{\theta^{i*}\right\}_{i\in I}, p^*\right)$ *is an equilibrium of an incomplete security market economy* $\left(\left\{e^i\right\}_{i\in I}, \left\{U^i\right\}_{i\in I}, x\right)$. *Then, (i) with some aggregation weight vector* $\lambda \in \mathbf{R}^I_+$, *we can construct a single-agent economy* $\left(\left(e^A_0, \tilde{e}^A_1\right), \tilde{U}^\lambda\left(c^A_0, \tilde{c}^A_1\right), \tilde{x}\right)$, *where* $\left(e^A_0, \tilde{e}^A_1\right) = \left(\sum_{i\in I} e^i_0, \sum_{i\in I} \tilde{e}^i_i\right)$, *such that the no-trade equilibrium* $\left(\tilde{c}^A, \theta^A, p\right) = \left(\tilde{e}^A, 0, p^*\right)$ *is an equilibrium; Furthermore, (ii) the optimal utility* $\tilde{U}^\lambda\left(e^A_0, \tilde{e}^A_1\right)$ *is related to the equilibrium consumption allocation* $\left\{c^{i*}\right\}_{i\in I}$ *in the original competitive economy as:*

$$\tilde{U}^\lambda\left(e^A_0, \tilde{e}^A_1\right) = \sum_{i\in I} \lambda_i \tilde{U}^1\left(c^{i*}_0, \tilde{c}^{i*}_1\right), \tag{9.14}$$

where $\tilde{c}^{i*} \equiv A_L c^{i*}$.

Proof. See Appendix A.2. □

Given the proposition, a version of the representative-agent asset pricing formula developed by Duffie (1996) can be proven in this incomplete market model:

Corollary 3: *Suppose $\tilde{U}^{\lambda}$ is differentiable at $\left(e_0^A, \tilde{e}_1^A\right)$. Then, the equilibrium asset prices are given by:*

$$p^* = \tilde{x}^{\top} \tilde{U}_1^{\lambda}\left(e_0^A, \tilde{e}_1^A\right) / \tilde{U}_0^{\lambda}\left(e_0^A, \tilde{e}_1^A\right), \tag{9.15}$$

where $\tilde{U}_1^{\lambda} \triangleq \partial \tilde{U}^{\lambda} / \partial \tilde{c}_1$ and $\tilde{U}_0^{\lambda} \triangleq \partial \tilde{U}^{\lambda} / \partial c_0$.

Proof. See Appendix A.3. □

9.4.3 The APT without Zero-Beta and Factor Portfolios

To show the usefulness of the simple approach proposed above, let me use the pricing rule obtained above to reformulate Ross' (1976, 1977) APT within the incomplete market model without zero-beta and factor-mimicking portfolios. In this case, there exist some unavoidable factor risks and, as a result, factor risk premia are neither unique nor observable.[4] Avoiding this difficulty, an incomplete-market version of the APT can be derived using the price functional constructed from pseudo state prices, as we shall show now.

I follow Dybvig and Ross (1988) in considering a special case without idiosyncratic risks.[5] Assume that returns on marketed securities, $r_n \equiv x_n/p_n - 1_S \in \mathbf{R}^S$, are generated by a linear G-factor model: for $n = 1, \cdots, N$,

$$r_n = \mathcal{E}\left[r_n\right] 1_S + \beta_{n1} h_1 + \cdots + \beta_{nG} h_G, \tag{9.16}$$

where $\mathcal{E}\left[\cdot\right]$ represents expectations with respect to Π; $h_k \in \mathbf{R}^S$ $(k = 1, \cdots, G)$ denote vectors of return-generating factors common to all the securities; and β_{nk} are factor loadings.

Gross return $1_S + r_n = (1 + \mathcal{E}\left[r_n\right]) 1_S + \beta_{n1} h_1 + \cdots + \beta_{nG} h_G$ is necessarily marketed whereas each component of this expression is not normally marketed in the presence of security market deficiency.[6] For example, the state-independent component, $(1 + \mathcal{E}\left[r_n\right]) 1_S$, is not attainable unless the riskless asset is marketed. Similarly, other individual components cannot be attained in the absence of factor portfolios.

To avoid this difficulty, let us represent the return vectors in the reduced form by multiplying transformation matrix A_L from the left as

$$\tilde{1}_S + \tilde{r}_n = (1 + \mathcal{E}\left[r_n\right]) \tilde{1}_S + \beta_{n1} \tilde{h}_1 + \cdots + \beta_{nG} \tilde{h}_G, \tag{9.17}$$

where tildes ($\tilde{\cdot}$) denote transformed vectors, e.g., $\tilde{1}_S = A_L 1_S$ and $\tilde{r}_n = A_L r_n$. Note from proposition 1 that, in the absence of arbitrage, it is valid that

[4] For the nonuniqueness of factor risk premia in the presence of unavoidable factor risks, see Ingersoll (1987, Chap.7).

[5] The discussion presented below can be easily extended to models with asset idiosyncrasy if the idiosyncratic risks are assumed approximately diversifiable as in Ross (1976, 1977).

[6] Note that $y \in \mathrm{span}(x)$ and $z \in \mathrm{span}(x)$ imply $y + z \in \mathrm{span}(x)$ whereas the inverse is not always true.

$\phi^\top\left(\tilde{1}_S+\tilde{r}_n\right)=\phi^\top\tilde{x}_n/p_n=1$ for any marketed securities n. Applying equation (9.17) to this relationship yields

$$\left(1+\mathcal{E}\left[r_n\right]\right)\phi^\top\tilde{1}_S+\beta_{n1}\phi^\top\tilde{h}_1+\cdots+\beta_{nG}\phi^\top\tilde{h}_G=1.$$

Rearranging this equation produces the incomplete market version of the APT as follows:

Proposition 5: *Suppose that security returns conform to (9.16). Then, in the absence of arbitrage, expected returns on the marketed securities are given by*

$$\mathcal{E}\left[r_n\right]=\chi_0+\chi_1\beta_{n1}+\cdots+\chi_G\beta_{nG},\ n=1,\cdots,N, \tag{9.18}$$

where χ_0 and χ_k $(k=1,\cdots,G)$ are uniquely given by $\chi_0=\left(1/\phi^\top\tilde{1}_S\right)-1$ and $\chi_k=-\phi^\top\tilde{h}_k/\phi^\top\tilde{1}_S$, respectively.

9.4.4 Short-Sale Constraints

The space-reducing method presented in the previous section can be applied to the case of short sale constraints:

$$\theta^i\geq 0\ \ \forall i\in I. \tag{9.19}$$

Indeed, the familiar supermartingale property under short sale constriants can be reproduced using psuedo concepts:

Proposition 6: *Under short-sale constraints (9.19), (i) the security price system is arbitrage-free if and only if there exists a pseudo state prices $\phi\in\mathbf{R}^S_{++}$, such that*

$$p\geq\tilde{x}^\top\phi, \tag{9.20}$$

and (ii) if the utility functionals are differentiable, the equilibrium security prices must satisfy

$$p^*\geq\tilde{x}^\top\tilde{U}_1^i\left(c_0^{i*},\tilde{c}_1^{i*}\right)/\tilde{U}_0^i\left(c_0^{i*},\tilde{c}_1^{i*}\right)\forall i\in I. \tag{9.21}$$

Proof: See Appendix A.4.□

9.5 Conclusions

In this paper, I have proposed a simple approach to the incomplete market analysis. The key idea is to transform the consumption/portfolio choice problem into a reduced consumption space reflecting the security market deficiency. The resultant canonical representation allows me to analyze asset pricing in precisely the same manner as in the case of complete markets. Securities with orthnormal payoffs in the reduced space (pseudo Arrow Debreu securities) are used to characterize the incomplete market equilibrium. The absence of arbitrage is equivalent to the unique existence of a martingale measure constructed from the pseudo state prices. The martingale measure is used to, following the martingale approach, form a single lifetime budget constraint for a static problem equivalent to the original dynamic problem. Based on this, the incomplete market versions of the representative-agent asset pricing formula and the APT are derived.

The analysis is, however, limited especially in the two following points. First, the model is restricted to the finite dimensional case. The logical next step is to extend this approach to infinite dimension case such as a Brownian filtration model treated by He and Pearson (1991b). Secondly, the representative-agnet asset pricing formula obtained above crucially depends on the one-commodity setting. When there are more than two commodities, it is well known that the incomplete market equilibrium is usually not constrained-Pareto optimal [see Cass (1992) and Duffie (1987)]. In contrast to the complete market case, this would make it difficult to reproduce the same formula under multi-commodity settings.

9.6 Appendix for Chapter 9

9.6.1 Appendix A.1: Proof of Proposition 3

To prove part 1 of proposition 3, note that state price vector ψ satisfies $p = x^\top \psi$ in the absence of arbitrage. Since $x = A_L^\top \tilde{x}$, this and equation (9.4) imply that

$$A_L \psi = \phi,$$

from which ψ can be obtained as a particular solution $A_L^\top \left(A_L A_L^\top\right)^{-1} \phi = A_L^\top \phi$ plus the general solution to $A_L \nu = 0$. Regarding part 2 of the proposition, from the definition of minimax martingale measures, the associate state price vector ψ^{i*} equals the vector of intertemporal marginal rate under the attained consumption allocation.□

9.6.2 Appendix A.2: Proof of Proposition 4

Proposition 3 can be proven using the same way as used by Duffie (1996): Suppose that $(\{c^{i*}\}_{i\in I}, \{\theta^{i*}\}_{i\in I}, p^*)$ is an equilibrium of an incomplete security market economy $(\{e^i\}_{i\in I}, \{U^i\}_{i\in I}, x)$. Let ϕ^* represents the pseudo state price vector under this equilibrium and $\bar{c}^{i*}$ the corresponding optimal composite consumption. Then, from proposition 2 and the saddle point theorem, there exists a Lagrange multiplier $\alpha_i \geq 0$ such that $\left(c_0^{i*}, \bar{c}^{i*}\right)$ solves a problem.

$$\sup_{c_0^i, \bar{c}_1^i} \tilde{U}^i \left(c_0^i, \bar{c}_1^i\right) + \alpha_i \left\{e_0^i - c_0^i + \phi^{*\top} \left(\bar{e}_1^i - \bar{c}_1^i\right)\right\}.$$

The positivity of α_i follows from the fact that $\tilde{U}^i$ is increasing in c_0^i.

To construct a utility functional for a single agent, set $\lambda_i = 1/\alpha_i$ as agent weights. The resulting aggregate utility $\sum_{i\in I} \lambda_i \tilde{U}^i \left(c_0^i, \bar{c}_1^i\right)$ indeed attains maximum at $\left(c_0^i, \bar{c}_1^i\right) = \left(c_0^{i*}, \bar{c}_1^{i*}\right)$, as follows:

$$\begin{aligned}
\sum_{i\in I} \lambda_i \tilde{U}^i \left(c_0^{i*}, \bar{c}_1^{i*}\right) &= \sum_{i\in I} \left[\lambda_i \tilde{U}^i \left(c_0^{i*}, \bar{c}_1^{i*}\right) + \lambda_i \alpha_i \left\{e_0^i - c_0^{i*} + \phi^{*\top} \left(\bar{e}_1^i - \bar{c}_1^{i*}\right)\right\}\right] \\
&\geq \sum_{i\in I} \left[\lambda_i \tilde{U}^i \left(c_0^i, \bar{c}_1^i\right) + \lambda_i \alpha_i \left\{e_0^i - c_0^i + \phi^{*\top} \left(\bar{e}_1^i - \bar{c}_1^i\right)\right\}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i\in I} \lambda_i \tilde{U}^i\left(c_0^i, \tilde{c}_1^i\right) + e_0^i - c_0^i + \phi^{*\top}\left(\tilde{e}_1^i - \tilde{c}_1^i\right) \\
&\geq \sum_{i\in I} \lambda_i \tilde{U}^i\left(c_0^i, \tilde{c}_1^i\right).
\end{aligned}
$$

Finally, let me prove that, given the aggregate utility constructed above and aggregate endowment vector $\left(e_0^A, \tilde{e}_1^A\right)$, $\tilde{U}^\lambda\left(c_0^A, \tilde{c}_1^A\right)$ attains maximum at $\left(c_0^A, \tilde{c}_1^A\right) = \left(e_0^A, \tilde{e}_1^A\right)$. Suppose that $\tilde{U}^\lambda\left(c_0^A, \tilde{c}_1^A\right)$ attains maximum at $\left(c_0^A, \tilde{c}_1^A\right) = \left(x_0^A, \tilde{x}_1^A\right) \neq \left(e_0^A, \tilde{e}_1^A\right)$, so that $\tilde{U}^\lambda\left(x_0^A, \tilde{x}_1^A\right) > \tilde{U}^\lambda\left(e_0^A, \tilde{e}_1^A\right)$. Then, from equation (9.13), there exists an underlying allocation $\left(\left\{x_0^i\right\}_{i\in I}, \left\{\tilde{x}_1^i\right\}_{i\in I}\right)$, where $\sum_{i\in I} x_0^i = x_0^A$ and $\sum_{i\in I} \tilde{x}_1^i = \tilde{x}_1^A$, such that

$$
\begin{aligned}
& \sum_{i\in I} \lambda_i\left[\tilde{U}^i\left(x_0^i, \tilde{x}_1^i\right) + \alpha_i\left\{e_0^i - x_0^i + \phi^{*\top}\left(\tilde{e}_1^i - \tilde{x}_1^i\right)\right\}\right] \\
> & \sum_{i\in I} \lambda_i\left[\tilde{U}^i\left(c_0^{i*}, \tilde{c}_1^{i*}\right) + \alpha_i\left\{e_0^i - c_0^{i*} + \phi^{*\top}\left(\tilde{e}_1^i - \tilde{c}_1^{i*}\right)\right\}\right],
\end{aligned}
$$

which contradicts the optimality of $\left(c_0^{i*}, \tilde{c}_1^{i*}\right)$. □

9.6.3 Appendix A.3: Proof of Corollary 3

From corollary 2, the consumption choice problem for the single agent constructed following proposition 4 can be reduced to:

$$
\sup_{\left(c_0^{A*}, \tilde{c}_1^A\right)} \tilde{U}^\lambda\left(c_0^A, \tilde{c}_1^A\right) + \alpha\left\{e_0^A - c_0^A + \phi^\top\left(\tilde{e}_1^A - \tilde{c}_1^A\right)\right\},
$$

where α represents a Lagrange multiplier. Suppose that $\tilde{U}^\lambda$ is differentiable. Then, from the first order condition and proposition 4, it is valid that

$$
\tilde{U}_1^\lambda\left(e_0^A, \tilde{e}_1^A\right) / \tilde{U}_0^\lambda\left(e_0^A, \tilde{e}_1^A\right) = \phi,
$$

which implies (9.15).□

9.6.4 Appendix A.4: Proof of Proposition 6

(i) To prove the "if" part, suppose that a price system p is arbitrage-free. Then, setting M as $M \triangleq \left\{\begin{pmatrix} -p^\top\theta \\ \tilde{x}\theta \end{pmatrix} \text{ where } \theta \in \mathbf{R}_+^N\right\}$, I have $M \cap \mathbf{R}_{++}^{L+1} = \{0\}$. The separating hyperplane theorem implies that there exists a linear functional, $F(y) = \varphi^\top y$, $\varphi = (\varphi_l)_{l=0,\cdots,L} \in \mathbf{R}_{++}^{L+1}$, such that

$$
F(y_M) \leq 0 \leq F(y), \forall y_M \in M, y \in \mathbf{R}_{++}^{L+1}.
$$

The first inequality can be rewritten as $-p^\top\theta + \phi^\top\tilde{x}\theta \leq 0$, where $\phi = (\varphi_l/\varphi_0)_{l=1,\cdots,L} \in \mathbf{R}_{++}^L$. From $\theta \geq 0$, this implies equation (9.20).

To prove the "only if" part, suppose that equation (9.20) holds valid. Consider a feasible portfolio $\theta \in \mathbf{R}_+^N$ such that $\tilde{x}\theta \geq 0$. I prove that this cannot be an arbitrage. First, it must hold valid that $-p^\top\theta \leq 0$ because $p^\top\theta \geq \phi^\top\tilde{x}\theta \geq 0$, where the first inequality comes from the assumption and the second from the positivity of θ. Next, suppose that this portfolio satisfies $p^\top \theta = 0$. Then, by assumption, I have $\phi^\top \tilde{x}\theta \leq 0$. This implies that either $\tilde{x}\theta = 0$ or at least one component of $\tilde{x}\theta$ is strictly negative. It follows that θ cannot be an arbitrage.

(ii) From the Kuhn-Tucker theorem, the first order conditions are given by

$$\theta\left(-p^* \tilde{U}_0^i + \tilde{x}^\top \tilde{U}_1^i\right) = 0; \; -p^* \tilde{U}_0^i + \tilde{x}^\top \tilde{U}_1^i \leq 0; \text{ and } \theta \geq 0,$$

which imply inequality (9.21).□

Bibliography

[1] Cass, D., 1992, Incomplete financial markets and indeterminacy of competitive equilibrium, in J. Laffont, ed.,Advances in Economic theory: Sixth world congress Vol.II (Cambridge University Press, New York, NY).

[2] Cox, J.C. and C.-F. Huang, 1989, Optimal consumption and portfolio policies when asset prices follow a diffusion process, Journal of Economic Theory 49, 33-83.

[3] Cox, J.C. and C.-F. Huang, 1991, A variational problem arising in financial economics, Journal of Mathematical Economics 20, 465-487.

[4] Duffie, D, 1987, Stochastic equilibria with incomplete financial markets, Journal of Economic Theory 41, 405-416.

[5] Duffie, D., 1996, Dynamic asset pricing theory 2nd edition (Princeton University Press, Princeton, New Jersey).

[6] Dybvig, P.H. and S. Ross, 1988, Arbitrage, The new Palgrave dictionary of economics.

[7] He, H. and N.D. Pearson, 1991a, Consumption and portfolio policies with incomplete markets and short-sale constraints: The finite-dimensional case, Mathematical Finance 1, 1-10.

[8] He, H. and N.D. Pearson, 1991b, Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case, Journal of Economic Theory 54, 259-304.

[9] Ingersoll, J.E.,Jr., 1987, Theory of financial decision making (Rowman and Littlefield, Totowa, New Jersey).

[10] Milne, F., 1981, Induced preferences and the theory of the consumers, Journal of Economic Theory 24, 205-217.

[11] Milne, F., 1988, Arbitrage and diversification in a general equilibrium asset economy, Econometrica 56, 815-840.

[12] Pliska, S.R., 1986, A stochastic calculus model of continuous trading: Optimal portfolios, Mathematics of Operations Research 11, 371-382.

[13] Ross, S., 1976, The arbitrage theory of capital asset pricing, Journal of Economic Theory 13, 341-360.

[14] Ross, S., 1977, Return, risk, and arbitrage, in I. Friend and J.L. Bicksler, eds., Risk and return in finance (Ballinger Publishing Co., Cambridge, Mass), 189-218.

[15] Ross, S., 1978, A simple approach to the valuation of risky streams, Journal of Business 51, 453-475.