博士論文

Applications of Stochastic Processes to Macroeconomics
（マクロ経済学への確率過程の導入）

猪瀬 淳也

APPLICATIONS OF STOCHASTIC PROCESS TO THE MACROECONOMICS

JUNYA INOSE

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Firstly, I owe my deepest gratitude to my adviser Prof. Makoto Nirei for supporting to finalize this dissertation, for his patience and deep knowledge. His insightful comments and suggestions helped me in writing and refining this dissertation. At the same time, I would like to express my gratitude to my ex-adviser, Prof. Hiroshi Yoshikawa for his guidance and instilling in me the study of macroeconomics, and various indispensable suggestions for theoretical analysis. I was greatly influenced by his broad interests and enthusiasm, and his zeal and generosity will affect all my future lifetime.

Besides my adviser and ex-adviser, I would like to thank the rest of my thesis committee: Prof. Hiroshi Ohashi, Prof. Shin-ichi Fukuda and Prof. Tsutomu Watanabe (in alphabetical order of the first name), for their insightful comments which deeply helped me to widen my research from various viewpoints.

I would also like to offer my sincere thanks to Prof. Kazuya Kamiya in Kobe University and all RIETI (The Research Institute of Economy, Trade and Industry) working group member, Prof. Takashi Unayama, Prof. Koichi Ando, Prof. Ryoji Hiraguchi, Prof. Masayuki Keida for their advice and fruitful discussion for my research. And finally, last but by no means least, I would like express appreciation to my wife Yuriko who was always my support in every moment.

Part I

INTRODUCTION

1 INTRODUCTION

1.1 MODERN MACROECONOMICS AND MICROFOUNDATION

Since the Lucas critique, it is widely believed by most of economists that macroeconomic theory must be based on microfoundations. Modern macroeconomics from Friedman, Phelps, Lucas to real business cycle theory (Kydland and Prescott [31]) is considered to be based on the principles of microeconomic theory. The optimization of rational economic agents such as households and firms play key role in the macroeconomic theory, and this concept strongly affects recent analytic framework as the real business cycle theory (RBC), endogenous growth theory, and dynamic stochastic general equilibrium theory (DSGE). The basic framework to analyze macroeconomic optimization is, in general, to analyze the optimization of rational economic agents who maximize the utility (in case of households) or profit (in case of firms) over an infinite horizon under rational expectation and calculation of dynamic optimization of, e.g., consumers, investors or managers.

However, the recent microfoundation does not set careful treatment on the aggregation. One good example would be the Dixit and Stiglitz [11]. In Dixit and Stiglitz, the representative agent is assumed with its functional type of the utility functions in a CES-type. The major reason that the functional type of the utility set as a CES type is the analytic convenience. Here one question arises; is there any reasonability for these assumptions? Also, is it really reasonable to assume the elasticity of substitution as constant across all goods? If the macroeconomics literally pursue the “microfoundation”, not only the introduction of the microeconomic concept into the macroeconomics, but also dealing the aggregation of microeconomic variables carefully are both important.

1.2 REVIEW OF RELATED LITERATURES

1.2.1 *Demand Aggregation*

In this subsection, we briefly review models of our concern. To begin with, a literature which describes the demand aggregation should be considered with high concern. In the demand theory, the uniqueness and stability of Walrasian equilibrium under standard dynamic adjustment processes is examined and discussed for a long time. As a corollary of the analysis, the Walrasian equilibrium is found to be unique if there is a unique solution for the equation: excess demand function equals to zero.

Sonnenschein [47] posed a question on the restrictions on the excess demand function under utility-maximizing rational agents, and the theorem, named as Debreu-Sonnenschein-Mantel Theorem (Sonnenschein [47], Debreu [7] and Mantel [34]), provided the negative answer. The theorem implies that an economy has an arbitrary number of equilibrium in general with arbitrary stability properties. MWG (Mas-Colell, Whinston, and Green [35]) described this property as "anything goes" in general equilibrium theory. Therefore, special assumptions on the preference, such as Cobb-Douglas or CES, is required to obtain comparative statics results and empirical contents.

Also, Debreu-Sonnenschein-Mantel Theorem establishes that the Law of Demand —when the goods price falls, the demand for the goods increase — does not apply even at the level of a single market. And yet Neoclassical macroeconomists consider it is valid to start with a model in which the entire economy is a single "representative agent". Because of the theorem, it soon turned out that it was not possible to extend the law for single individuals to the Law of Demand to apply on the market level without unrealistic assumptions; to assume all individuals have the same preferences.

Among the sequence of the progress following the result of Sonnenshein, a research on the law of demand aggregation of sufficiently diverse individual demands may be one of the significant turning points, which was pioneered by Hildenbrand [16]. After his work, evidence and further application of the "Law of Demand" has been considered both in the theoretical and empirical aspects (see, for example, Hildenbrand and Jerison [18], Hardle, Hildenbrand, and Jerison [15], and Hildenbrand and Kneip [19]).

One of the standard model for aggregating consumer demand in the literature of the second line of research is known as μ model (see, for example Hildenbrand [17]). First define household i 's demand function as $f^i(p, \omega^i)$ where p is a price vector for l goods, and ω^i is a disposable income of the household i. To define a micro economic model of a large and heterogeneous population of household, one has first to specify the space $\mathscr{F}$ of admissible demand function. The space of household characteristics is defined by the Cartesian product

$$R_+ \times \mathscr{F}$$

Every household is described by a point (ω^i, f^i) in the space $R_+ \times \mathscr{F}$ with the distribution μ. In terms of distribution μ the market demand is defined by,

$$F(p) \equiv \int_{R_+ \times \mathscr{F}} f(p, x)\, d\mu$$

In the μ model, the heterogeneous households are divided according to its income and preference structure ($f^i(p, \omega^i)$). Then, the admissible demand functions in $\mathscr{F}$ are assumed to be parametrized by a parameter α in some set $\mathscr{A}$. Thus, instead of $f \in \mathscr{F}$, the demand function is assumed to be written as f^α with $\alpha \in \mathscr{A}$. As a standard assumptions for the μ model, the following 4 assumptions are assumed: (i) μ us a probability measure on the σ- field of Borelian subsets of $R_+ \times \mathscr{A}$, (ii) there exists mean income and the mean income is finite, (iii) f^α are continuous in (α, p, x) and continuously differentiable in p and x, (iv) average Sultsky substitution matrix is negative semi definite. The household attribute to divide heterogeneous households turns more detail in the following papers (see, for example, Hildenbrand [17]) and household attributes, such as age and employment status or household size, are employed for the empirical calculation.

Also, a literature starting from Becker [3] could be another good reference for our work. Becker proved that basic features such as the demand function, which decreases with respect to the price and supply functions which increases with respect to the price can be derived as a consequences of a budget-constrained agent's randomly chosen behavior. The detailed relation with our work is described in the introduction of the chapter 2, however, our work could be considered as a combination of literature of Hildenbrand and Becker.

1.2.2 *Utility Function Representation and the Separability*

In the part 2 of this dissertation, we address the microfoundation of the functional type of the utility function. As already described in the previous subsection, macroeconomists set special assumptions to the preferences, or the utility function, like Cobb-Douglas. However, such assumptions are generally chosen with considering only the analytic convenience and not considering any microfoundations, just as introduced in the 1.1 of this dissertation. We address this problem with considering separability and discrete choice.

To begin with, we set our first focus on the allocation of the total expenditure; the separable problem. In principle, each consumer has to deal with the problem to allocate his/her income between saving and consumption, or purchasing durable and non-durable goods simultaneously. All of the parts of this allocation problem may interact, and therefore the change in future wage may cause the change in current saving plan or the purchase of the durable goods. However, if we allow all interaction at the same time, the allocation problem cannot be solved because of its complexity and therefore the simplification is required, either by aggregation or separation. The separation of the decision-making leads that the ultimate determinants like assets, wage rates, prices, interest rates are related to the total expenditure, but not to each group expenditures directly. An importance for detecting the separability for certain goods group to examine the structure of the utility function was recognized for a long time. It has its origin in Stotz [48], Gorman [14], Goldman and Uzawa [13] and other related works in the same literature, and related developments are also available (see, for example, Deaton and Muellbauer [6]).

Another focus shall be set in the probabilistic consumer choice and the discrete choice, especially a brand of commodity purchases. The major reason to set the focus is simple; after allocating the expenditure, consumer have only to consider the brand to be purchased, and most of goods in the real world is literally discrete and probabilistic. The probabilistic consumer theory, or the probabilistic choice system, describes the observable distribution of demands by a population of consumers, and assume the hypothesis of random preference maximization which postulates that the distribution of demands in a population is the results of individual preference maximization (see McFadden [37] for details). The discrete choice models with differentiated products and the characteristics approach (see for example Lancaster [32]) are described by Anderson, Palma, and Thisse [2] and it was revealed that these literatures have some common parts both in technical and conceptual aspects.

Recent developments treat these two different literatures simultaneously, like considering the separability under the discrete choice (see, for example, Smith, Abdoolakhan, and Taplin [46]).

1.2.3 *Firm Value and Retained Earnings*

In the part 3 of this dissertation, we address the problem of the corporate finance, with setting our focus on the retained earnings. The paper which closely relates to our research is Bolton, Chen, and Wang [4]. The major reason we spare the last part of this dissertation will be mentioned in 1.3 and the conclusion of this dissertation.

To begin with, a reference on the Modigliani and Miller [41] should be set at first. As is well known, i) in the perfect market without tax, the corporate value is irrelevant to the capital structure and ii) with tax, levered firm increases its corporate value due to tax deduction of the debt cost. Since the Modigliani and Miller [41], a sizable literatures has investigated to understand firms' financing policies. Early standard model of investment under uncertainty assume the frictionless capital credit markets. This assumption allows firms for smooth financing (e.g., funding or borrowing) from the market and cash reserves does not affect the firm's value. These models faced questions by many empirical studies and this leads to a literature treating uncertainties or frictions. A literature to treat uncertainty reveals that if a firm faces liquidity risk, the firm accumulates large cash balances to avoid bankruptcy. Another, but related literature treating frictions devotes main efforts to understand the effects of frictions, such as corporate taxes, credit (for example, Hugonnier, Malamud, and Morellec [22]) and equity market frictions, etc.

Another related and important literature to our work is the dividend policy, sometimes mentioned as the "dividend puzzle". According to Matos [36],

... although there is no consensus in the marketplace on the need and importance of payout policies, most managers and some academics believe the policies affect the value of firms. Based on the empirical studies, the answer to the dividend puzzle — namely, to understand why firms insist on paying dividends if they are supposed to be irrelevant to the value of the firm — seems to be that the payment of dividends has a natural market among the inframarginal investors who can make some tax-based arbitrage profit.

In the part 3 of this dissertation, we address both the financial structure problem and the dividend puzzle by introducing retained earnings, which relates closely with dividend, explicitly into current dynamic capital structure model.

The major reason to focus the retained earnings is current corporate financing structure in macro level. Despite many related literatures, not many researches set its focus on the retained earnings which attract interests in the global, especially in advanced economies. For instance, OECD [43] pointed out the importance to consider the retained earnings as follows;

In advanced economies between 1995 and 2010, it is estimated that on average 66% of corporate investments were financed by shareholder capital in the form of retained earnings.

Although the retained earnings occupy majority of the actual corporate finance, the current corporate financial theory does not provide any theoretical framework to explain the phenomenon. In the part 3 of this dissertation we address this issue by modifying the dynamic capital structure model. The retained earnings are usually defined by a cumulative corporate profit that are not paid out to the shareholders as a dividends or share. The inclusion of the retained earnings to the currently established model enables us to describe the dynamics and long-term firm's growth policy by introducing dividend rate and other related variables.

1.3 contents

Based on this consciousness, this dissertation is composed of part 1, 2 (demand aggregation) and part 3 (stochastic process in the corporate finance). In the part 1 of this dissertation, we address the robustness of the assumption of the "representative agent". In most of the recent macroeconomics model, the assumption of the "representative agent" is frequently used. Despite the frequency of usages of the assumption, the robustness or the validity of the assumption is not examined with the microfoundation. The part 1 addresses this issue by formulating consumer's utility in probability density function. This expression enables us to describe the consumer's attribute (mainly the budget allocation to each goods group like food, leisure, telecommunications, etc.).

In the part 2 of this dissertation, we examine the microfoundation of the utility function (especially CES utility function) within certain goods group, especially goods group whose choice can be described in the discrete choose model. Formulating the results of the part 1 and 2 to describe the way to describe macroeconomic demand with the exact microfoundation.

On the contrary to the part 1 and 2, part 3 addresses the issue on the corporate finance. Before moving onto the detailed explanation, let me describe the background for focusing the corporate finance. As is described in the first part of this chapter, the maximization problem in standard macroeconomic model is regarding households and firms. As the the part 1 and 2 mainly treated the maximization problem for households, part 3 addresses maximization for firms. Also, at the last half of the part 2, we also addressed the firm's profit maximization problem under the monopolistic competition and stochastic process. So basically, the description of the maximization process both for households and firms are provided in part 1 and 2. However, in the standard macroeconomic model, we need another important part, financing. There are several financing paths in the macroeconomics. One major path is to set households also as an investor, and also there are several literatures to include banks into the model. No matter which financial paths the model takes, the financing process must be included to the model so that the model is to be closed. In the financing process, the aggregation is also calculated with the "analytic convenience" and therefore we address the financing process in the part 3 of this dissertation.

As described in above, this dissertation covers essential part in basic macroeconomic models; utility description and maximization of households (part 1 and first half of part 2), profit maximization in monopolistic competition (last half of part 2) and financing process (part 3) in stochastic process. Although such model does not introduce monetary or expectation so far, it is possible to introduce such function technically.

Part II

REPRESENTATIVE AGENT AND STOCHASTIC PROCESS

2 REPRESENTATIVE AGENT IN A FORM OF PROBABILITY DISTRIBUTION

2.1 SUMMARY

This chapter provides methodology for the demand aggregation with rigorous micro foundation by using probability distribution function. In the model setting, a set of agents defined by similar type of attributes is approximated as one representative agent with probability error. By introducing several assumptions on the probability process, a restriction condition to assume single representative agent as a whole economy is provided for the following 3 cases as: i) normal diffusion with Markov process, ii) normal diffusion with non-Markov process, and iii) anomalous diffusion. In case of the normal diffusion, no restriction condition for the shape of the probability density function of the consumption bundle is required because the Gaussian distribution is a stable distribution (i.e., a linear combination of two independent random variables with Gaussian distribution has the same distribution). On the other hand, in case of the anomalous diffusion, some restricting condition on the parameter of the probability process is required to maintain such stability.

2.2 INTRODUCTION

A struggle on describing heterogeneity and its demand aggregation based on rigorous micro foundation has a long history. The standard theory on demand aggregation has developed until the early 80s (as a reference, see for example, Deaton and Muellbauer [6], Houthakker and Taylor [21], etc.). The standard aggregation theory sets its focus on the market transaction under privately consumption (e.g., no public goods, no externalities), and provide several approaches to discuss aggregation.

The first line of research could be characterized for considering the condition to describe the aggregated behavior of a group as a single decision maker. These progresses exerted influence on the theoretical development of RBC theory.

A second line of research could be characterized as a struggle on finding structural features of the aggregated demand. The famous original problem has been initially raised by Sonnenschein [47], Mantel [34] and Debreu [7], and now known as the Debreu-Sonnenschein-Mantel Theorem. They raised a problem on whether the structure of the individual demand described by the utility maximization is preserved by demand aggregation. The relation between the Debreu-Sonnenschein-Mantel Theorem and descending the "Law of Demand" by Hildenbrand [16] is as mentioned in 1.2 Review of Related Literatures.

Third line of research is starting from Becker [3]. Becker pointed out that "households may be irrational and yet markets quite rational". This means that we should not attribute any observed irrationality of market participants to market, nor attribute any rationality of markets to the participants, the individuals. Becker derived such result with considering households who distributes along its budget constraint. As long as opportunities were initially restricted to the budget line, the average consumption of many households would be close to the mid-point of the budget constraint, with different households uniformly distributed around the mid-point.

In this chapter, we extend traditional and recent works on demand aggregation under various household attributes by formulating consumer's utility in probability density function. Our methodology especially highlighted the consumer heterogeneity in the aspect of the "deep parameter" for the purpose of smooth aggregation. Here it should be noted that, in the literature of Hildenbrand, a demand function of each household is a starting point, and sophisticated technical conditions on preferences or utility functions is abandoned. However, in this chapter, we both employ the preference, or utility functions, and the core concept of μ model, to divide household's purchasing based on their attributes. In the μ model, the households are assumed to distribute in their characteristic space. We introduce the concept of the household distribution into each household characteristic set, i.e., we firstly divide all household into several groups by its characteristics and analyze the distribution along its budget constraint. The major reason to divide households is that if we set all household into one group, the distribution of household along the budget constraint may yield multi-peak distribution. However, by dividing many households into several groups with respect to its characteristics, it might be possible to consider the single-peak distribution which can be described in a simple equation.

The purpose of this chapter is to provide redefinition of the restriction condition for the representative agent frequently assumed in RBC model. In RBC model, the competitive equilibrium of the market economy is achieved under a resource constraint with maximizing representative household's expected utility. Although this assumption was path-breaking for the development of macro economics, this also generated many critics which mainly focusing on i) abstracting heterogeneity of firms and households and ii) constraints for all agents to act optimally in all markets and at all times (see, for example, Kirman [28]). Following these critics, some research papers with counterarguments are also published (see, for example, Kiyotaki [29]). These discussions are, as a matter of cause, based on the market completeness. However, if we introduce some distortion or constraint as an assumption, such completeness may break up and the foundation of the representative agent collapses. To address this matter, here we provide a restricting condition for the approximation of representative agent without assuming the market completeness. The contribution of this paper in this field is to provide the restriction condition for the representative agent with a stochastic process model. As we set our major purpose to the description of the representative agent, position of the research literature is rather different from that of Hildenbrand [16] or Becker [3] because such literature sets its focus on the analysis of the "Low of Demand" or macro-rationality and micro-irrationality. Lastly, it should also be noted that what we provide is not a complete model to calculate general equilibrium under dynamic optimization, but one tool which may be used to calculate the utility function, or the "deep parameter", of consumers. Therefore, we do not set any focus on the profit maximization of firms, market clearing conditions, etc. in this chapter, but confine our interest only into the demand aggregation to provide better approximated description.

The reminder of the chapter is organized as follows: Section 2 firstly examines issues in construction problem setup for grasping micro-foundation based aggregated demands. Thereafter propose a key concept for introducing methodology of probability distribution into the consumer choice problem. Section 3 introduces several patterns for describing dynamics and provides conditions to maintain assumption of the representative agent. Section 4 provides technical support for conducting an empirical test. Section 5 concludes.

2.3 MICRO FOUNDATION AND PROBLEM SETUP

2.3.1 *Definition of the Micro Foundation*

The concept of Pareto optimality is rigorously defined in microeconomics (see, for example, Mas-Colell, Whinston, and Green [35]). Consider an economy consisting of consumers ($i = 1, ..., I$) and goods ($l = 1, ..., L$). Consumer i's preferences over consumption bundles $x_i = (x_{1i}, ..., x_{Li})$ in his consumption set $X_i \subset R^L$ are represented by the preference relation. Also, define the price vector of each good as $\boldsymbol{p} = (p_1, ..., p_L)$ and initial endowment as ω.

Hereafter we set our focus on the amount of utility along the budget constraint line (dashed line in Figure 2.1). The utility becomes maximum at the point of Walras equilibrium ($\mathbf{x}^*$), and decays as the point goes farther from equilibrium point, as long as the utility function is strictly concave (Figure 2.1 above). In this setup, the amount of utility along the budget constraint line could be drawn intuitively as written in Figure 2.1 below. Here it should be emphasized that the coordinate of the Figure 2.1 below is $\boldsymbol{x_b}$ which represents the budget constraint line in 2 goods economy with satisfying $\boldsymbol{p} \cdot \boldsymbol{x_b} = \omega$.

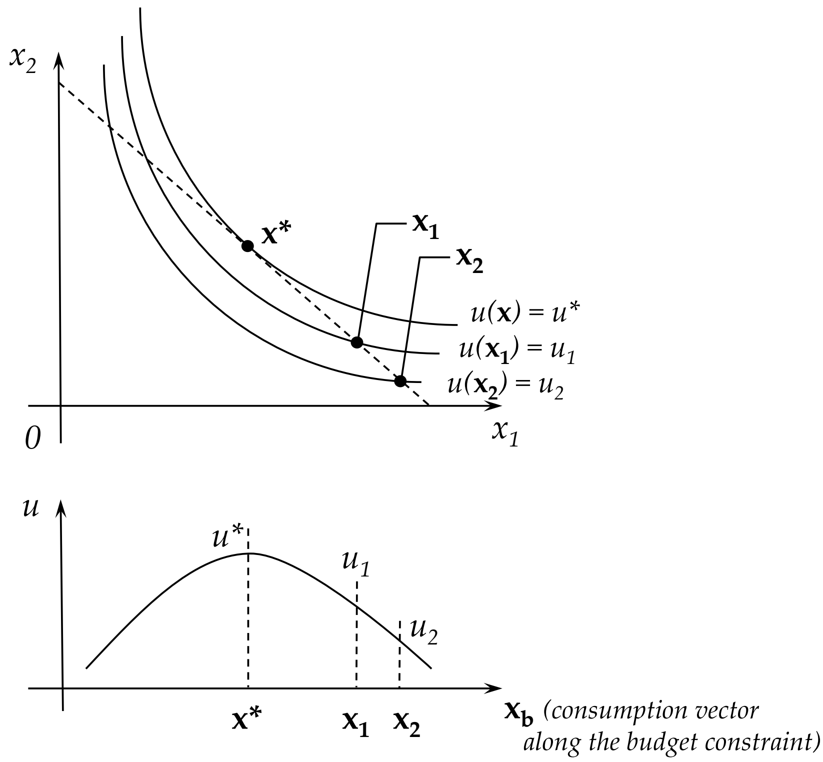

Figure 2.1: Utility along the Budget Constraint Plane

The similar discussion could be extended to the multi-goods (e.g., assume L goods in the economy) situation. In the L goods situation, the budget constraint $\boldsymbol{p} \cdot \boldsymbol{x_b} = \omega$ provides a hyper plane $\boldsymbol{x_b}$ within L-1 dimensional phase space, and the utility along the budget constraint could be described as a function in a form that $u : R_+^{L-1} \to R$.

According to the standard μ model, the observable household characteristics is parametrized by μ. Define the set of the households who have a characteristics μ as $I(\mu)$ and also define the number of households who have characteristics μ as $N(\mu)$, and consider consumption bundle of these households. If we employ the amount of income as one of the characteristics of the households, the distribution of the consumption bundle who has μ characteristics is allocated along the budget line (dashed line in Figure 2.2). Also, in the standard μ model, the utility function (or parametrizing factor of the utility function α) can change independently from other household attribute. This leads that the shape of utility function can differ each other even if the amount of income is the same. In general, observable cloud (i.e., distribution of households along the budget constraint) of the consumption bundle with μ character has some kind of distribution written in Figure 2.2, and each consumption bundle of household i, $i \in I(\mu)$ maximizes each utility function $u_i(x)$.

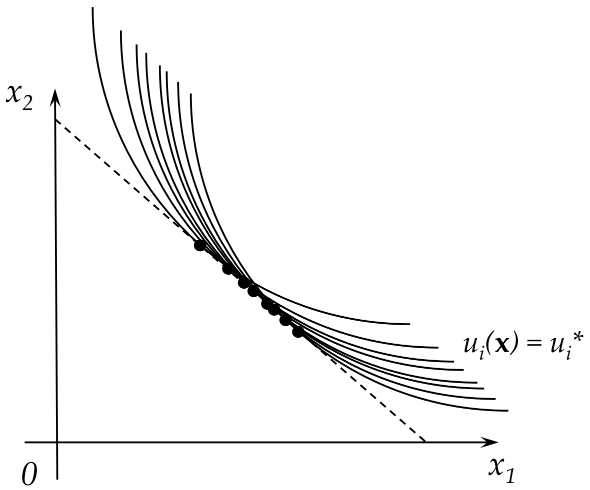

Figure 2.2: Observable Cloud of Consumption Bundles under the Same Household Characteristics

2.3.2 *Describing Probability Density Function of the Observable Cloud of Consumption Bundles*

One standard way to analyze the distribution of the observable clouds on budget constraint plane is to introduce the concept of probability distribution into the system. Such approach is frequently used in the field of a discrete choice as the Random Utility Model. In the standard analysis of the Random Utility, the observer usually constrains its focus to some concrete discrete event like the choice of the transport options such as car and train. Here we introduce its basic concept to the observable cloud of consumption bundles with dividing all households with respect to its characteristics. If all of the distributions of the cloud could be well approximated as some function among some set of the households with similar characteristics μ, the whole demand within nation could also be well approximated by integrating these attributes. In other word, let us consider a system that a household with characteristics μ draw a lottery to determine the budget allocation on budget constraint plane, which obeys a distribution $P_{\mu_i}(\mathbf{x_b})$. The major purpose of our work is to describe the dynamics of $P_{\mu_i}(\mathbf{x_b})$ for every consumer characteristic, or heterogeneity. Here it should be noted that the characteristics μ have to include the initial endowment ω so that all households with the same characteristics are allocated on the same budget constraint plane. Then we consider the probability process of the consumption bundle $\mathbf{x_b} = \{\mathbf{x_b}(t)\}_{t>0}$ within the probability space $\{\Omega, \mathscr{F}, \mathbb{P}\}$ under price taking situation. Let $\mathbf{x_b}$ be an Ito process driven by the standard Wiener process $\mathbf{R_t}$ and described by the stochastic differential equation in R_+^{L-1} as $d\mathbf{x_b}(t) = a(\mathbf{x_b}(t)) \cdot dt + b(\sigma) \cdot d\mathbf{R_t}$.

Definition.
(1) All households are divided into N types as $\{\mu_i\}_{i \leq N}$ according to its attributes.
(2) The budget constraint for the household type μ_i is defined as $\mathbf{p} \cdot \mathbf{x_b} = \omega_i$.
(3) The probability density function of the observable clouds for the household type μ_i is described as $P_{\mu_i}(\mathbf{x_b})$.

Assumption 1. *$^{\exists}N$ s.t. $for^{\forall}i$, $P_{\mu_i}(\mathbf{x_b})$ is a continuous function and has a single peak in $L-1$ dimensional hyper plane*

Assumption 2. *$^{\exists}N$ s.t. $for^{\forall}i$, all consumption bundles within the household type μ_i has an ergodic property, i.e.,*

$$\frac{1}{T}\sum_{t=1}^{T}\mathbf{x_b}(t) \xrightarrow{p} E\left[\mathbf{x_b}\right],\ T \to \infty$$

where $E\left[\mathbf{x_b}\right]$ is a convergence value of the spatial average as sample number $n \to \infty$.

The economic meaning of these 2 assumptions are to assume representativeness for a group of similar household attributes as an approximation. If we divide adequate number of consumers by its attributes, the assumptions allows us to describe one representative, or typical, consumer for each segment of consumers with good approximation. For example, we may be able to assume a typical consumption bundle (with error) for the household whose head is 20 years old man, single, living in urban area, or the household whose head is 50 years old man, with wife and 2 children, living in suburb, etc.

In addition, the *Assumption 2* could be understood as the most essential assumption for this paper. To assume ergodic property is equal to abandon treating fixed effects for each household, and use a probability density function alternatively. This assumption, or approximation, may arise a critique for losing strictness for the aggregation. However, let us show our several counterarguments in advance. First of all, our fundamental motivation is to provide new methodology to approximate a wide variety of households, and at the same time, to provide theoretical methodology to evaluate the validity of the approximation empirically. Therefore, if the adequate amount of data is available, we can examine the validity of this approximation empirically. Secondly, even under the assumption of the ergodic property, we allow not only normal diffusion process but so-called anomalous diffusion in the following sections. This expansion may take the theoretical framework closer to the real economy.

Meanwhile, let us also discuss the advantage for setting the *Assumption 2* at the same time. The core advantage to employ this assumption is that we could expect asymptotic property for estimating probability density function. In case of rigorous micro-foundation, the functional shape of the utility function are assumed with mainly focusing on the mathematical convenience to solve equations. However, this assumption does not necessarily approximate real economy well. On the other hand, our methodology enables us to approximate actual utility function itself for each household attribute so long as the previous assumptions could be well achieved as an approximation. Moreover, we can expect our fitted probability function, or inversely calculated representative utility function, asymptotically equals to the actual utility function if we could obtain plenty amount of micro data from consumer survey, or any other micro panel data.

The intuitive understanding for this approximation is described in the Figure 2.3. Most theories with rigorous micro foundation assume that all of the household's utility function can be different and also can move with errors. However, every consumption bundle is settled in the maximizing point of each utility function and no noise is allowed for achieving utility maximizing condition. On the other hand, the employment of the *Assumption 1* and *Assumption 2* enables us to set a representative utility function for the selected households with similar attributes. Instead of describing household's heterogeneity by the difference of its utility function, we allow the existence of error from stationary point which corresponds to the utility maximizing point for the households.

Now let us compare these two models. The first model (Figure 2.3. above) corresponds to a model without any approximation and holds as long as we stand on the micro-foundations. On the other hand, the second model (Figure 2.3. middle) with single, or representative utility function could be regarded as a model with certain approximation of the first model. Here let us recall the model of Becker [3]. He derived that, as long as opportunities were initially restricted to the budget line, the average consumption of many households would be close to the mid-point of the budget constraint, with different households uniformly distributed around the mid-point. In our model, we assume one representative utility function (u_R) for each household group, and assume all households within certain group distribute around the Walras equilibrium point of the representative utility function.

2.3.2.1 *Describing Dynamics of Observable Clouds*

To describe the dynamics of the observable clouds, let us consider 3 types of the stochastic process; (1) Markov process, (2) Non-Markov

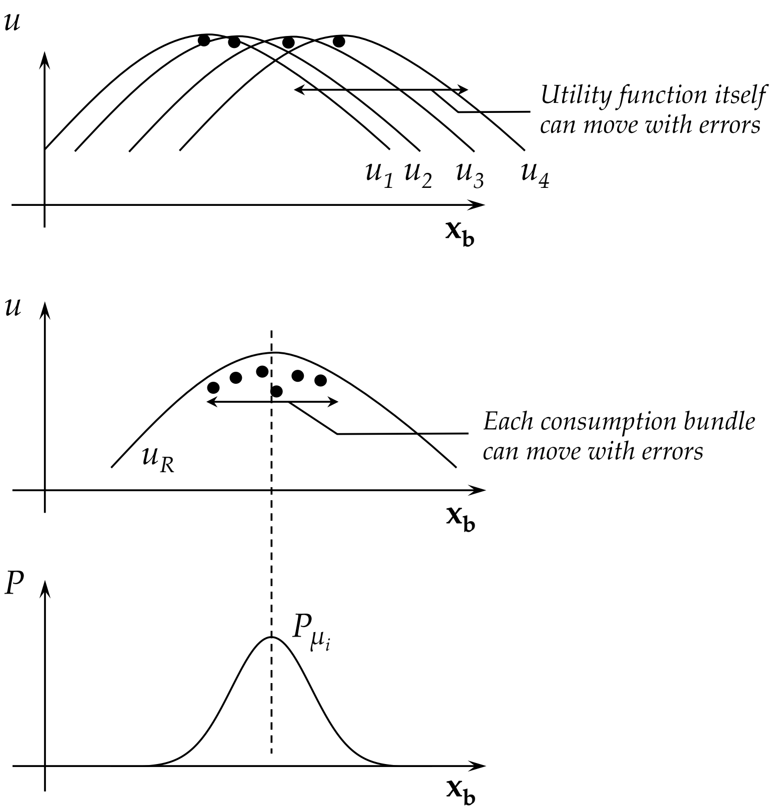

Figure 2.3: Aggregation of Heterogeneous Utility Functions and Probability Density Function of the Consumption Bundles

process and (3) Anomalous diffusion. It would require less explanation for the consideration of the first one, Markov process as we already set such stochastic process for the dynamics of $\mathbf{x_b}(t)$. Then, the next question is that whether the assumption of the Markov process is enough for describing economic dynamics of the observable clouds. In general, a economy without any friction can be well described in the standard Markov process. On the other hand, an effect of friction plays important role in economics because in some case, the friction, or slack, plays key role for the description of the economics which pace down a speed of relaxation after drastic change of external environment (e.g., labor market slack which make wage growth slow, etc.). Although such example is commonly used in the literature of the labor market or search process, it might be worth to consider similar characteristics to our problem setup. Therefore, we introduce the (2) Non-Markov process for the reference.

Lastly, the introduction of the (3) Anomalous diffusion is worth for consideration. Firstly, let us explain the definition of the Anomalous diffusion. As is described in the Appendix, the anomalous diffusion shows the non-linear growth of the mean square displacement with time. In the anomalous diffusion case, it is known that the relation between the mean square displacement and time changes as $<(\Delta x)^2> \propto 2D\Delta t^{\alpha}$. Here in the standard diffusion process, the mean square displacement is expected to show the linear growth with respect to time, i.e., $\alpha = 1$, and process with $\alpha \neq 1$ is defined as anomalous diffusion. Now let us bring our focus back onto the economics. In sub-diffusive case ($0 < \alpha < 1$), the consumption bundle diffuse slowly, and therefore households with similar consumption bundles tend to stay still close after several periods. On the other hand, in enhanced diffusion case ($\alpha > 1$), the diffusion of the consumption bundles are very quick, and households with similar consumption bundles may stay far away even after some short periods. We do not evaluate actual economic activity in this paper because of the lack of adequate data set of consumption bundles, however, it is meaningful to leave a methodology which can evaluate the actual dynamics of consumption bundles.

The standard Markov process can be described in the following Fokker-Planck equation, which describes the dynamics of $P_{\mu_i}(\mathbf{x_b})$. However, in case of the non-Markov process, it is difficult to describe in the general form. Therefore, we introduce the Kramers' equation which enables us to describe the dynamics of $P_{\mu_i}(\mathbf{x_b}, \mathbf{v_b})$ with considering the "memory" effect. The detailed explanation and derivation of these equations are left in Appendix A. Lastly, the anomalous diffusion is defined by fractional Fokker-Planck equation and for details, see Metzler [39].

Fokker-Planck equation:

$$\frac{\partial}{\partial t}P_{\mu_i} = \left(-\sum_{j=1}^{L} \frac{\partial}{\partial x_j}\alpha_{1j}(\mathbf{x_b}) + \frac{1}{2}\sum_{j=1}^{L}\sum_{k=1}^{L} \frac{\partial^2}{\partial x_j \partial x_k}\alpha_{2jk}(\mathbf{x_b}) \right) P_{\mu_i}(x_1, x_2, ..., x_L, t) \tag{2.1}$$

Kramers equation:

$$\frac{\partial}{\partial t}P_{\mu_i}(\mathbf{x_b}, \mathbf{v_b}, t) = \left[-\sum_{j=1}^{L-1} \frac{\partial}{\partial x_{bj}} v_{bj} + \sum_{j=1}^{L-1} \frac{\partial}{\partial v_{bj}} \left(-\frac{\partial}{\partial x_{bj}} u_R(x) + \gamma v_{bj} \right) + D\sum_{j=1}^{L-1} \frac{\partial^2}{\partial v_{bj}^2} \right] P_{\mu_i}(\mathbf{x_b}, \mathbf{v_b}, t) \tag{2.2}$$

Now let us move our focus back to our assumptions. First and foremost, in the derivation of the Fokker-Planck equation, we need to assume the process to be Markov (see appendix for the details) and therefore the effect of friction or any other term which violate a Markov process assumption is ignored.

The Kramers equation is mainly composed of two parts of effects: i) an effect from external force and friction, and ii) an effect from diffusion. In our economic model, the first part represents a effect of representative utility function (u_R) for each household group. Next, the second part represents the effect to diffuse from initial consumption bundle. This effect could be seen when there is a blind purchase, unexpected consumption, etc. In the field of behavioral economics, topics like emotional effects in economic decisions or non-rational decision making are established, and the ideal optimization process is abandoned in this theory. There are several motivations to abandon the ideal optimization. For example, facts like i) no exact optimizing strategy is known in many real-world situations or, ii) even if an optimizing strategy exists, unrealistic amounts of knowledge on every alternatives or consequences might be required, could be good motives.

The approximation this model employs could also be well explained as to assume i) the fluctuation of the utility function among each agent, and ii) deviation from ideal optimization (e.g., effect of blind purchase or unexpected consumption) as i.i.d., and treat sum of these two effects as single error.

2.4 DEMAND AGGREGATION

2.4.1 *Dynamics of Each Consumption Bundle under Representative Utility Function*

For describing dynamics of the consumption bundles under given potential, or utility function, using a methodology for calculating dynamics of particle under physical potential instead of analyzing dynamics of distribution as a whole plays key role. In this subsection, we firstly revisit standard theory for describing particle dynamics in the field of physics, and secondly provide concept to connect methodology in physics with economic phenomenon. The derivation of the following generalized Langevin equation is written in the Appendix.

generalized Langevin equation:

$$\frac{d}{dt}\mathbf{v_b}(t) = -\int_{-\infty}^{t} \gamma(t-t')\mathbf{v_b}(t')dt' + \nabla u_R(\mathbf{x_b}) + \mathbf{R}(t) \qquad (2.3)$$

This concept is firstly developed in the field of statistical physics (see, for example, Mori [42] and Kubo [30]), and here let us spare several lines to consider economic meaning of this equation. The left hand side of the generalized Langevin equation equals to a rate of acceleration for the movement of the consumption bundles. The first term of the right hand side could be interpreted as an effect of a relaxation against drastic transformation of the exterior environment (e.g., tax revisions, monetary policy change, etc.). In general, most of economic models has a feature to relax gradually to the stationary distribution, and this term governs the speed of relaxation from the aspect of consumption adjustment. Next, the second term represents an effect to confine their consumption bundles around typical household's one. This effect is related to our *Assumption 1*, and it is intuitively rational to assume a typical consumption bundle within limited type of households. The third term represents the noise which yields deviation from typical household's consumption bundle.

2.4.2 *Restriction toward Aggregation*

For the demand aggregation, whether the distribution has a stability or not plays crucial role. If the distribution before aggregation does not have a stability, the error term of the aggregated variable (e.g., aggregated consumption) does not have simple distribution (like Gaussian). However, in the standard macroeconomics, many empirical studies assume the error term to distribute in Gaussian or any other frequently used distribution which generally has stable property. On the other hand, if all of the distribution before aggregation has a similar type of stability, the error term of an aggregated variable has also similar type of distribution, and this leads that the aggregated demand could be described with similar methodology as used so far.

As mentioned above, the condition that all of the distribution before aggregation has a similar type of stability is just a sufficient condition but not a necessary condition for the existence of the representative agent. However, it might be reasonable to assume that every household follows the same nature and therefore the basic distribution is common for every household. Then the shape and structure of the distribution is the matter and we would like to examine the condition for such case in the followings.

2.4.2.1 *Case of Normal Diffusion under Markov Process*

In the case of normal diffusion under Markov process, there are no restriction for the demand aggregation, because the Gaussian process generally has a stability. Therefore, in this model setting, the typical assumption for the representative agent is satisfied.

2.4.2.2 *Case of Non-Markov Process*

If we set focus on the steady state distribution, the difference in its process does not affect its result. However, if we set focus on the relaxation process, its dynamics differs dramatically. The Langevin equation in Markov process (A.47) has a shape of homogeneous differential equation. This corresponds that the relaxation process follows exponential function just as calculated in many economic models. On the other hand, if we introduce generalized memory term in the Langevin equation, its solution no longer follows usual exponential relaxation but shows slow relaxation process because of having previous memories.

2.4.2.3 *Case of Anomalous Diffusion*

In case of the anomalous diffusion, whether the probability density function calculated from the previous Langevin equation (in general, fractional Fokker-Planck equation, see Metzler [39]) shows stability or not is of high concern. Now consider to calculate the distributions of consumption bundles for all of the household attribute μ. If any of the probability density function does not present additive property, the description of the aggregated demand may become complicated.

To consider these issues, an approach to consider the stability in non-trivial Lévy process plays key role because Lévy process includes several probability process which shows anomalous diffusive property (e.g., Lévy flight). A useful diagram for judging stability of a non-trivial Lévy process is known as "Takayasu Diamond" (Takayasu [49], detailed explanation is written in the Appendix), whose basic concept was originally derived by Levy [33] and Khintchine [27]. The definition of the stable distribution is given by Feller [12] as:

Definition 7. *Let X, X_1 , . . . , X_n be independent random variables with a common distribution R. The distribution R is stable if and only if for $Y_n \equiv X_1 + X_2 + \ldots + X_n$ there exist constant c_n and ϵ_n such that*

$$Y_n \stackrel{d}{=} c_n X + \epsilon_n \tag{2.4}$$

where $\stackrel{d}{=}$ indicates that the random variable of both sides have the same distribution.

In general, if we add two or more probability variable which follows similar type of probability distribution, the added variable does not necessarily follow the same probability distribution. The stable distribution is a set of distributions which achieves this property, i.e., the sum of two or more stable probability variable with similar probability distribution follows the same distribution.

2.5 TOWARD AN EMPIRICAL TEST

2.5.1 *Calibration of the Representative Utility Function*

Calculating the representative utility function from the distribution of consumption bundles is difficult in general, however, it becomes possible with adequate amount of the data set. The first and foremost condition to guarantee the validity of this approximation is that the probability density function of consumption bundles during the data acquisition fluctuates only around the error term, and no dynamical change of the probability density function is observed. If this approximation could be regarded as rational, we can use steady state Fokker-Planck equation or Kramers equation and able to calculate representative potential inversely.

2.5.2 *Evaluating an Assumption of Non-Markov Property*

A methodology for an empirical test to verify the validity of the assumption of the representative agent could be provided via these theoretical frameworks. Firstly, the assumption of the "Non-Markov property" of each household could be confirmed by calculating mean square displacement $< X^2(t) >$. According to Vinales and Desposito [50], mean square displacement for times $t \gg \tau$ under generalized Langevin equation with Mittag-Leffler function is calculated as:

$$< X^2(t) > \approx 2\gamma D t^2 E^1_{2-\mu,3}[-(\omega_\mu t)^{2-\mu}] + 2\gamma D \frac{\tau^\mu}{\gamma_\mu} \left\{ 1 - E^1_{2-\mu,1}[-(\omega_\mu t)^{2-\mu}] \right\}. \tag{2.5}$$

where γ_μ and ω_μ are constant.

Therefore, if we could obtain panel data of household's consumption breakdown, we can evaluate the existence of memory term in actual consumption bundles by calculating time variation of the mean square displacement.

2.5.3 *Evaluating an Assumption of Anomalous Diffusion*

Secondly, the assumption of normal diffusion could be evaluated by observing the shape of a probability density function of any attribute of consumers. For instance, an analysis of consumption bundles of 30-40 years old, urban living, married households may be able to reveal the existence of anomalous diffusion. In general, the probability density function under anomalous diffusion presents fat tail (not decay in exponential). The solution of steady state probability density function under harmonic potential (in our model, to assume $u_R(\mathbf{x_b})$ as $u_R(\mathbf{x_b}) \sim -\frac{1}{2}\lambda x^2$) is calculated in Jerison [24] and the result shows asymptotic power law behavior:

$$P^{eq}_{\mu_i}(\mathbf{x_b}) \approx \frac{D\gamma}{\mu\lambda|\mathbf{x_b}|^{1+\mu}} \tag{2.6}$$

where γ denotes friction coefficient and μ is a exponent of a characteristic function ($p(k)$) of the noise variable ($p(k) = exp(-D|k|^\mu)$). Therefore, if we could obtain the probability density function of consumption bundles for any type of consumer attribute, we can confirm the existence of anomalous diffusion by evaluating a tail of the probability density function whether to obey exponential decay or power-law.

2.6 DISCUSSION

In this chapter we established new methodology for approximating huge number of consumers. As is described in the introduction of this chapter, our purpose is to provide a theoretical framework to consider the validity of the assumption of the representative agent. According to our result, if all assumptions hold as a good approximation and the consumption bundles of each household attributes acts like standard Brownian motion, the assumption of the representative agent holds and aggregation process becomes quite simple. In addition, even if the consumption bundles of each household attribute follow non-Markov process, or anomalous diffusion, the assumption of the representative agent still holds within limited parameter space. Then the problem is that whether the actual households act as expected, at least in good approximation, or not. We need empirical analysis to answer this question. It should be emphasized that for the empirical analysis of this work, we need huge sample size to obtain appropriate evaluation. Here we describe two major reasons we need huge sample.

First and foremost, the basic concept to divide households based on their attributes increases the request on the sample size. For example, we decide to divide all households with their attribute like i) the age of a head of the household (e.g., divide age group in 5 criteria), ii) living place (e.g., divide living place in 2 criteria like urban and suburb), iii) income level (divide in 5 criteria), then the households are divided into 50 groups. As we have to check each distribution of the consumption bundles, the sample size of each group should be at least more than several hundreds.

Second reason is the evaluation of the distribution function. As is written in the equation (2.6), we have to analyze the tail distribution of the consumption bundles of each household group when we consider the case of anomalous diffusion. Whether the tail decreases in exponent or power law is one of the major differences, we also require huge sample size for each household group to evaluate the tail dynamics.

As is discussed so far, the empirical analysis on this issue is not simple and straight forward. However, we believe that such empirical analysis can bring fair evaluation of the assumption of the representative agent, and hope that this theoretical framework encourages future applications both in theoretical and empirical works.

Part III

UTILITY FUNCTION AND STOCHASTIC PROCESS

3 LOYALTY AND CONSUMPTION: A CES REPRESENTATION

3.1 SUMMARY

This chapter offers a new interpretation of the elasticity of substitution in the CES utility function under the discrete choice and separability. We model an economy with one discrete choice goods group and one composite good under diverse consumers. The results from our theoretical analysis illustrate the relation between the distribution of loyalty of each good and goods demands. Moreover, the origin of the elasticity of substitution of the CES utility function is described based on our assumptions. According to our results, the elasticity of substitution σ does not change even if the distribution of loyalty differs by each good. On the other hand, coefficients of the demand of good i (X_i^σ) varies according to each good's attractiveness. We also consider the production under this economy and obtain the result as the increase in productivity leads the decrease in price. This effect is the same as the standard Melitz model (Melitz [38]).

3.2 INTRODUCTION

The analysis on consumer behavior has a long history both in empirical and theoretical aspects, and the utility function plays one of the most key roles to describe micro- and macroeconomic consumer behavior. One of the most standard methodology to describe the representative consumer is to assume its utility functions as a CES-type utility function (see for example Dixit and Stiglitz [11]). But here one question arises: is there any justification for these assumptions? Also, is it really reasonable to assume the elasticity of substitution as constant across all goods? In this paper, we explore the answers to these questions by considering the micro-foundation of the CES utility function under the separability and discrete choice. The related literatures are written in the 1.2.2.

In this chapter, we consider the aggregated demand under the separability and the discrete choice with micro-foundation. The reasons for considering discreteness under the separability could be summarized as the following two points. First point is that the introduction of this model enables us to consider the distribution of attractiveness of each good, especially the relation between the diversity of demand of each consumer and the total demand of its good. In the real economy, the loyalty, or attraction consumers feel, for certain goods differs by consumers, and its difference can be described by a distribution. According to our result, the shape of the distribution of consumer's loyalty for each good is essential for describing the demand of goods, and maintaining or increasing its loyalty plays key role for the better profits of firms. As a second point, this model also reveals another interpretation of the elasticity constant σ of the CES utility function with micro foundation, which is generally set as a deep parameter. Under our assumption, the elasticity constant σ has a micro foundation and has a meaning which closely relates to the distribution of loyalties.

In our model, each consumer purchases a product which is most attractive for him/her. In such case, the firm must increase consumers who like their product the best and this may lead to the market strategy, i.e., customer segmentation, promotion strategy, etc. Under such literature, a strategy to provide a product which offers high loyalty only for certain consumer segment, but not for other usual segment may be able to provide better profit for certain market circumstances. In addition, this theoretical framework also enables us to analyze firm's strategies to increase R&D expenditure to achieve product innovation, or disruptive innovation which lead to rapid increase in loyalty and better profit.

The reminder of the paper is organized as follows: Section 2 firstly outline our model by defining micro foundations and introducing several approximations, followed by the application of the Houthakker [20] into our model. Section 3 shows brief numerical calculation results for reference and section 4 concludes.

3.3 THE MODEL

3.3.1 *Definition of the Variables*

Let us construct our economic model. In this paper, we first start from focusing a certain goods group with considering the separability. Let x_{iA} be m dimensional sub-vector of the consumption vector x_i of consumer i so that $x_i = (x_{iA}, x_{i\bar{A}})$. x_{iA} is then said to be strongly *separable* if the utility function takes in the form

$$u_i = f_i\left(u_{iA}(x_{iA}) + u_{i\bar{A}}(x_{i\bar{A}})\right) \tag{3.1}$$

where $u_{iA}(x_{iA})$ is the sub-utility function associated with x_{iA}, and f is some monotone increasing function. The consumer i chooses a good under the consumption vector x_i with the discrete choice, i.e., each consumer purchases one unit of the commodity which offers the greatest utility. In this case, $u_{iA}(x_{iA})$ can be written in the form;

$$u_{iA}(x_{iA}) = max\left\{u_{iA1},\, u_{iA2},\, ...,\, u_{iAm}\right\} \tag{3.2}$$

where $\boldsymbol{u_{iA}} = (u_{iA1}, u_{iA2}, ..., u_{iAm}) \in \mathbb{R}^m$ is a stochastic variable and represents the utility vector of consumer i when the consumer purchases each good under the commodity subgroup.

3.3.2 *Economy with 2 Discrete Goods and a Composite Good*

For the simplicity, let us consider the economy with 2 discrete goods ($A1$ and $A2$) and one composite good N. The effect of pricing can be taken into account by substitution between discrete goods and continuous good in case of the strong separability. p_{A1}, p_{A2} and p_N are prices of good A1, A2 and a composite good, and ω_i is a budget for the consumer i. The decision for purchasing good A1 or A2 by consumer i can be described as follows;

Lemma 1. *Under the strong separability, the consumer i purchases good A1 if and only if;*
(1-1) The utility for purchasing the good A1 is greater than the utility for purchasing a composite good additionally, and
(1-2) The utility for purchasing good A1 is greater than the sum of (a) the utility purchasing good A2 and (b) the utility for purchasing a composite good with reserved money.

Proof. Firstly, the strong separability is generally defined as follows in this economy.

$$u(x_{iA1}, x_{iA2}, x_{iN}) = u(x_{iA1}, x_{iA2}) + u(x_{iN}).$$

Also, as the goods subgroup $x_{iA}$is discrete choice goods group, each consumer basically chooses one good from goods subgroup (A1, A2). However, in this economy, we also allow not to purchase anything from goods subgroup (A1, A2) and use all budget for purchasing the composite good for more generalization.

Also, the utility for purchasing q_{iN} amount of the composite good is described by $u_{iN}(q_{iN})$ for each consumer i. Then, the utility for purchasing (i) good A1, (ii) good A2 and (iii) purchasing nothing from goods subgroup (A1, A2) can be described in the form of indirect utility function as follows;

$$\begin{cases} u_{iA1} + u_{iN}\left({}^{(\omega_i - p_{A1})}/_{p_N}\right) & (i) \\ u_{iA2} + u_{iN}\left({}^{(\omega_i - p_{A2})}/_{p_N}\right) & (ii) \\ u_{iN}\left({}^{\omega_i}/_{p_N}\right) & (iii) \end{cases} \tag{3.3}$$

According to the discrete choice model, the good A1 will be purchased only in the case that the indirect utility in case of (i) is larger than that of (*ii*) and (*iii*), and this leads to the descriptions in the **Lemma 1.**

$$\begin{array}{llll} (1-1) & u_{iA1} & +u_{iN}\left(^{(\omega_i - p_{A1})}/_{p_N}\right) > & u_{iN}\left(^{\omega_i}/_{p_N}\right) \\ (1-2) & u_{iA1} & +u_{iN}\left(^{(\omega_i - p_{A1})}/_{p_N}\right) > & u_{iA2} + u_{iN}\left(^{\omega_i - p_{A2}}/_{p_N}\right) \end{array} \tag{3.4}$$

■

Also, we need a strong assumption regarding the income effect to proceed calculations.

Assumption 1. *The utility function for the composite good $u_{iN}(q_N)$ is same for all consumers in the whole economy, i.e., $u_{iN}(q_{iN}) = u_N(q_{iN})$ for $^\forall i$.*

The Lemma and Assumption lead us to describe the actual demand of good A1. Firstly, we assume the indirect utility of the discrete goods ($\boldsymbol{u_{iA}} = (u_{A1}, u_{A2}) \in \mathbb{R}^2$) acts as stochastic and i.i.d. for each consumer. If we define the distribution of consumers in the 2 dimensional phase space as $\varphi(u_{A1}, u_{A2})$, the consumers allocated in the region $u_1 > u_2$ will choose the good 1 in discrete goods group and *vise versa*. Also, as is described in the **Lemma 1**, the (indirect) utility for purchasing good 1 has to be larger than purchasing composite good with same price. Then the demand of good 1 can be described as the integration of $\varphi(u_{A1}, u_{A2})$ for gray colored region in Figure 3.1. This leads to the expression as;

$$X_1 = \int_0^\infty \rho(\omega) d\omega \int_{\delta_1}^\infty du_{A1} \int_0^{u_{A1}-\delta_{12}} \varphi(u_{A1}, u_{A2}) du_{A2} \tag{3.5}$$

where $\rho(\omega)$ is the income distribution in this economy, and the demand for the good 2 can be described symmetrically. Here we define indirect utility $\delta_m(\omega, p_{Am}, p_N)$ and $\delta_{mn}(\omega, p_{Am}, p_{An}, p_N)$ as follows to describe equations simply. Here it should be noted that δ_{mn} could be interpreted as the one of the major sources of the loyalty. As is described in the Figure 3.1, the demand on good 1 will increase when δ_{12} decreases. The value of δ_{12}, or more generally δ_{mn}, affects the demand directly.

$$\begin{array}{rcl} \delta_m(\omega, p_{Am}, p_N) & \equiv & u_N\left(^{\omega}/_{p_N}\right) - u_N\left(^{(\omega - p_{Am})}/_{p_N}\right) \\ \delta_{mn}(\omega, p_{Am}, p_{An}, p_N) & \equiv & -u_N\left(^{\omega - p_{Am}}/_{p_N}\right) + u_N\left(^{\omega - p_{An}}/_{p_N}\right) \; (= -\delta_m + \delta_n) \end{array} \tag{3.6}$$

The schematic image of the effect of introducing the pricing into this model is described in the Figure 3.1. Next, we consider the whole utility in this economy. By using the distribution function $\varphi(u_1, u_2)$,

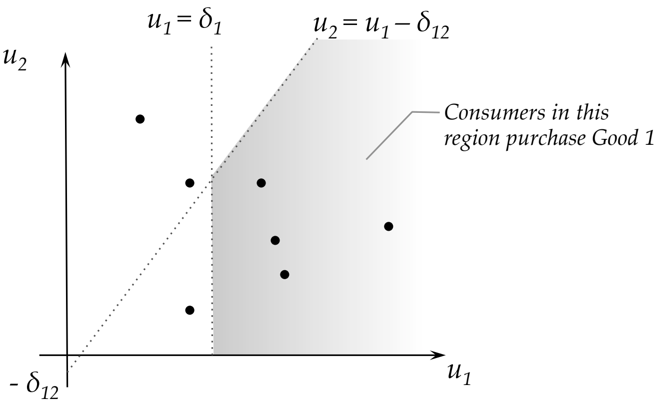

Figure 3.1: Distribution of Utility Values and Product Choice

the utility of whole economy could be described as;

$$\begin{aligned} U &= \int_0^\infty \rho(\omega)d\omega \int_{\delta_1}^\infty u_{A1}du_{A1} \int_0^{u_{A1}-\delta_{12}} \varphi(u_{A1},\ u_{A2})du_{A2} \\ &+ \int_0^\infty \rho(\omega)d\omega \int_{\delta_2}^\infty u_{A2}du_{A2} \int_0^{u_{A2}-\delta_{21}} \varphi(u_{A1},\ u_{A2})du_{A1} \end{aligned} \tag{3.7}$$

As we can describe demand of good i (X_i) and utility of whole economy U in terms of φ, we can specify utility function $U(X_1,\ X_2)$ by assuming some functional shape of φ.

3.3.3 *CES*

One of the most standard, frequently used functional type of the utility function in the macroeconomics would be the CES function. Therefore, in this section, we introduce the distribution function φ to retrieve aggregated utility function in the CES form as to be power function, just as assumed in Houthakker [20].

Firstly, let us assume the functional type of the distribution function $\varphi(u_1,\ u_2)$.

Assumption 2. *The distribution function $\varphi(u_1,\ u_2)$ follows the functional form:*

$$\varphi(u_{A1}, u_{A2}) = \begin{cases} A\,(u_{A1}+d)^{-\alpha}\,(u_{A2}-\delta_2)^{\beta_1} \\ \quad (u_{A1} > \delta_1,\ \delta_2 < u_{A2} < u_{A1}-\delta_{12},\ \alpha > 0,\ \beta_1 > 0) \\ A\,(u_{A2}+d+\delta_{12})^{-\alpha}\,(u_{A1}-\delta_1)^{\beta_2} \\ \quad (u_{A2} > \delta_2,\ \delta_1 < u_{A1} > u_{A2}-\delta_{21},\ \alpha > 0,\ \beta_2 > 0) \end{cases} . \tag{3.8}$$

Assumption 3. *Income of all consumers are the same across the economy, i.e.,* $\rho(\omega) = D(\omega - \omega_0)$ *where* $D(\omega - \omega_0)$ *is Dirac's delta-function.*
Under these assumptions, the demand of good i ($i = 1, 2$) could be

straightly shown as;

$$\begin{aligned} X_i &= A \int_{\delta_i}^{\infty} du_{Ai} \int_{\delta_j}^{u_{Ai}-\delta_{ij}} du_{Aj} \left(u_{Ai} + d\right)^{-\alpha} \left(u_{Aj} - \delta_j\right)^{\beta_i} \\ &= \tfrac{A}{1+\beta_i} \int_{\delta_i}^{\infty} du_{Ai} \left(u_{Ai} + d\right)^{-\alpha} \left(u_{Ai} - \delta_i\right)^{\beta_i} \\ &= \tfrac{A}{1+\beta_i} \left(d + \delta_i\right)^{-\alpha+\beta_i+2} \int_0^1 t_{Ai}^{\alpha-\beta_i-3} \left(1 - t_{Ai}\right)^{\beta_i+1} dt_{Ai} \\ &= \tfrac{A}{1+\beta_i} \left(d + \delta_i\right)^{-\alpha+\beta_i+2} B(\alpha - \beta_i - 2, \beta_i + 2) \end{aligned} \tag{3.9}$$

The normalization term A can be calculated by $X_1|_{\delta_1=0} + X_2|_{\delta_2=0} = 1$ and this leads to;

$$\frac{1}{A} = d^{-\alpha+2} \left\{ \frac{d^{\beta_1} B(\alpha - \beta_1 - 2, \beta_1 + 2)}{1 + \beta_1} + \frac{d^{\beta_2} B(\alpha - \beta_2 - 2, \beta_2 + 2)}{1 + \beta_2} \right\} \tag{3.10}$$

Also, the utility of the whole economy for purchasing discrete goods (defined as U_{d0}) could be also shown as;

$$\begin{aligned} U_{d0} &= A \sum_{i=1,2,\, i\neq j} \int_0^{\infty} \rho(\omega) d\omega \int_{\delta_i}^{\infty} du_{Ai} \int_0^{u_{Ai}-\delta_{ij}} du_{Aj} u_{Ai} \left(u_{Ai} + d\right)^{-\alpha} \left(u_{Aj} - \delta_j\right)^{\beta_i} \\ &= X_1|_{\alpha\to\alpha-1} + X_2|_{\alpha\to\alpha-1} - d(X_1 + X_2) \end{aligned} \tag{3.11}$$

The relation between the aggregated demand and utility in the whole economy can be obtained by eliminating the function δ (δ_1, δ_2 and δ_{12}) from the equation (3.10) and (3.11). Here let us employ two more assumptions to make equations simpler.

Assumption 4. d*,* $\delta_{12} \ll \delta_1$*,* δ_2 *holds with good approximation.*

Under this assumption, the integration in (3.10) and (3.11) can be described simply as;

$$\begin{aligned} X_i &\sim \tfrac{A}{1+\beta_i} B\left(\alpha - \beta_i - 2, \beta_i + 2\right) \delta_i^{-\alpha+\beta_i+2}, \; (i = 1, 2) \\ U_{d0} &\sim \tfrac{A}{1+\beta_1} B\left(\alpha - \beta_1 - 3, \beta_1 + 2\right) \delta_1^{-\alpha+\beta_1+3} + \tfrac{A}{1+\beta_1} B\left(\alpha - \beta_2 - 3, \beta_2 + 2\right) \delta_2^{-\alpha+\beta_2+3} \end{aligned} \tag{3.12}$$

So as the utility function exists in the real value, $\alpha - \beta_2 - 3 > 1$ for $i = 1, 2$ is required additionally. Deleting δ_i from X_i and U_{d0} to obtain the relation between aggregated demands and the utility for the discrete choice goods in the whole economy as:

$$\begin{aligned} U_{d0} &= A^{\frac{1}{\alpha-\beta_1-3}} B\left(\alpha-\beta_1-3, \beta_1+2\right) \left(\frac{1+\beta_1}{B(\alpha-\beta_1-3,\beta_1+2)}\right)^{\frac{-1}{\alpha-\beta_1-3}} X_1^{\sigma_1} \\ &\quad + A^{\frac{1}{\alpha-\beta_2-3}} B\left(\alpha-\beta_2-3, \beta_2+2\right) \left(\frac{1+\beta_2}{B(\alpha-\beta_2-3,\beta_2+2)}\right)^{\frac{-1}{\alpha-\beta_2-3}} X_2^{\sigma_2} \end{aligned} \tag{3.13}$$

where $\sigma_i = {}^{\alpha-\beta_i-3}/_{\alpha-\beta_i-2}$. Furthermore, let us assume further condition for the exponent in (3.8).

Assumption 5. *The exponent of the utility function defined in the equation (3.8) is the same across the good 1 and 2, i.e., $\beta_1 = \beta_2 = \beta$.*

Proposition 1. *The utility of the whole economy becomes CES under the Assumption 1-5.*

Proof. Plugging **Assumption 5.** into (3.13) and re-defining the utility of the whole economy U_d as $U_d = U_{d0}^{1/\sigma}$ leads to the following:

$$U_{d0} = C\left\{X_1^{\sigma} + X_2^{\sigma}\right\}^{1/\sigma} \tag{3.14}$$

where $C^{\sigma} = A^{\frac{1}{\alpha-\beta-3}} B\left(\alpha-\beta-3, \beta+2\right) \left(\frac{1+\beta}{B(\alpha-\beta-3,\beta+2)}\right)^{\frac{-1}{\alpha-\beta-3}}$.

3.3.4 *Meaning of Assumptions for the CES*

The necessary assumptions for the CES utility function to be a better approximation are that i) the distribution of consumer's utility can be well approximated in (3.8), ii) the price of each good itself is much higher than the price differences ($\delta_{12} \ll \delta_1, \delta_2$), and iii) the shape of the distribution of utilities are almost similar among goods in the choice set ($\beta_1 = \beta_2 = \beta$). The first assumption has to be confirmed with some marketing technology like a conjoint analysis. If the utilities for a certain good distributes in power low among potential consumers, this assumption can be regarded as rational for discussing approximations. This assumption reflects features of consumers under the discrete choice model, while the assumption ii) and iii) reflect features of goods of our interest. The assumption ii) and iii) becomes reasonable if the loyalties of goods are almost the same across the goods sub-vector. In other words, this assumption may not hold if there are appreciable product differentiation, especially in the field of the monopolistic competition with various goods loyalties.

3.4 PROFIT MAXIMIZATION AND PRICING UNDER THIS ECONOMY

3.4.1 *Production under This Economy*

In this section let us consider the profit maximization under the economy with demand and utility function written in (3.12) and (3.13). Suppose there are two monopolistic firms which produces non-storable discrete good 1 and 2 in the economy. Each firm can change its own price but can not affect to the price of the competitor. There is only one factor (capital) for the production of each good. If one denote the production by firm m as q_m, the cost function to produce q_m amount of good m as cq_m^γ and profit as π_m, it is easy to calculate the equilibrium as follows.

Profit Maximization
The firm m follows the profit maximization problem;

$$\begin{aligned} \max_{p_m} \pi_m &= \max_{p_m} \left\{ p_m q_m - c q_m^\gamma \right\} \\ s.t.\ X_i &= \tfrac{A}{1+\beta_i} \left(d+\delta_i\right)^{-\alpha+\beta_i+2} B(\alpha-\beta_i-2, \beta_i+2) \end{aligned} \tag{3.15}$$

Market Clearing Condition
As the good 1 and 2 are non-storable, the market clearing condition for each goods is simply described as $X_m = q_m$.

Functional Type of the δ_1
Based on **Assumption 1-3** and market clearing condition, the profit maximization condition of the firm 1 can be written as

$$\frac{\partial}{\partial p_1}\pi_1 = q_1 + p_1\frac{\partial q_1}{\partial p_1} - c\gamma q_1^{\gamma-1}\frac{\partial q_1}{\partial p_1} \tag{3.16}$$

where

$$q_1 = \frac{A}{1+\beta_1} B\left(\alpha-\beta_1-2, \beta_1+2\right)\delta_1^{-\alpha+\beta_1+2}. \tag{3.17}$$

To calculate $\frac{\partial q_1}{\partial p_1}$, we need to assume the concrete functional type of the $\delta_1(p_1)$. As defined in the section 3.2., $\delta_1(p_1) = u_N(\omega/p_N) - u_N(\omega-p_{A1}/p_N)$. The right hand side of this equation means the difference of the level of utility when the amount of the composite good is ω/p_N and $\omega-p_{A1}/p_N$. To consider its perceptual difference, it might be worth to employ the famous literature constructed by Weber and Fechner (see, for example, Weber and Schneider [51]). The most important law in behavioral psychology is Weber-Fechner's law. Here this law describes the relation between the magnitude of physical stimulus and psychological sense. If we employ this famous analogy, the indirect utility $\delta_1(p_1)$ could be described in the logarithm form as follows:

Assumption 6. *$\delta_1(p_1)$ is described in the form $B \times log\frac{\omega}{\omega - p_{A1}}$ with some constant B.*

Here it should be noted that we have not examined the existence of the direct utility function which retrieve the functional form of $\delta_1(p_1)$ in the **Assumption 6.**

Under the **Assumption 1-3.** and the **Assumption 6.**, $\frac{\partial q_1}{\partial p_1}$ can be calculated as

$$\frac{\partial q_1}{\partial p_1} = \frac{-A}{(1+\beta_1)(\omega - p_{A1})}\left(B \times log\frac{\omega}{\omega - p_{A1}}\right)^{-\alpha+1}. \tag{3.18}$$

By plugging 3.14 and 3.15 into 3.13, this maximization condition is solved and equilibrium price p_1^* can be calculated.

3.4.2 *Numerical Calculation*

Although it is possible to calculate the equilibrium with these assumptions, theoretical analysis may turn out to be complicated and hard to understand its functional features. To help its understandings on actual relations in this economy, it would be useful to calculate concretely with certain parameter value. The table 1 is the summary of the Key Parameters for the numerical calculations. Also it should be noted that we set p_N as a numeraire, so we do not set any measures for the p_N.

Parameters	Value
γ	0.8
ω	10,000
c	2.0
d	1~10
B	1
α	15~20
β_1, β_2	1~α -5

Table 3.1: Summary of Key Parameters

This table shows key variables and these values used in the numerical calculations. The parameter which determines the shape of the distribution of utilities (β_1, β_2) varies by cases.

The following numerical calculations are conducted based on the relation before the **Assumption 3**. to measure the rationality of following assumptions. As is already shown in the (3.10) or (3.12), the change in any parameter in $(\alpha, \beta_1, \beta_2)$ leads to the change in equilibrium price and also leads to the change in elasticity.

Before we move onto the concrete analysis of the numerically calculated result, let us discuss the relation between the goods loyalty and parameter value α and β_i. In the world of high α, the number consumers who feels quite high utility for holding that goods group decreases rapidly. This yields that the loyalty for holding the goods group is moderate in the high α, and vise versa. Also, in case of high β_i economy, the number of consumers who feels high utility rather holding good i decreases very rapidly. This also yields that, contrary to the case of α, the loyalty for holding good i becomes moderate in the low β_i economy.

The numerical calculation results under this parameter setting are shown in the Figure 3.2 to 3.9. The Figure 3.2 and 3.3 illustrates the equilibrium price of good 1 (p_1) and good 2 (p_2) as a function of the parameter for the substitution (β_1, β_2) in rather high goods loyalty $\alpha(= 15)$ with small d. As clearly described in the Figure 3.2, the equilibrium price p_1^* becomes lower when the loyalty of good 1 is higher $(\beta_1 \to 0)$. This means that if the productivity of the firm 1 is higher, the equilibrium price p_1^* becomes lower. One point to be noted is that the difference in the productivity does not makes any difference in this case, i.e., the difference of the β_2 does not affect to the equilibrium price of good 1. Another point to be noted is that the level of equilibrium price is not so low compared to the budget line. In any goods, the equilibrium price range is about 1000 to 4000 but the total budget is 10000 in this calculation.

The Figure 3.4-3.5 illustrates the equilibrium price of good 1 (p_1) and good 2 (p_2) as a function of the parameter for the substitution (β_1, β_2) in rather high goods loyalty $\alpha(= 15)$ with large d. The major difference from the Figure 3.2-3.3 is the level of the p_1 and p_2. As d takes a role to inflate the amount of utility of the purchased good, the purchased price is also inflated and becomes much closer to the total budget. Also, because of the shift of the utility value by d, the gradient of the equilibrium price with respect to the loyalty (i.e., $\partial p_i / \partial \beta_i$) becomes gradual compared to that of $d = 1$.

The Figure 3.6-3.7 illustrates the equilibrium price of good 1 (p_1) and good 2 (p_2) as a function of the parameter for the substitution (β_1, β_2) in rather low goods loyalty $\alpha(= 20)$ with small d. The major

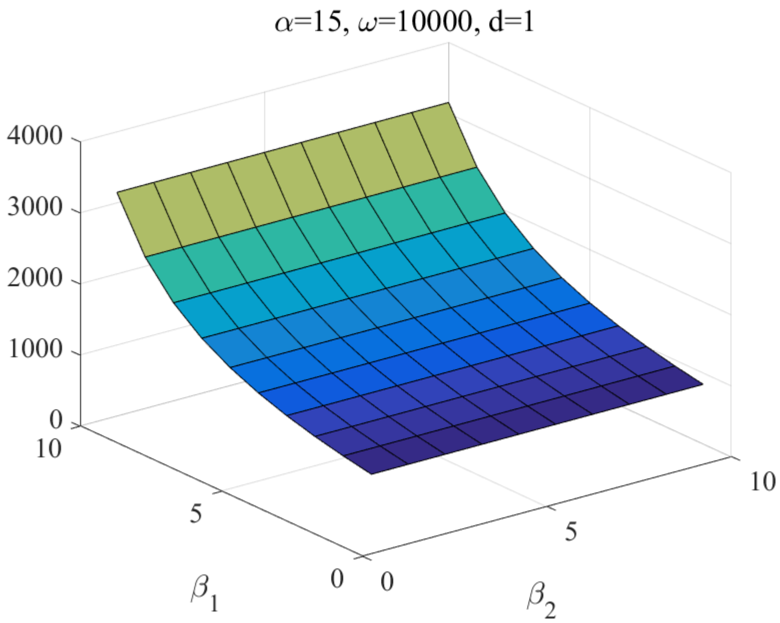

Figure 3.2: Numerical Calculation Result of p_1 in medium α and small d

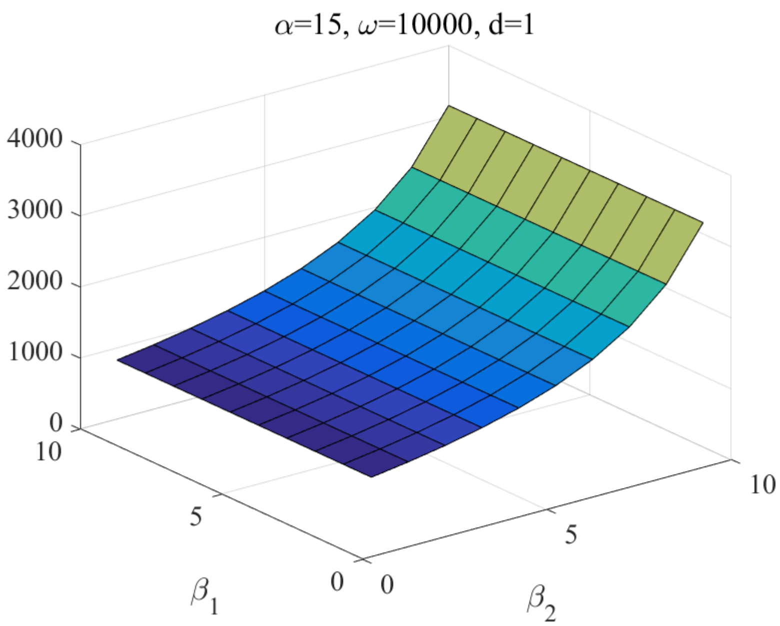

Figure 3.3: Numerical Calculation Result of p_2 in medium α and small d

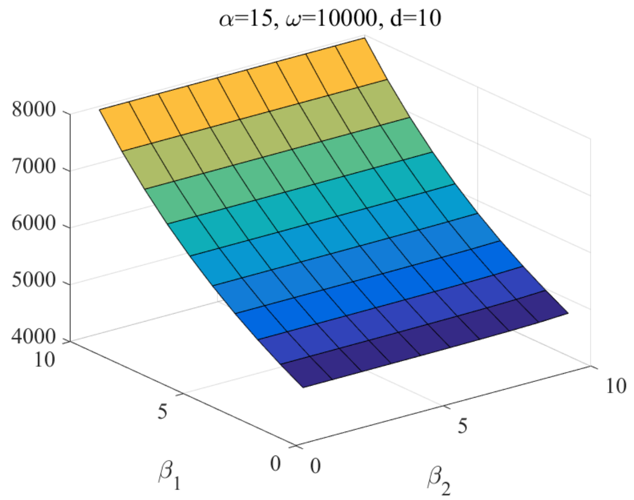

Figure 3.4: Numerical Calculation Result of p_1 in medium α and large d

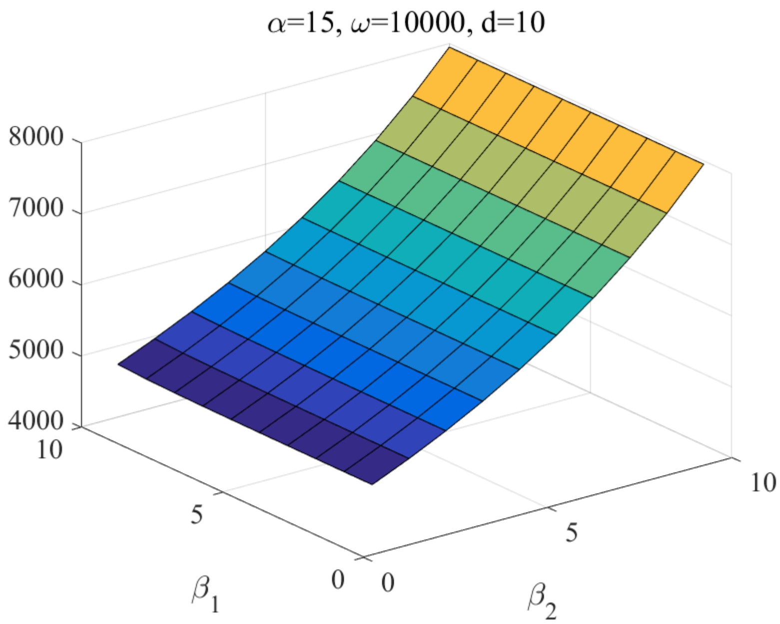

Figure 3.5: Numerical Calculation Result of p_2 in medium α and large d

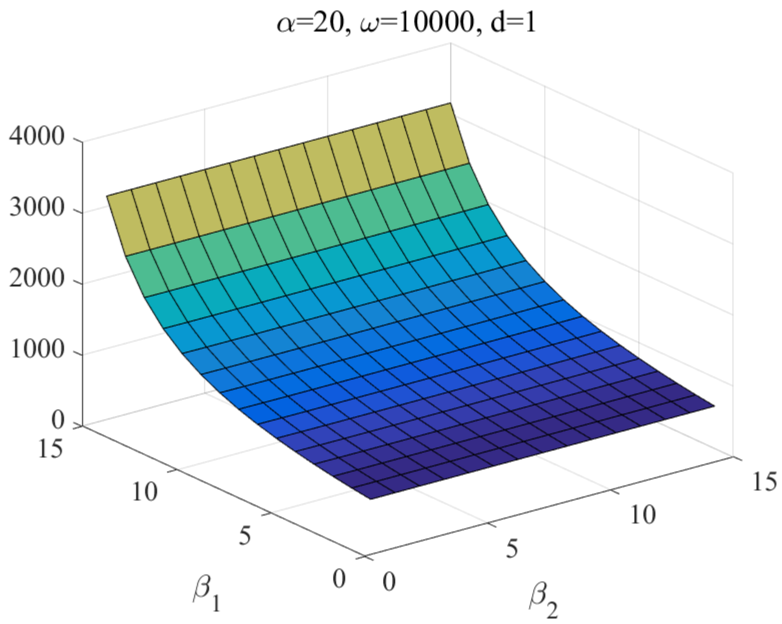

Figure 3.6: Numerical Calculation Result of p_1 in large α and small d

difference from the Figure 3.2 is only the range of the β_1 and β_2, due to the increase in α.

The Figure 3.8-3.9 illustrates the equilibrium price of good 1 (p_1) and good 2 (p_2) as a function of the parameter for the substitution (β_1, β_2) in rather low goods loyalty $\alpha(= 20)$ with large d. It is worthwhile to compare this result with Figure 3.4. In the Figure 3.5, there is obvious distortion around $(\beta_1, \beta_2) = (15, 1)$. This distortion is because of the existence of the normalization term A. The increase in β_2 leads to the increase in A, and this effect was reflected to the calculation results especially in low β_1 region.

These results are the same as that of the general monopolistic competition model like Melitz [38]. In the standard Melitz model, the demand of good m in industry l is described $X_{lm} = A_l p_{lm}^{-\frac{1}{1-\rho_l}}$ and firm's profit maximization condition yields $D_{lm} = p_{lm} X_{lm} = \rho_l^{\frac{\rho_l}{1-\rho_l}} A_l^{\frac{1}{1-\rho_l}} \omega_l^{-\frac{\rho_l}{1-\rho_l}} \theta_{lm}^{\frac{\rho_l}{1-\rho_l}}$ where $A_l = \beta_l Y P_l^{-\frac{\rho_l}{1-\rho_l}}$, ω_l is a wage and θ_{lm} is a productivity. Using these 2 relations to obtain $p_{lm}^{-\frac{\rho_l}{1-\rho_l}} = \rho_l^{\frac{\rho_l}{1-\rho_l}} A_l^{\frac{\rho_l}{1-\rho_l}} \omega_l^{-\frac{\rho_l}{1-\rho_l}} \theta_{lm}^{\frac{\rho_l}{1-\rho_l}}$, and as $\rho_l^{\frac{\rho_l}{1-\rho_l}} A_l^{\frac{\rho_l}{1-\rho_l}} \omega_l^{-\frac{\rho_l}{1-\rho_l}}$ is the same in given industry l, we can simply describe the relation of price and productivity as $p_{lm}^{-1} = c\theta_{lm}$. Under this condition, the increase in productivity leads the decrease in price.

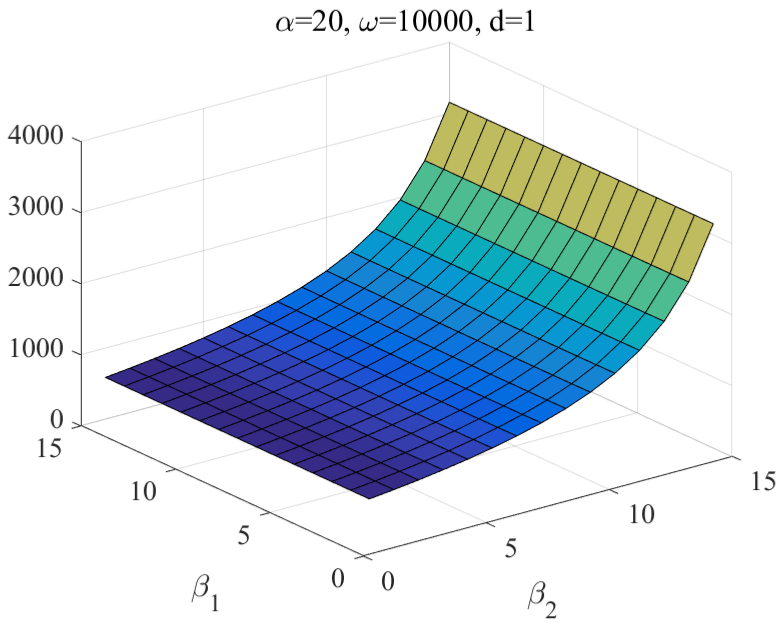

Figure 3.7: Numerical Calculation Result of p_2 in large α and small d

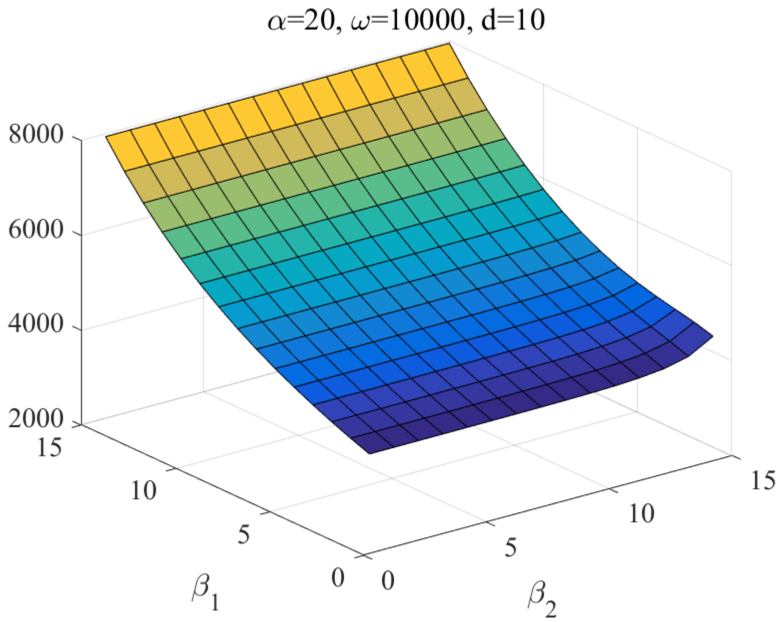

Figure 3.8: Numerical Calculation Result of p_1 in large α and large d

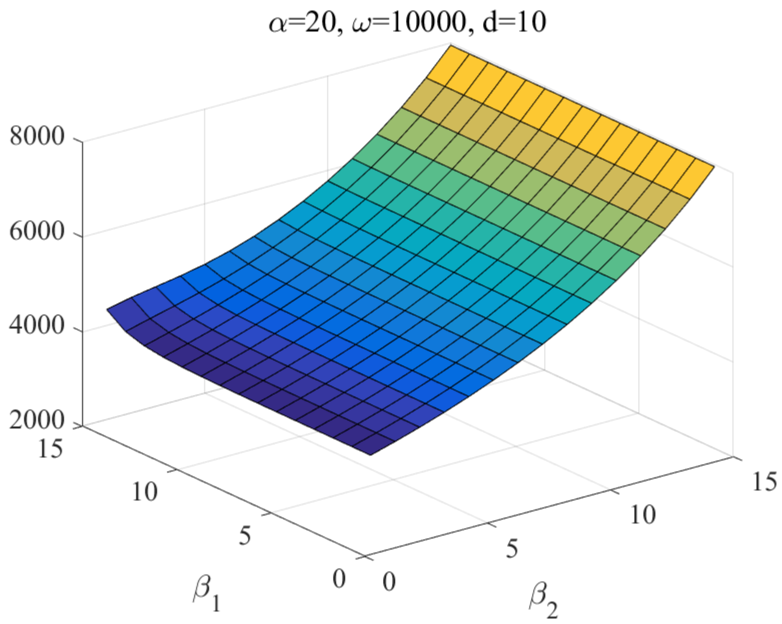

Figure 3.9: Numerical Calculation Result of p_2 in large α and large d

3.5 DISCUSSIONS

One of the major features of our model deriving such results lies in the budget constraint in this economy. In this paper we established new methodology for deriving CES utility function with microfoundation. As is already pointed out, there are 5 major assumptions required to retrieve CES utility function, and these assumptions would not be valid at least in case of the market with various level of goods loyalties. Taking this result into account, the validity to expand the CES utility function into the whole economy may change by industry, i.e., an industry with poor differentiate goods may obtain good approximation by CES utility function, however, an industry with strong differentiation may not. In our model the product differentiation affects mainly to the coefficient of X_1^{σ} just described in the equation (3.12). In case of the Dixit-Stiglitz lite (Dixit and Stiglitz [11]), utility of the representative consumer is described in CES based form for the whole goods in the economy. However, if this CES assumption may not hold for several goods subgroups or industries, this may lead the model to differ from the reality as a macro economy.

Also, in our model, each consumer purchases a product which is most attractive for him/her. In such case, the firm must increase number of consumers who like their product the best and this may lead to the market strategy of the firm, i.e., customer segmentation, promotion strategy, etc. Under such literature, we may be able to analyze that a strategy to provide a product which offers high loyalty for certain consumer segment, but not for other segment (like a good for only the professional, or say, the geek) might be able to provide better profit under certain market circumstances. In addition, this theoretical framework also enables us to analyze firm's strategies to increase R&D expenditure to achieve product innovation, or disruptive innovation which lead to rapid decrease in β.

Regarding the model development, we mostly follow Houthakker [20] for the calculation. Our model set its focus on the utility function, and Houthakker [20] on the production function. Although each focus is different, these two analyze analogous functions, CES and Cobb–Douglas. In that sense, the theoretical development of our model is limited to the derivation of not Cobb-Douglas but CES utility function. However, our key contribution would be found in i) development and interpretation of the assumptions to trace the analysis of Houthakker [20] in the aspect of utility function, and ii) derivation of the firm's profit maximization condition under such literature. Also, if we review a literature to analyze the shape of production functions, Jones [26] would be another important paper. Jones [26] assumed the distribution of idea to be Pareto and derived two results; the Cobb-Douglas global production function and labor-augmenting technical change. The major reason we chose not Jones [26] but Houthakker [20] as our theoretical foundation is that the model of Houthakker [20] is higher affinity on the theoretical extension to the CES, as Houthakker [20] treats the distribution more explicitly.

Lastly, we would like to point out that our approach allows us to provide new methodology to clarify assumptions to retrieve the CES utility function, and also hope that this theoretical framework encourages future applications in empirical works.

Part IV

CORPORATE FINANCE AND STOCHASTIC PROCESS

4 FIRM VALUE AND RETAINED EARNING

4.1 SUMMARY

We propose a model of dynamic investment, financing and risk management with retained earnings. The key contribution of this chapter is to provide dynamic model which explicitly include the retained earnings. To consider the retained earnings explicitly, we described the dynamics of both asset and liability section of the balance sheet, i.e., cash holdings and (physical) property in the asset section, and stock and retained earnings in the equity section. Our key results are: (1) the decrease in the friction on the re-investment (γ_{min}) enables a firm to adjust more rapidly to increase its cash-capital ratio when its cash-capital ratio is low, and (2) the firm's decision is also affected by the economic situation like a recession or upturn.

4.2 INTRODUCTION

The basic literature of our concern is as mentioned in the 1.2.3 of this dissertation. Papers most closely related to ours is Bolton, Chen, and Wang [4], Bolton, Chen, and Wang [5] and Decamps et al. [8]. Our research and Bolton, Chen, and Wang employ simple AK model and assume the firm's cumulative productivity evolution with a standard Brownian motion under the risk-neutral measure. Bolton, Chen, and Wang assume financial frictions directly instead of explicitly modeling an agent problem. The frictions assumed in Bolton, Chen, and Wang are related to the features of the optimal contract that motivate effort. Specifically, they assume that the firm maintains a cash balance and that it is costly to issue new equity when the firm runs out of cash. In addition, they assume that it is costly to keep cash inside the firm instead of paying it out to shareholders. On the other hand, Decamps et al. [8] analyzed model of a firm facing internal agency costs of free cash flow and equity issuance cost such as professional fee, commissions, etc.

The feature of our model is that we consider a firm facing explicit external financing cost and allowed to reserve retained earnings with the payout policy taking into account. Here we introduced 3 frictions in our model, 1) a friction on the equity issuance cost, 2) a friction on the re-investment and 3) a friction on the under-investment. The first friction is the same as Decamps et al. [8] and we introduced this friction to distinguish the financing cost for the external financing and the retained earnings. The second friction is about the ratio of the usage of the profit, and this leads to the dividend policy. In general, a profit is used whether to pay as a dividend or to re-invest to the firm's activity. According to Matos [36],

... the managers may use dividend changes to reflect a possible change in their own expectations about the future earnings of firms, due to inside information. Dividend changes could then be seen as a simple mechanism to adapt the market value to the new prospective of corporations' insiders. Accordingly, the market value of corporations would react to the announcement of dividend changes reflecting their informational content. This leads that

the dividend policy also works as a signal to the market for the managers, and then there is an incentive for the managers as an agent to send a good signal for the corporate management and growth. This friction is a distortion on choosing the ratio of re-investment, and the detail is written in 4.5.2. Lastly, the third friction is about the under-investment and the second hand market. The outline and motivation for introducing this friction is written in 4.5.4. Also, the analysis on the heterogeneity in the drift of productivity shock would be another feature of our analysis. In the numerical calculation, Bolton, Chen, and Wang [4] sets the drift of productivity shock as homogeneous across firms. Bolton, Chen, and Wang [5] changed this setting by considering it as a state variable, but the drift term has only 2 options to take. As our concern is macroeconomic dynamics under heterogeneous agent, it is natural to eliminate the restriction on the level of the drift of productivity shock. To consider the aggregation, it would be reasonable to consider that there is a distribution on the drift of productivity shock, or growth, and calculate macroeconomic variable by integrating across the distribution. So here it should be again emphasized that we consistently focus the heterogeneity of agent and calculate macro variable by integrating across the heterogeneity.

The reminder of this chapter proceeds as follows. Section 3 sets up our baseline model. Section 4 presents the model solution in the first best benchmark. Section 5 conducts quantitative analysis and Section 6 concludes.

4.3 THE MODEL

4.3.1 *Definition of the Variables*

Firstly, let us introduce the core definitions and variables on our model of dynamic capital structure choice with uncertain productivity shock. The basic framework closely follows Bolton, Chen, and Wang [4]. In our model, we add a subsection of the "Firm's Asset and Capital" and consider the capital structure explicitly.

4.3.1.1 *Production and Investment*

We consider a financially constrained firm with stochastic productivity evolution, as considered in Bolton, Chen, and Wang [4]. Firstly, we describe the firm's physical production and investment process.

$$dK_t = (I_t - \delta K_t)\, dt \tag{4.1}$$

Secondly, we assume the firm's revenue at time t to evolve proportionally to its capital stock, just assumed in the simple AK model. The dynamics of A is governed by two terms; a constant growth term described in μ and a stochastic term.

$$dA_t = \mu dt + \sigma dZ_t \tag{4.2}$$

where Z is a standard Brownian motion.

4.3.1.2 *Firm's Asset and Capital*

Consider a firm whose asset is composed of cash inventory (W_t) and property (K_t), e.g., productive facilities, and its financial resource is composed of stock, or namely shareholder's equity (S_t) and retained earnings (E_t). Although other parts are almost same as Bolton, Chen, and Wang [4], the introduction of the retained earnings (E_t) explicitly is original for our work. The introduction of the retained earnings This definition of variables leads an identical equation for the asset section and shareholder's equity section of the balance sheet as;

$$W_t + K_t = S_t + E_t \ \ (dW_t + dK_t = dS_t + dE_t)\,. \tag{4.3}$$

The schematic image of the asset and equity section is in the Figure 4.1. The firm uses its property (K_t) for the production and reserves the cash inventory (W_t) for the risk management. The financial resource of these asset is composed of the shareholder's equity (S_t) and retained earnings (E_t).

.

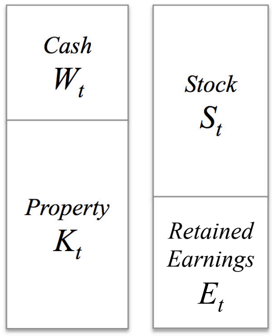

Figure 4.1: Simplified Balance Sheet without Liability

4.3.1.3 *Profit and Firm Value Maximization*

The increase in firm's cash flow (dW_t) and net income ($Y_t dt$) during a time period dt can be described as

$$\begin{aligned} dW_t &= K_t dA_t + rW_t dt - I_t dt - G\left(I_t, K_t\right) dt + \frac{1}{p} dS_t - dL_t \\ Y_t dt &= K_t dA_t + rW_t dt - \delta K_t dt - G\left(I_t, K_t\right) dt - \left(1 - \frac{1}{p}\right) dS_t \end{aligned} \tag{4.4}$$

where r is an interest rate, δK_t $(\delta \geqq 0)$ represents a depreciation of the physical stock K_t, $G\left(I_t, K_t\right)$ is the additional adjustment cost, which is an increasing function of the investment, dL_t is the dividend process and $p\ (> 1)$ is equity issuance cost as a friction for each dollar of new shares issued as is assumed in Decamps et al. [8]. The description of the firm's cash flow and net income is basically the same as other literature like Bolton, Chen, and Wang [4]. As supposed in other literature such as Demarzo, Fishman, and Wang [9], we assume that the adjustment cost satisfies $G\left(0, K_t\right) = 0$, is smooth and convex in investment I_t, and is homogeneous of degree one in I_t and K_t. We assume that there is no tax non-operational revenue nor expenditure for the firm at first.

Then, as our model setting our focus on the retained earnings, the net income is distributed to the shareholder as a dividend (dL_t) and reserved in the firm as the retained earnings (dE_t).

$$Y_t dt = dL_t + dE_t \tag{4.5}$$

where L_t is the cumulative dividend process. Here it should be noted that, in the standard model, the firm is not allowed to spare the retained earnings and therefore $Y_t dt = dL_t$ in general.

The firm value is calculated as 1) the sum of expected present values of future dividends minus 2) the sum of the expected present value of future gross issuance process, following Decamps et al. [8].

$$V = max\left[E_0 \int_0^{\tau} e^{-rt}(dL_t - dS_t)\right] + e^{-r\tau}\left(lK_\tau + W_\tau\right) \qquad (4.6)$$

where τ is the liquidation time, E_0 is the expectation operator induced by the firm's maximization process starting at $t = 0$ and l is capital liquidation value. If $\tau = \infty$, then the firm will not choose to liquidate. In general, a firm will choose to liquidate when the cost of financing is too high, or it faces the failure of heirs, etc.

4.3.2 *Optimal Choice; Capital Expansion and Payout Policy*

The firm can choose the ratio of the shareholder's equity S_t and retained earnings E_t. We assume that the firm chooses the ratio of shareholder's equity (dS_t) and retained earnings (dE_t) during a time period dt so that the summation of each value equals to the summation of dW_t and dK_t. Defining the ratio of dS_t and dE_t at time t as $\alpha_t$to describe the dependence as follows;

$$dE_t = \alpha_t\left(dW_t + dK_t\right),\ dS_t = \left(1 - \alpha_t\right)\left(dW_t + dK_t\right). \qquad (4.7)$$

If the ratio of S_t and E_t converges to the certain value, the ratio of dS_t and dE_t also converges to the same value and therefore $E_\infty = \frac{\alpha_\infty}{1-\alpha_\infty} S_\infty$.

Secondly, a firm also reserves a right to choose its payout policy directly by setting the ratio of the dL_t and dE_t, whose summation equals to the firm's net income, according to (4.5). If we define the ratio of the dL_t and dE_t as β_t $(0 \leq \beta_t \leq 1)$ for every time t, the retained earnings and dividend process during a time period dt is described as a function of the firm's net profit and β_t as;

$$dE_t = \beta_t Y_t dt,\ dL_t = (1-\beta_t)\, Y_t dt. \tag{4.8}$$

4.3.3 *Cash Inventory Dynamics and the Firm Value*

Henceforward, we first calculate the increase of the cash inventory dW_t. By using (4.7) and (4.8), $dS_t = (1-\alpha_t)\,\frac{\beta_t}{\alpha_t} Y_t dt$. Defining $\frac{\beta_t}{\alpha_t}$ as γ_t, the firm's net profit can be described in;

$$Y_t dt = \frac{1}{1+\gamma_t\left(1-\frac{1}{p}\right)(1-\alpha_t)} \left\{K_t\left(\mu dt + \sigma dZ_t\right) + rW_t dt - \delta K_t dt - G\left(I_t, K_t\right) dt\right\} \tag{4.9}$$

Also, the firm's cash inventory can be calculated by using (4.4) and (4.8) as;

$$\begin{aligned} dW_t &= Y_t dt - dK_t + dS_t - dL_t \\ &= \gamma_t Y_t dt - \left(I_t - \delta K_t\right) dt \end{aligned} \tag{4.10}$$

Next, we can calculate 1) the expected present values of future dividends minus 2) the expected present value of future gross issuance process, $dL_t - dS_t \equiv f\left(W_t, K_t\right) dt$ by using (4.4), (4.7) and (4.8) as

$$f\left(W_t, K_t\right) dt = (1-\gamma_t)\, Y_t dt. \tag{4.11}$$

(4.1), (4.4) and (4.11) yields the following Hamilton-Jacobi-Bellman Equation;

$$
\begin{aligned}
rV\left(W_t, K_t\right) \;=\; & \max_{\alpha_t,\gamma_t,I_t}\left[\frac{1-\gamma_t}{1+\gamma_t\left(1-\frac{1}{p}\right)(1-\alpha_t)}\left\{\mu K_t + rW_t - \delta K_t - G\left(I_t, K_t\right)\right\}\right. \\
& + \left[\frac{\gamma_t}{1+\gamma_t\left(1-\frac{1}{p}\right)(1-\alpha_t)}\left\{\mu K_t + rW_t - \delta K_t - G\left(i_t\right)\right\} - \left(I_t - \delta K_t\right)\right] V_W \\
& \left. + \left(I_t - \delta K_t\right) V_K + \frac{1}{2}\left\{\frac{\gamma_t \sigma K_t}{1+\gamma_t\left(1-\frac{1}{p}\right)(1-\alpha_t)}\right\}^2 V_{WW}\right]
\end{aligned}
\tag{4.12}
$$

Here V_K, $V_W$and V_{WW} represents $\frac{\partial V}{\partial K}$, $\frac{\partial V}{\partial W}$ and $\frac{\partial^2 V}{\partial W^2}$ respectively. If we set γ_t as 1 and p as 1, the HJB equation (4.12) becomes equal to that of Bolton, Chen, and Wang [4]. Also, if we set γ_t as 0, (4.12) becomes a classical description of the firm value.

4.3.4 *Simplification and Hamilton-Jacobi-Bellman Equation*

Calculations so far revealed that our assumptions lead the firm's maximization problem to be described by 2 stochastic variables (W_t and K_t) and 3 parameters to be defined by maximization condition (α_t, β_t and I_t). However, the firm's maximization problem can be reduced to a 1 stochastic variable problem by exploiting homogeneity, i.e., writing the firm value $V\left(K, W\right) = K \cdot v\left(w\right)$ where $w = {}^{W}\!/_{K}$. By defining ${}^{W_t}\!/_{K_t}$ as w_t and similarly for other variables, the key equations for the time development of the stochastic variable (4.9) can be re-described as:

$$
\begin{aligned}
dw_t \;=\; & d\left(\frac{W_t}{K_t}\right) = \frac{dW_t}{K_t} - w_t\left(i_t - \delta\right) dt \\
=\; & \frac{\gamma_t Y_t dt}{K_t} - \left(i_t - \delta\right) dt - w_t\left(i_t - \delta\right) dt \\
=\; & \left[\frac{\gamma_t}{1+\gamma_t\left(1-\frac{1}{p}\right)(1-\alpha_t)}\left\{\mu + rw_t - \delta - g\left(i_t\right)\right\} - \left(1 + w_t\right)\left(i_t - \delta\right)\right] dt \\
& + \frac{\gamma_t \sigma}{1+\gamma_t\left(1-\frac{1}{p}\right)(1-\alpha_t)} dZ_t
\end{aligned}
\tag{4.13}
$$

here we also assumed that $G_t(I_t, K_t)$ to be ${}^{G_t(I_t,K_t)}\!/_{K_t} = g\left({}^{I_t}\!/_{K_t}\right) = g\left(i_t\right) = {}^{\theta i_t^2}\!/_{2}$. Similarly, the firm's dividend minus issuance process can be reduced into a form by dividing K_t as

$$\begin{aligned}\frac{1}{K_t} f\left(w_t\right) dt &= \frac{1}{K_t}\left(1-\gamma_t\right) Y_t dt \\ &= \frac{1-\gamma_t}{1+\gamma_t\left(1-\frac{1}{p}\right)\left(1-\alpha_t\right)}\left\{\mu + r w_t - \delta - g\left(i_t\right)\right\} dt \\ &\quad + \frac{\left(1-\gamma_t\right)\sigma}{1+\gamma_t\left(1-\frac{1}{p}\right)\left(1-\alpha_t\right)} dZ_t\end{aligned} \tag{4.14}$$

Also, taking note that the marginal q is $V_K(K,W) = v(w) - w\frac{\partial v}{\partial w}$,and the marginal value of cash $V_W = \frac{\partial v}{\partial w}$ and $V_{WW} = \frac{\partial^2 v}{\partial w^2}/K_t$. According to (4.12) and (4.13), the firm value $v\left(w_t\right)$ satisfies the following the Hamilton-Jacobi-Bellman equation:

$$\begin{aligned} r v\left(w_t\right) &= \max_{\alpha_t,\gamma_t,i_t}\Bigg[\left(i_t-\delta\right) v\left(w_t\right) + \frac{1-\gamma_t}{1+\gamma_t\left(1-\frac{1}{p}\right)\left(1-\alpha_t\right)}\left\{\mu + r w_t - \delta - g\left(i_t\right)\right\} \\ &\quad + \left[\frac{\gamma_t}{1+\gamma_t\left(1-\frac{1}{p}\right)\left(1-\alpha_t\right)}\left\{\mu + r w_t - \delta - g\left(i_t\right)\right\} - \left(1+w_t\right)\left(i_t-\delta\right)\right] v^{(1)}\left(w_t\right) \\ &\quad + \frac{1}{2}\left\{\frac{\gamma_t\sigma}{1+\gamma_t\left(1-\frac{1}{p}\right)\left(1-\alpha_t\right)}\right\}^2 v^{(2)}\left(w_t\right)\Bigg]\end{aligned} \tag{4.15}$$

where $v^{(i)}$ represents $\frac{\partial^i v}{\partial w^i}$. This equation equals to the neoclassical benchmarks when there is no friction, i.e., $\gamma_t = 0$ ($\Longleftrightarrow \beta_t = 0$: no retained earnings) and $p = 1$.

4.3.5 *Tax Distortion*

The effect of tax distortions has been considered since Modigliani and Miller [41]. The major reason to consider the tax distortion is that if the model includes debt for its financial source, the tax benefit of debt appears because the interest is paid before the taxation. In our model we did not include the debt and therefore no tax distortion is expected. If we are to consider the taxation in this economy, the equation (4.9) becomes;

$$Y_t dt = \frac{1-\kappa}{1+\gamma_t\left(1-\frac{1}{p}\right)(1-\alpha_t)}\left\{K_t\left(\mu dt+\sigma dZ_t\right)+rW_t dt-\delta K_t dt-G\left(I_t,K_t\right)dt\right\} \tag{4.16}$$

where κ is the corporate tax rate ($0 \leq \kappa \leq 1$). The derivation of the Hamilton-Jacobi-Bellman equation is straight forward based on this equation. For the simplicity, we set κ to be zero as the taxation does not provide any distortion in our model.

4.3.6 *Restriction on the Investment Rate*

In the previous literature like Bolton, Chen, and Wang [4] and Decamps et al. [8], the increase in the firm's operating revenue is calculated as $K_t dA_t$. In the standard AK model, the operating revenue of the firm is described as $A_t K_t$ and the increase at time t corresponds to $d\left(A_t K_t\right) = K_t dA_t + A_t dK_t$. The second term corresponds to the growth in capital (dK_t) but the introduction of the second term makes the comparative statics difficult as the productivity diverges as long as $\mu > 0$. Therefore, in this dissertation, we assume only the term $K_t dA_t$ for the simplicity.

4.3.7 *Liquidation and Refinancing*

In the previous literature like Bolton, Chen, and Wang [4], the firm chooses whether liquidation or refinancing when it runs out of cash. If the firm chooses to increase its shareholder's equity, it must pay the financing costs which is defined as p in our model. The firm may choose to liquidate when, for example, the financing cost is too high or the return on capital is too low. In the following analysis, we mainly focus not on the refinancing case but the liquidation case for the following reasons. First, liquidation is one of the major issues in many articles, Hugonnier and Morellec [23] and DeMarzo and Sannikov [10]. Although the choice of the liquidation/refinancing differs by parameter values, it might be reasonable to consider that the financially constrained firms does not have an opportunity for refinancing when it runs out of cash. In most cases, the risk premium for financially constrained firms are high, and not so many investors are willing to provide additional investment when firms fail to manage their cash holdings.

4.3.8 *Restrictions in Internal Financing Region*

In the liquidation case, the firm's policy functions and the firm value shows different dynamics in two regions: (i) an internal financing region, and (ii) payout region. The payout region is achieved when the firm holds cash stock W greater than or equal to the upper barrier $\overline{W}$ which is determined endogenously. And the internal financing region is achieved when the firm holds cash stock W lower than the upper barrier $\overline{W}$ and higher than the lower barrier $\underline{W}$. In the liquidation case, lower barrier becomes $\underline{W} = 0$. Here the liquidation value is assumed to be proportional to the firm's capital; $V(\underline{W}) = lK$. We therefore have

$$v(0) = l. \tag{4.17}$$

The firm value in higher barrier $\overline{W}$ is defined through value matching, smooth pasting and super contact conditions. Here the value matching condition means the firm value in internal financing region $v_i(w)$ and payout region $v_p(w)$ matches at the higher barrier $\overline{w}$; that is $v_i(\overline{w}) = v_p(\overline{w})$. Also, the smooth pasting condition means the derivative of the firm value also matches at the higher barrier; that is $v_i'(\overline{w}) = v_p'(\overline{w})$. Lastly, the super contact condition means the second derivatives of the firm value match at the boundary; $v_i''(\overline{w}) = v_p''(\overline{w})$. In the payout region, the firm value linearly increases as w increases. Moreover, the firm value must be continuous before and after the cash distribution. So the firm value in the payout region can be described as

$$v_p(w) = v_p(\overline{w}) + w - \overline{w}. \tag{4.18}$$

These conditions leads to;

$$v_i'(\overline{w}) = 1 \tag{4.19}$$

and

$$v_i''(\overline{w}) = 0. \tag{4.20}$$

4.4 MODEL SOLUTION

4.4.1 *First Best Benchmark*

Before considering the firm's dividend policy with issuance cost, we analyze the benchmark case which does not assume such friction. In the case without the financial friction, the equity issuance cost $p = 1$ and there is no noise to the productivity, i.e., $\sigma = 0$. Also, because there are no issuance costs in the benchmark economy, cash reserves do not increase the firm value. It is therefore optimal for the firm not to hold any cash for all the time $w_t = 0\ for\ ^{\forall}t \geqq 0$. Under such benchmark case, we can easily retrieve the Modigliani and Miller logic. If we go back to the equation (4.1) and assume i to be constant across time, the firm's physical property evolves as $K_t = K_0 e^{(i-\delta)t}$. Next, 1) the expected present values of future dividends minus 2) the expected present value of future gross issuance process evolves as the following;

$$dL_t - dS_t = \frac{\alpha_t - \beta_t}{\alpha_t} \left(\mu + rw - \delta - g\left(i\right)\right) K_t dt \tag{4.21}$$

where $0 \leq \alpha_t, \beta_t \leq 1$. As $\frac{\alpha_t - \beta_t}{\alpha_t}$ is increasing with respect to α_t and decreasing with respect to β_t, the firm chooses α_t to be 1 and β_t to be 0 to maximize its firm value. This leads to the firm's value to be

$$\begin{aligned} V_{FB} &= \max_{i_t} \int_0^\infty e^{-rt} \left(dL_t - dS_t\right) \\ &= \max_{i_t} \int_0^\infty \left(\mu - \delta - g\left(i\right)\right) K_0 e^{(i-\delta-r)t} dt \\ &= \max_{i_t} \left[\tfrac{K_0}{r+\delta-i} \left(\mu - \delta - g\left(i\right)\right)\right] \end{aligned} \tag{4.22}$$

Here we assumed that $r + \delta - i$ to be positive. In this framework, the Tobin's average q calculated is described as;

$$q_{FB} = \max_{i} \left[\frac{\mu - \delta - g\left(i\right)}{r + \delta - i}\right]. \tag{4.23}$$

This expression slightly different from other literature such as Bolton, Chen, and Wang [4] and Demarzo, Fishman, and Wang [9]. In the standard previous literature, the numerator of the average Tobin's q is $\mu - i - g\left(i\right)$, despite our model is $\mu - \delta - g\left(i\right)$. The fundamental difference in the expression of numerator is the determining process of the dividend. In the previous literature, the dividend is calculated based on the cash flow of each period. On the other hand, in our model, dividend is calculated based on the profit of the period. This difference arises because our model sets its focus on the retained earnings. In the account process, retained earnings is defined as the profit after dividend payment. To describe this feature, we set firms decision process to determine its dividend not from the current cash flow but from firm's net profit, or profit after taxation.

4.4.2 *Comparative Statics*

The Hamilton-Jacobi-Bellman equation (4.15) describes the optimization control of the firm value under uncertainty and frictions at every time t. In this section we consider the comparative statics under the financial frictions of the Hamilton-Jacobi-Bellman equation. As the second order differential equation is highly complicated, it is difficult to provide theoretical solution of the equation. However, it is still possible to provide several parameter restrictions with simple calculations.

The first issue to be considered in this section is the range of γ_t. As γ_t is defined as $\frac{\beta_t}{\alpha_t}$ and $0 \leq \alpha_t, \beta_t \leq 1$, it is natural to set the range of γ_t to be $0 \leq \gamma_t < \infty$. However, according to (4.14) the source of the firm value turns to be negative when $\gamma_t > 1$. It is still possible if $dL_t - dS_t$ tentatively turns to be negative due to large productivity shocks or any other accidents, but such action is not sustainable and allowed only in short term. Therefore, in the comparative statics, it is reasonable to set γ_t to be $0 \leq \gamma_t \leq 1$ and therefore $\alpha_t > \beta_t$.

The second issue is the dynamics of the cash holdings, i.e., (4.13). In the comparative statics, it is also reasonable to assume $E_0 \left[dw_t\right]_{t\to\infty}$ to be zero. This restriction actually provides the condition for the investment rate i_t to satisfy;

$$\frac{\gamma_\infty}{1+\gamma_\infty\left(1-\frac{1}{p}\right)\left(1-\alpha_\infty\right)}\left\{\mu + rw_\infty - \delta - g\left(i_\infty\right)\right\} = \left(1+w_\infty\right)\left(i_\infty - \delta\right). \tag{4.24}$$

(4.24) provides the relation between i_∞ and w_∞. The appropriate cash holding rate w_∞ can be calculated numerically from (4.15) and using (4.23) to provide the appropriate rate of investment in $t \to \infty$. However, here the problem occurs as there are two independent methodologies to determine investment rate, one is from maximization condition of (4.15) and another is the restriction on the dynamics of cash holdings (4.24).

The third issue is the first order condition of (4.15). It is straightforward to calculate the first order condition of (4.15) for 3 variables,

i_t, α_t and γ_t. The first order condition for the three variables are as follows;

$$
\begin{aligned}
[i_t]: \quad & i_t = -\frac{1}{\theta}\frac{v-(1+w_t)v^{(1)}}{1-\gamma_t+\gamma_t v^{(1)}}\left\{1+\gamma_t\left(1-\frac{1}{p}\right)(1-\alpha_t)\right\} \\
[\alpha_t]: \quad & (1-\gamma_t+\gamma_t v)\left(\mu+rw_t-\delta-g\left(i_t\right)\right) \\
& +\frac{\gamma_t^2\sigma^2}{1+\gamma_t\left(1-\frac{1}{p}\right)(1-\alpha_t)}v^{(2)}=0 \\
[\gamma_t]: \quad & \left(\mu+rw_t-\delta-g\left(i_t\right)\right)\left\{v^{(1)}-1-\left(1-\frac{1}{p}\right)(1-\alpha_t)\right\} \\
& +\frac{\gamma_t\sigma^2 v^{(2)}}{1+\gamma_t\left(1-\frac{1}{p}\right)(1-\alpha_t)}=0
\end{aligned}
\tag{4.25}
$$

In previous papers such as Bolton, Chen, and Wang [4], there is only one first order condition and it is feasible to plug into the HJB equation. On the other hand, in our model, the first order conditions are very complicated, and it might not be feasible to calculate analytically. Therefore, in the next session, we try to solve these equations numerically with newly proposed methodology.

4.5 QUANTITATIVE ANALYSIS; NUMERICAL SOLUTION FOR THE HJB EQUATION

The analytical solution of the HJB equation shall be calculated by solving second order differential equation with first order conditions (4.25). However, it is difficult to analyze the second order differential equation (4.15) under (4.25) analytically. Therefore here we conduct numerical analysis to to solve HJB equation and policy functions for α_t, β_t and i_t.

Here the difference between the analysis of Bolton, Chen, and Wang [4] should be emphasized. In standard economic model, a solution after substituting first order conditions into (4.15) are calculated numerically. As a difference from the model of Bolton, Chen, and Wang [4], our model includes 3 parameters (i.e. α_t, β_t and i_t) as control variables and therefore the introduction of the first order condition to the HJB equation becomes much complicated. Taking such situation into consideration, we developed a methodology to calculate both the solution of HJB equation and policy functions.

Parameters	Symbol	Value
Risk-free rate	r	6%
Rate of depreciation	δ	10.07%
Risk-neutral mean productivity shock	μ	18%
Volatility of productivity shock	σ	9%
Adjustment cost parameter	θ	1.5
Capital liquidation value	l	0.9
Equity issuance cost	p	1.05

Table 4.1: Summary of Key Parameters
This table shows key variables and these values used in the numerical calculations.

Table 1 summarizes the parameter values for the numerical analysis. Most variables are based on Bolton, Chen, and Wang [4] and Decamps et al. [8]. Also, Table 2 summarizes conditions for the range of control variables, which is mainly described in 4.4.

Parameters	Symbol	Value
Capital Expansion Policy	α_t	$10\% \leq \alpha_t \leq 100\%$
Payout Policy	β_t	$\alpha_t \times \gamma_{min} < \beta_t < \alpha_t$
Investment	i_t	$-100\% \leq i_t \leq (r+\delta)$

Table 4.2: Restrictions on Control Variables
This table summarizes the restrictions on control variables used in the model. Regarding the range of the investment, the first best benchmark calculation exerts relation as $i_t < r + \delta$.

4.5.1 *Brief Protocol of the Numerical Calculation*

The numerical calculation aims to solve the second order differential equation (4.15) under first order conditions (4.25) and boundary conditions (4.17), (4.19) and (4.20) to determine the level of α_t, β_t and i_t. However, there are 2 major difficulties for solving this problem. The first problem is that when we plug first order conditions into the HJB equation, the equation becomes too complicated and cannot be solved analytically. The second problem is about the boundary conditions. If the initial values $v(0)$, $v^{(1)}(0)$ is set exogenously, it is straight forward to solve the differential equation by using a method like the Runge-Kutta. However, we only have explicit initial value as $v(0) = l$, and $v^{(1)}(0)$ is not defined. Instead, we have to find a value $v^{(1)}(0)$ which satisfies (4.19) and (4.20) for certain $\overline{w}$.

To solve the first problem, we simply maximize the right hand side of the (4.15) under the restrictions of control variables (Table 2) in every w_t. In detail, 1) we first set the initial values $v(0) = l$, $v^{(1)}(0)$ and then search a set of control variables $(i_t, \alpha_t, \gamma_t)$ which maximize the right hand side of the (4.15) at $w_t = 0$ through the MATLAB function; fmincon. 2) Then we solve the differential equation (4.15) by the Runge-Kutta 4th order to obtain $v(h)$, $v^{(1)}(h)$ and $v^{(2)}(h)$ where h is the step size and here defined as $1/10{,}000$. We iterate these 2 steps to solve until 1) $v^{(1)}(w_t)$ becomes less than 1 (4.19) or 2) $v^{(2)}(w_t)$ becomes 0 (4.20). Here it should be noted that $v^{(2)}(w_t)$ is negative throughout $0 < w_t < \overline{w}$, as the marginal value of cash is expected to decrease as w_t increases. This is mainly because of the benefit to hold the cash. In general, the benefit for the firm to hold the cash is to avoid bankruptcy due to the productivity shock. However, this merit decreases as the cash to capital ratio increases, and therefore the $v^{(1)}(w_t)$ is expected to decrease with respect to w_t.

4.5.1.1 *Adjustment Process for* $v^{(1)}(0)$

After we reach condition whether 1) $v^{(1)}(w_t)$ becomes less than 1 or 2) $v^{(2)}(w_t)$ becomes less than 0, we iterate the calculation by adjusting the value $v^{(1)}(0)$. We employ the adjustment process of $v^{(1)}(0)$ as follows. First, we set the lower and higher barrier of $v^{(1)}(0)$ as $v_l(1)$ and $v_h(1)$. Then we start first iteration ($n = 1$) with setting $v^{(1)}(0)|_{n=1}$ as $(v_h(1)+v_l(1))/2$. If $v^{(1)}(w_t)$ becomes less than 1 before $v^{(2)}(w_t)$ becomes larger than 0, we start next iteration with setting $v_l(2)$ as $(v_h(1)+v_l(1))/2$ and $v_h(2)$ as $v_h(1)$. Otherwise, which means if $v^{(2)}(w_t)$ becomes larger than 0 before $v^{(1)}(w_t)$ becomes less than 1, we start next iteration with setting $v_h(2)$ as $(v_h(1)+v_l(1))/2$ and $v_l(2)$ as $v_l(1)$. In both cases, $v^{(1)}(0)|_{n=2}$ is set as $(v_h(2)+v_l(2))/2$ in the second iteration. When we iterate these processes for N times, the range of the high barrier and lower barrier of $v^{(1)}(0)|_N$ decreases with a speed of $(v_h(1)-v_l(1))/2^N$.

To understand the meaning of this adjustment process, let us describe the convergence process in different $v^{(1)}(0)$. If $v^{(1)}(0)$ is exactly equal to the appropriate value, the $v^{(2)}(w_t)$ becomes 0 and $v^{(1)}(w_t)$ becomes 1 at $w_t = \overline{w}$. In case $v^{(1)}(0)$ is higher than the appropriate value, it would be natural to consider that $v^{(1)}(w_t)$ remains high even $v^{(2)}(w_t)$ becomes 0, as the starting value is higher than the appropriate value. On the other hand, in case $v^{(1)}(0)$ is lower than the appropriate value, $v^{(1)}(w_t)$ reaches to the target value 1 faster than $v^{(2)}(w_t)$ becomes 0 for the similar reason. These intuitive descriptions can be achieved when i) $v^{(1)}(w_t)$ is monotonously decreasing and starting from a certain positive value, ii) $v^{(2)}(w_t)$ is monotonously increasing and starting from a certain negative value, and iii) these basic functional form does not change as different $v^{(1)}(0)$.

4.5.1.2 *Calculations for $w_t = 0$*

To illustrate steps to calculate $v(w_t)$ in concrete, let us describe these steps in equations. Here we assume that the firm values and its derivatives ($v(w_t)$, $v^{(1)}(w_t)$ and $v^{(2)}(w_t)$) are already calculated, but i) a set of control variables $(i_t, \alpha_t, \gamma_t)$ and ii) firm values and its derivatives in next step ($v(w_t + h)$, $v^{(1)}(w_t + h)$, and $v^{(2)}(w_t + h)$) is to be calculated.

To begin with, we have to consider the first step, i.e., $w_t = 0$. The initial value $v(0) = l$ is exogenously determined. On the other hand, we have to estimate $v^{(1)}(0)$ through the iteration process written above. First we set $v_l(1)$ and $v_h(1)$ exogenously, such as $v_l(1) = 1$ and $v_h(1) = 10^3$. Then set $v^{(1)}(0)\,|_{n=1}$ as $(v_h(1)+v_l(1))/2$ and proceed calculation. Here the problem is how to set $v^{(2)}(0)$, $\alpha_t(0)$, $\gamma_t(0)$, and $i_t(0)$ as a solution. In principle, variables $v^{(2)}(0)$, $\alpha_t(0)$, $\gamma_t(0)$, and $i_t(0)$ are determined to minimize the following function and HJB equation ($F(\alpha_t(0), \gamma_t(0), i_t(0)) = rv(0)$):

$$
\begin{aligned}
F(\alpha_t(0), \gamma_t(0), i_t(0))\,|_{w_t=0} &= (i_t(0) - \delta)\, v(0) \\
&+ \frac{1-\gamma_t(0)}{1+\gamma_t(0)\left(1-\frac{1}{p}\right)(1-\alpha_t(0))} \{\mu - \delta - g(i_t(0))\} \\
&+ \left[\frac{\gamma_t(0)}{1+\gamma_t(0)\left(1-\frac{1}{p}\right)(1-\alpha_t(0))} \{\mu - \delta - g(i_t(0))\} - (i_t(0) - \delta) \right] v^{(1)}(0) \\
&+ \frac{1}{2} \left\{ \frac{\gamma_t(0)\sigma}{1+\gamma_t(0)\left(1-\frac{1}{p}\right)(1-\alpha_t(0))} \right\}^2 v^{(2)}(0)\,.
\end{aligned}
\tag{4.26}
$$

As all variables except for $v^{(2)}(0)$, $\alpha_t(0)$, $\gamma_t(0)$, and $i_t(0)$ are already determined, we can easily find a solution $\alpha_t^*(0)$, $\gamma_t^*(0)$, and $i_t^*(0)$ through MATLAB fmincon function. However, the problem is that $v^{(2)}(0)$ can only calculated with using (4.15) and a solution $\alpha_t^*(0)$, $\gamma_t^*(0)$, and $i_t^*(0)$, but the value of $v^{(2)}(0)$ also affects the solution $\alpha_t^*(0)$, $\gamma_t^*(0)$, and $i_t^*(0)$. Therefore we iterate this process for 10 times and wait until $v^{(2)}(0)$, $\alpha_t^*(0)$, $\gamma_t^*(0)$, and $i_t^*(0)$ converges for certain values.

Then we can proceed the calculation to calculate $v(h)$, $v^{(1)}(h)$ and $v^{(2)}(h)$ through using the following method.

4.5.1.3 *Calculations of $w_t + h$ under Given w_t*

Here we consider to calculate the set of control variables which maximize the following function $F(\alpha_t, \gamma_t, i_t)$ under the given $v(w_t)$, $v^{(1)}(w_t)$ and $v^{(2)}(w_t)$ and restrictions on control variables.

$$
\begin{aligned}
F\left(\alpha_t,\gamma_t,i_t\right) &= \left(i_t-\delta\right) v\left(w_t\right) \\
&+\tfrac{1-\gamma_t}{1+\gamma_t\left(1-\frac{1}{p}\right)\left(1-\alpha_t\right)}\left\{\mu+rw_t-\delta-g\left(i_t\right)\right\} \\
&+\left[\tfrac{\gamma_t}{1+\gamma_t\left(1-\frac{1}{p}\right)\left(1-\alpha_t\right)}\left\{\mu+rw_t-\delta-g\left(i_t\right)\right\}-\left(1+w_t\right)\left(i_t-\delta\right)\right] v^{(1)}\left(w_t\right) \\
&+\tfrac{1}{2}\left\{\tfrac{\gamma_t\sigma}{1+\gamma_t\left(1-\frac{1}{p}\right)\left(1-\alpha_t\right)}\right\}^2 v^{(2)}\left(w_t\right)
\end{aligned}
\tag{4.27}
$$

If we define the solution of (4.22) as $\left(i_t^*,\alpha_t^*,\gamma_t^*\right)$, the firm values and its derivatives in next step ($v\left(w_t+h\right)$, $v^{(1)}\left(w_t+h\right)$ and $v^{(2)}\left(w_t+h\right)$) can be calculated through the Runge-Kutta 4th order. Before we move on to the concrete calculation, let us define a new function $H\left(=v^{(2)}\left(w_t\right)\right)$ as follows:

$$
\begin{aligned}
H\left(v\left(w_t\right),v^{(1)}\left(w_t\right)\right) &= \left\{\tfrac{2\left(1+\gamma_t^*\left(1-\frac{1}{p}\right)\left(1-\alpha_t^*\right)\right)}{\gamma_t^*\sigma}\right\}^2 \times \\
&\left[\left(i_t^*-\delta-r\right) v\left(w_t\right)+\tfrac{1-\gamma_t^*}{1+\gamma_t^*\left(1-\frac{1}{p}\right)\left(1-\alpha_t^*\right)}\left\{\mu+rw_t-\delta-g\left(i_t^*\right)\right\}\right. \\
&\left.+\left[\tfrac{\gamma_t^*}{1+\gamma_t^*\left(1-\frac{1}{p}\right)\left(1-\alpha_t^*\right)}\left\{\mu+rw_t-\delta-g\left(i_t^*\right)\right\}-\left(1+w_t\right)\left(i_t^*-\delta\right)\right] v^{(1)}\left(w_t\right)\right]
\end{aligned}
\tag{4.28}
$$

If we follow the standard approximation in derivations, $v^{(0)}\left(w_t+h\right)-v^{(0)}\left(w_t\right)$ can be approximated as $h\times v^{(1)}\left(w_t\right)$ and $v^{(1)}\left(w_t+h\right)-v^{(1)}\left(w_t\right)$ can be approximated as $h\times v^{(2)}\left(w_t\right)$. Here, in this dissertation, we use Runge-Kutta 4th order for the better approximation. With defining $u\left(w_t\right)\equiv v^{(1)}\left(w_t\right)$ to obtain the solution as:

$$\begin{aligned}
u\left(w_t+h\right) &= u\left(w_t\right)+\tfrac{h}{6}\left(j_{1t}+2j_{2t}+2j_{3t}+j_{4t}\right) \\
v\left(w_t+h\right) &= v\left(w_t\right)+\tfrac{h}{6}\left(k_{1t}+2k_{2t}+2k_{3t}+k_{4t}\right) \\
j_{1t} &= hH\left(w_t, v\left(w_t\right), u\left(w_t\right)\right) \\
k_{1t} &= h\cdot u\left(w_t\right) \\
j_{2t} &= hH\left(w_t+\tfrac{h}{2}, v\left(w_t\right)+\tfrac{k_{1t}}{2}, u\left(w_t\right)+\tfrac{j_{1t}}{2}\right) \\
k_{2t} &= h\cdot\left(u\left(w_t\right)+\tfrac{j_{1t}}{2}\right) \\
j_{3t} &= hH\left(w_t+\tfrac{h}{2}, v\left(w_t\right)+\tfrac{k_{2t}}{2}, u\left(w_t\right)+\tfrac{j_{2t}}{2}\right) \\
k_{3t} &= h\cdot\left(u\left(w_t\right)+\tfrac{j_{2t}}{2}\right) \\
j_{4t} &= hH\left(w_t+h, v\left(w_t\right)+k_{3t}, u\left(w_t\right)+j_{3t}\right) \\
k_{4t} &= h\cdot\left(u\left(w_t\right)+\tfrac{j_{3t}}{2}\right).
\end{aligned} \tag{4.29}$$

Iterating this procedure from $w_t = 0$ to $w_t = \overline{w}$ to calculate the solution of $v\left(w_t\right)$. Here it should be emphasized that what we calculate here is not w_{t+1} but $w_t + h$ for any given w_t. This means that what we analyze here is not the dynamics of the firm value but just a solution of HJB equation. If we are to analyze the exact dynamics of the firm value and cash-capital ratio, we had better analyze dw_t simultaneously.

Also, it should be noted that the control variable γ_t lies in the denominator. In this regard, the variable γ_t cannot be zero as long as we calculate the solution from the equations above. It is in some sense natural as, in the classical firm valuation, the firm value linearly increases when w_t increases, and there is no necessity to consider $v^{(2)}\left(w_t\right)$. However, this point makes it difficult to calculate the firm value dynamics in general. To avoid this difficulty, we set $\gamma_{min} > 0$ as a minimum value of the γ_t and assume that the γ_t stays within $\gamma_{min} \leq \gamma_t \leq 1$. Then another problem arises: how much is the value of the γ_{min}?

4.5.2 *A Friction on the Re-investment*

Before we consider the condition for γ_{min}, let us consider the feature of the γ_t in more detail. The equation (4.10) yields

$$Y_t dt = \left(dW_t + dK_t\right) + \left(dL_t - dS_t\right) = \gamma_t Y_t dt + \left(1-\gamma_t\right) Y_t dt \ . \tag{4.30}$$

This relation describes the feature of the γ_t well. In general, the firm's profit can be distributed to whether 1) to re-invest and increase the capital, or 2) to pay as a dividend. The case $\gamma_t = 0$ means that there are no re-investment and all profits are paid as a dividend. If we set the γ_{min} as positive, we do not allow the firm to pay all profits as a dividend, but force to re-invest at least in certain ratio.

Figure 4.2 represents the historical change of γ_t by the different firm size in Japan. According to the Financial Statements Statistics of Corporations by Industry, Ministry of Finance, Japan, the γ_t in big firms (capital stock is more than 1 billion yen) is around 53% in 2017, and the γ_t in small firms (capital stock is less than 10 million yen) is around 98%. Especially in case of small firms, it might be reasonable to set $\gamma_{min} > 0$. In general, if a firm pays all of its profit out as dividends or does not reinvest back into the business ($\gamma_t = 0$), investors tend to consider earnings growth of the firm might suffer. If such understandings is common among the market, it might be natural for managers to set some lower limit on the γ_t to maintain the stock price.

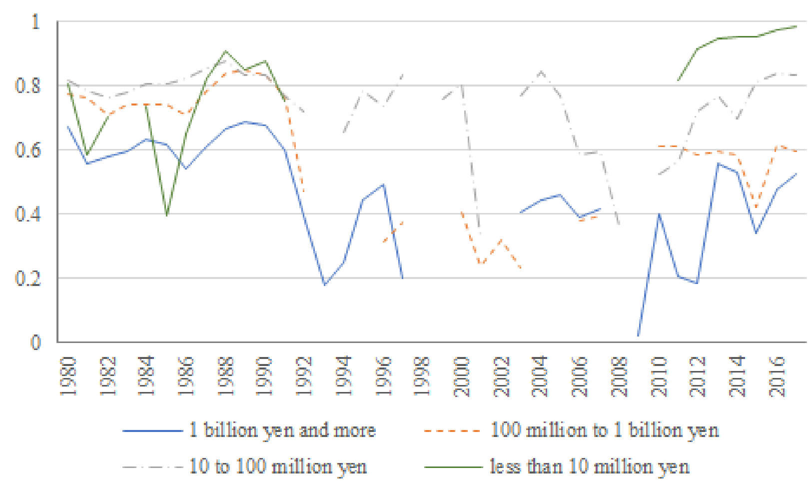

Figure 4.2: Historical Change of γ_t by the Firm Size

4.5.3 *Major Calculation Results for Internal Financing Region*

4.5.3.1 *Replication of Bolton, Chen, and Wang [4]*

Figure 4.3 plots simulated results of A) Firm Value-Capital Ratio, B) Marginal Value of Cash, C) Investment-Capital Ratio, D) dw_t, E) Policy Function of α and F) Policy Function of γ under the restrictions of $\alpha = \gamma = 1$ for any w_t. It can be easily confirmed that these conditions replicate conditions of Bolton, Chen, and Wang [4]. According to Bolton, Chen, and Wang [4], the marginal value of cash reaches a value of 30 as w approaches zero, and endogenous payout boundary $\overline{w} = 0.22$. These values are replicated in our analysis.

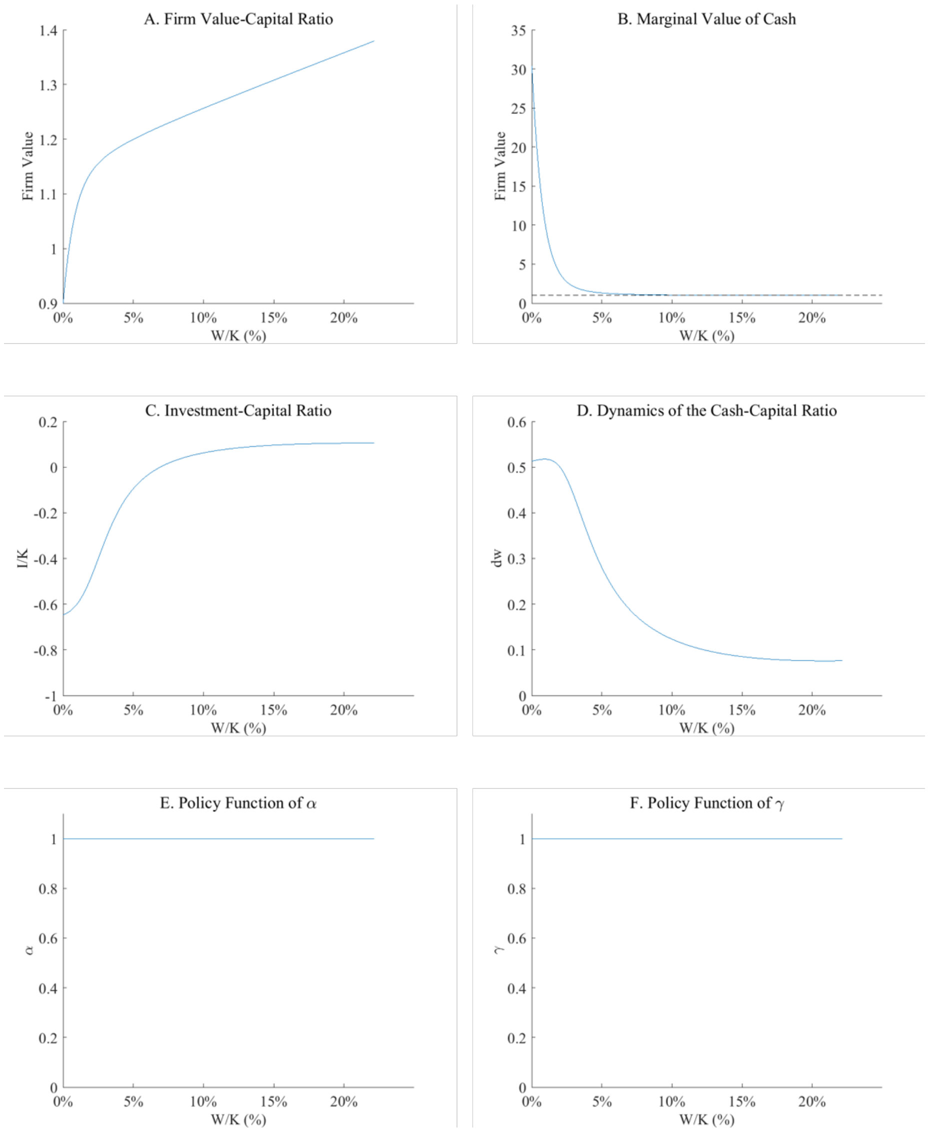

Figure 4.3: Firm Value and Related Variables under Bolton (2011) conditions

4.5.3.2 *Case of* $\gamma_{min} = 1.0$

Figure 4.4 plots simulated results of similar variables under $\gamma_{min} = 1.0$. As γ_{min} is assumed to be 1, the policy function of γ becomes constant at $\gamma = 1$, and the firm maximizes its value by adjusting α_t $(= \alpha(w_t))$ and i_t $(= i(w_t))$. The major difference with respect to Bolton, Chen, and Wang [4] is the change in $\alpha(w_t)$. $\alpha(w_t)$ becomes α_{min} in low w_t and α_{max} in high w_t. This means that when the firm's cash-capital ratio is low, the firm increases the ratio of shareholder's equity to finance adequate amount of cash from the stock market. However, as there is an equity issuance cost for increasing the shareholder's equity, the firm changes its financial source only to the retained earnings ($\alpha(w_t) = \alpha_{max} = 1$) when the cash-capital ratio becomes adequately high. Due to the difference in the policy function of

$\alpha(w_t)$, the marginal value of cash around $w \sim 0$ and endogenous payout boundary slightly changes from that of Bolton, Chen, and Wang [4]. Although there are several small differences in the policy function of $i(w_t)$, dw_t and other related variables, most changes are minor ones and we do not emphasize these differences here.

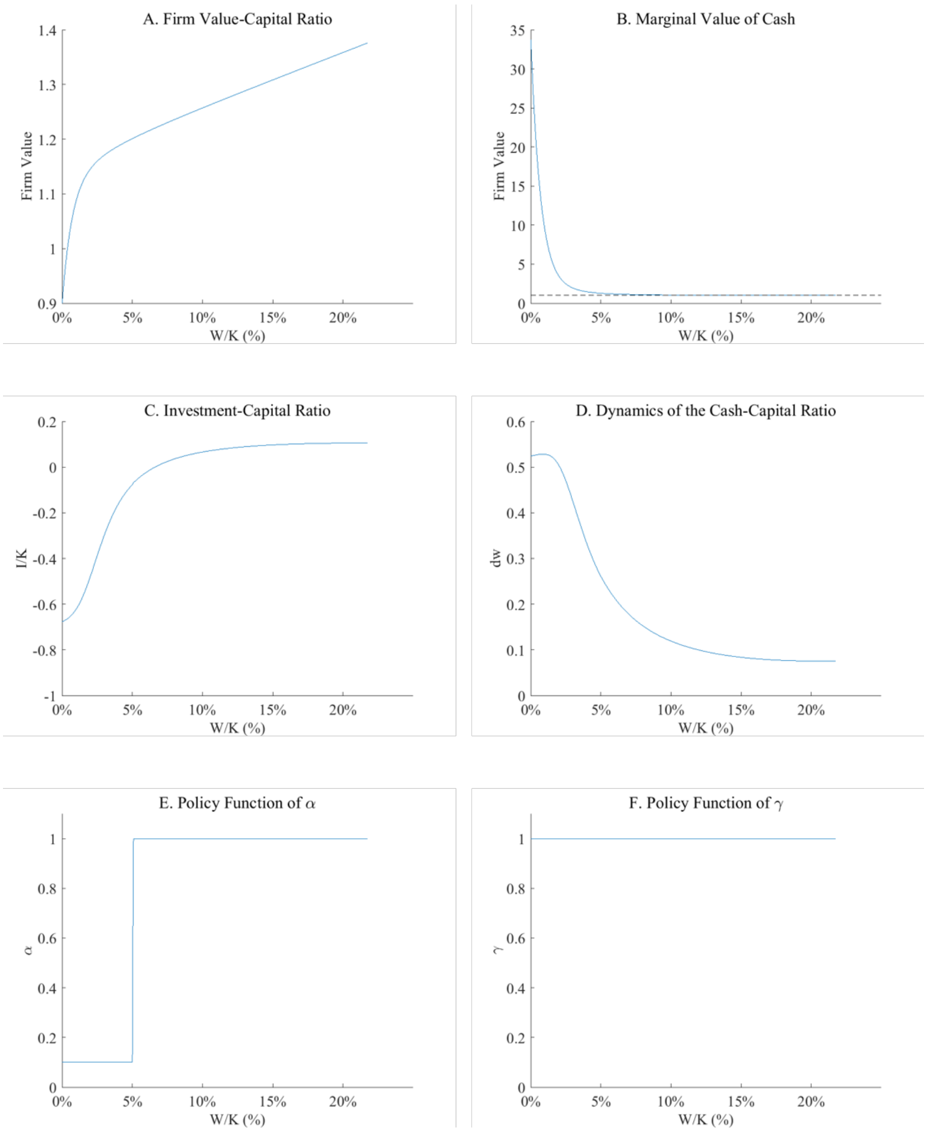

Figure 4.4: Firm Value and Related Variables under $\gamma_{min} = 1.0$

4.5.3.3 *Case of* $\gamma_{min} = \mathbf{0.90}$

Figure 4.4 plots simulated results of similar variables under $\gamma_{min} = 0.90$. First and foremost, let us discuss on the shape of the policy function of $\gamma(w_t)$. The shape of the policy function of $\gamma(w_t)$ follows 1) $\gamma(w_t)$ stays constant for almost all region at $\gamma(w_t) = \gamma_{min}$, but 2) suddenly shows an increase when $w_t \sim \overline{w}$. Here let us consider the intuitive understandings of these two features. To understand these features, it might be better to revisit (4.15) again. If we approximate $\frac{1}{1+\gamma(w_t)\left(1-\frac{1}{p}\right)(1-\alpha(w_t))} \sim 1$ for simplicity, the first order derivative for the right hand side of (4.15) becomes:

$$(\mu + rw_t - \delta - g(i_t)) \left\{ v^{(1)} - 1 - \left(1 - \frac{1}{p}\right)(1 - \alpha_t) \right\} + \gamma(w_t)\sigma^2 v^{(2)}. \tag{4.31}$$

Especially in the low w_t region, $v^{(2)}$ is around -10^4 and therefore the solution for the first order condition (4.31) is much smaller than γ_{min}. However, $v^{(2)}$ decreases as the increase in w_t and finally reaches 0 when $w_t \to \overline{w}$. Under such situation, the solution for the FOC (4.31) increases and when it exceeds γ_{min}, the policy function of $\gamma(w_t)$ becomes non-constant solution. The numerical analysis states that when $w_t \to \overline{w}$, $\gamma(w_t) = \gamma_{min}$ can no longer becomes the best solution.

Here it should also be noted that when γ_{min} decreases, 1) the marginal value of cash at $w_t \to 0$ increases, and 2) endogenous payout boundary $\overline{w}$ decreases, compared to the case of $\gamma_{min} = 1$.

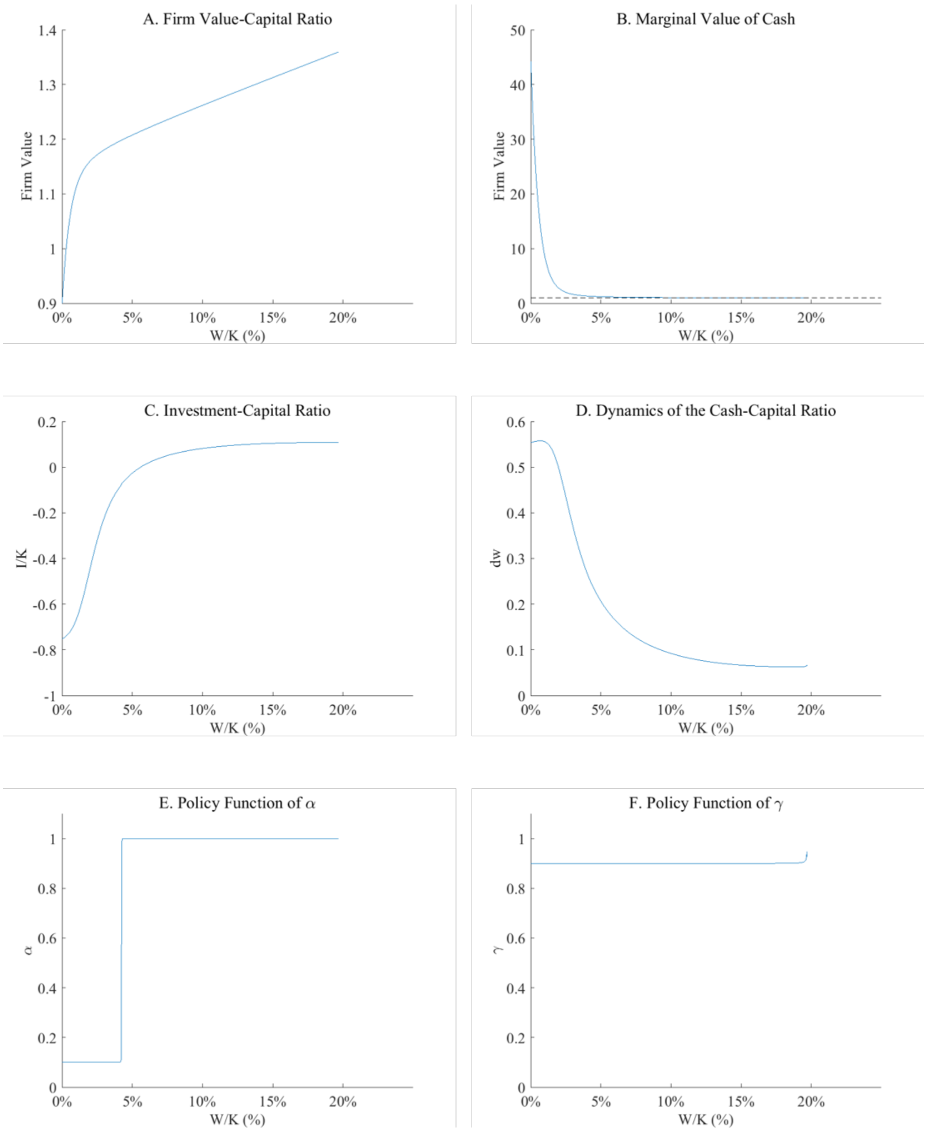

Figure 4.5: Firm Value and Related Variables under $\gamma_{min} = 0.9$

4.5.3.4 *Case of* γ_{min} = **0.70**

Next, Figure 4.6 plots similar functions under $\gamma_{min} = 0.70$. In general, most features are same as the case $\gamma_{min} = 0.90$. The major differences are 1) decrease in $\overline{w}$, 2) increase in $v^{(1)}(0)$, 3) decrease in $i(0)$ and 4) increase in $dw_t|_{w_t=0}$. These facts leads that if γ_{min} decreases, the firm tries to increase its cash-capital ratio more rapidly by increasing the rate of under-investment ($i(w_t) < 0$).

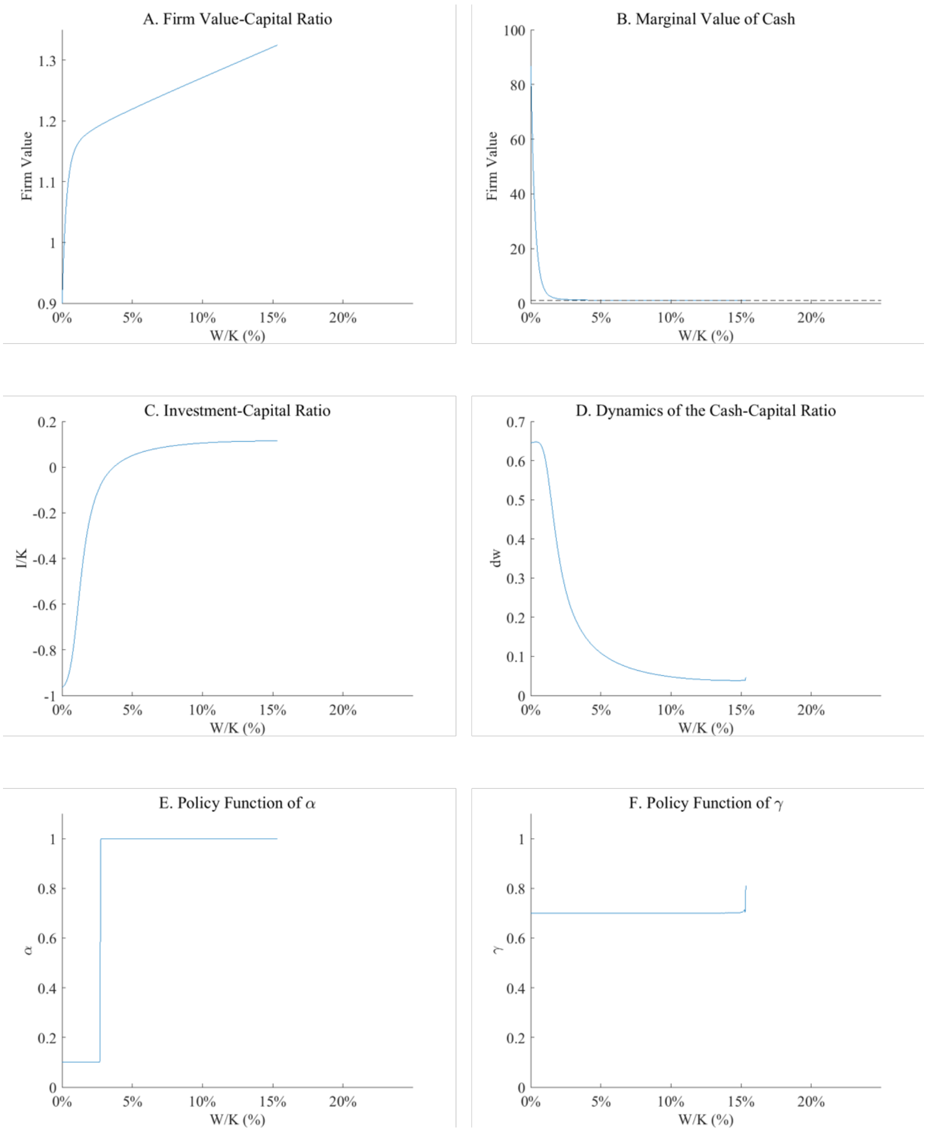

Figure 4.6: Firm Value and Related Variables under $\gamma_{min} = 0.7$

4.5.3.5 *Case of* γ_{min} = **0.50**

Lastly, Figure 4.7 plots similar functions under $\gamma_{min} = 0.50$. When γ_{min} decreases as 0.5, the investment-capital ratio at $w_t = 0$ decreases until i_{min} and other features are the same as other cases.

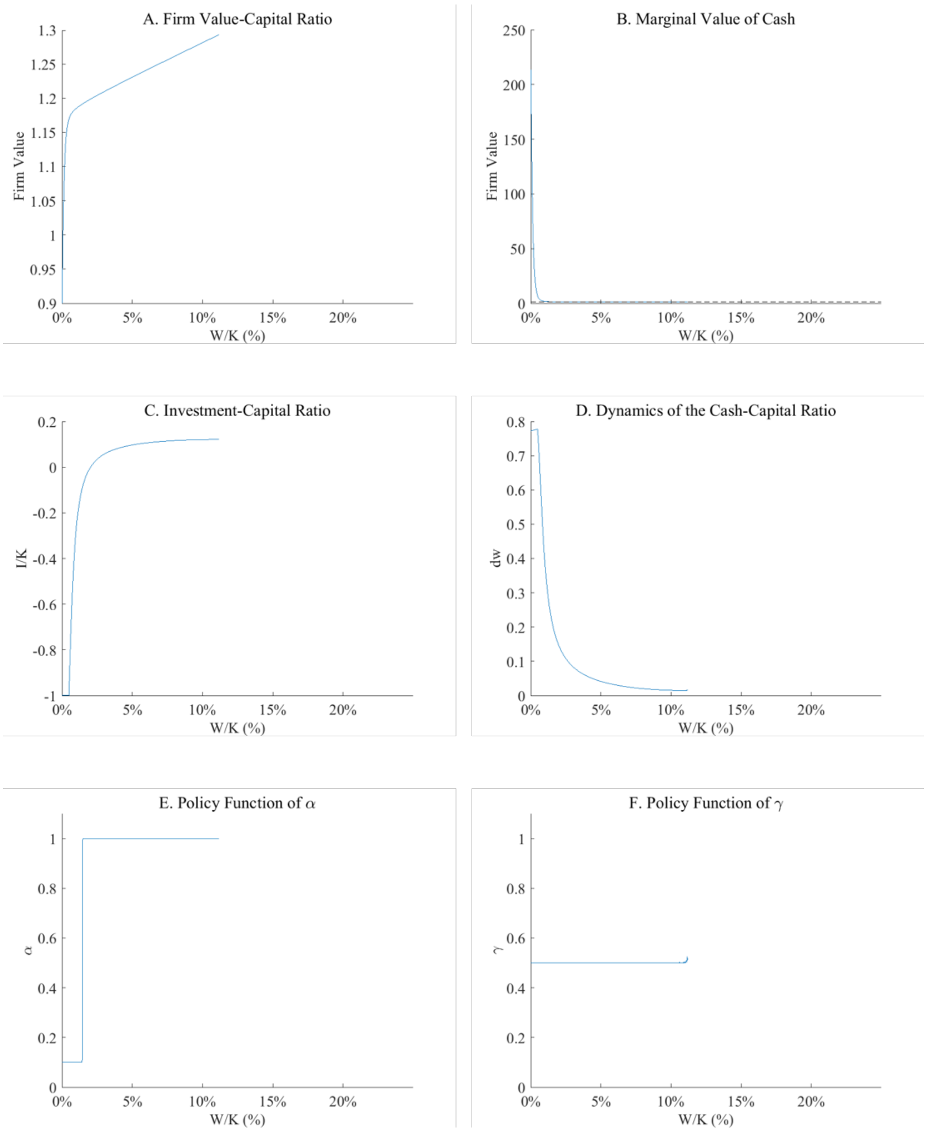

Figure 4.7: Firm Value and Related Variables under $\gamma_{min} = 0.5$

4.6 DISCUSSION

In this chapter we proposed a dynamic operational model which includes the retained earnings and cash holdings explicitly by considering the balance sheet of each firm. The first contribution of this chapter is to establish a methodology which can solve the HJB equation (4.15) under complicated first order equations (4.25) which cannot solve analytically. Also, the major findings of this analysis is that whether the firm accumulates the retained earnings or not is determined by the level of the cash-capital ratio w_t and a friction on reinvestments, γ_{min}.

First and foremost, the relation between γ_{min} and retained earnings should be discussed. As is shown in Figure 4.4 - 4.7, the policy function $\gamma(w_t)$ stays constant for most regions. Also, the decrease in γ_{min} leads to 1) decrease in $\overline{w}$, 2) increase in $v^{(1)}(0)$, 3) decrease in $i(0)$ and 4) increase in $dw_t|_{w_t=0}$. These changes indicates that the decrease in γ_{min} enables a firm to adjust more rapidly to increase its cash-capital ratio when its cash-capital ratio is low. As described above, the origin of γ_{min} is a friction on the re-investment. The increase in γ_{min} allows the firm to increase its re-investment and the firm does not have to pay all profits as a dividend. Here it should be emphasized that the adjustment process becomes rapid as γ_{min} decreases. Also, it might be reasonable to expect that the firm value becomes linear when $\gamma_{min} \to 0$. Actually, the HJB equation becomes equal to the neoclassical benchmark when $\gamma(w_t) = 0$. This fact leads that when the firm allows itself to pay much dividend without considering any reactions of investors, the firm value becomes close to the neoclassical benchmark. In other words, the existence of the friction of re-investments alienates firms from the benchmark.

Also, it can be considered that the firm's decision is also affected by the economic situation. According to the data in Japanese industry (Figure 4.2), 2 major findings were available: 1) γ_{min} tends to decrease in the recession (e.g., the financial crisis) and increase in the upturn, and 2) the small firms tend to retain more earnings. The first finding yields that the firm tries to convince investors for their growth potential especially when the economic situation is in upturn. When the economy is good, it is reasonable for the firm to claim that it has better growth opportunity than competitors, otherwise the firm loses the market share. On the other hand, in the bad economic situation, such claim may sound incredible for investors as many firms have already lost their opportunities.

The second finding is considered mainly due to the effect of the listing. In general, large firms tend to take itself public and finance from the market. On the other hand, small firms cannot take itself public. If the firm is not a listed company, a pressure from investors becomes moderate. When the pressure from investors is low, the firm may face two options: 1) to decrease dividend and retain more earnings, and 2) to decrease γ_{min} and operate the firm close to the neoclassical benchmark. The first option does not seem to maximize the firm value in most cases, but this situation may change if the firm is an owner managed firm. If the manager is also the owner of the firm, the manager can use its retained earnings at will. Instead, the manager may rather prefer to increase the retained earnings for tax reasons. If the owner managed firm pays dividend to the owner itself, the dividend becomes object of a taxation. On the other hand, if the owner managed firm retains its earnings, no (or less) taxation will occur. So, we can guess that one of the reason the ratio of retained earnings of small firms is large is the taxation.

Lastly, we see a few directions for related future research. One possible direction would be the assumption on the production function which is currently assumed as homogeneous of degree 1 and does not satisfy diminishing returns. If we allow the firm value to be the function of W_t and K_t and calculate the 2 variable Hamilton-Jacobi-Bellman equation, we can assume production function with diminishing returns and such expansion should be considered in the next work. Also, to introduce the debt into this model could be another direction for future research.

Part V

CONCLUSION

5 CONCLUSION

In this dissertation, we addressed the aggregation in macroeconomics especially with setting our focus on the statistical process.

With the results in part 1 and 2, we succeeded to clarify the microfoundation and assumptions to retrieve the representative agent and CES utility function which are frequently used in many macroeconomic literatures. In both researches, we provided sequences of assumptions which can be verified with empirical analysis. Here it should be emphasized that if succeeding research reveals all assumptions to be reasonable based on the actual market data, our work is meaningful in a aspect of providing rigorous microfoundation to the macroeconomic frameworks. However, if succeeding research reveals some assumptions are not reasonable and does not satisfied in usual economy, a severe problem arises. Such results may raise a warning for several macroeconomic literatures that the sticking to the analytic convenience might cause the following calculations in vain. At any rate, it is necessary for many macroeconomic literatures that whether such assumptions are all correct or not with using empirical data, and the significance of our research lies on it.

In the part 3 of this dissertation, we addressed the statistical process of the corporate finance, especially setting our focus on the retained earnings. The major reason for setting our focus to the retained earnings is the relation between the firm's profit and growth. If the market is completely perfect, the current profitability does not affect firm's financial needs as long as the bankers/equity providers can appropriately judge the financing program. However, the market is not always perfect and in such case, the firm's growth is restricted mainly to the internal financing, i.e., to use the retained earnings. Also, if we are to consider the long range growth, how much the firm should take its leverage is also a key variable. With keeping the profit rate constant, firstly the firm has to determine the leverage within capital stock, i.e., the ratio of increase in shareholder's equity and retained earnings. Next, the firm also has to determine the leverage between the debt and equity. Although we only considered the leverage within capital stock in the part 3 research, the basic concept can be applied to the debt and then we can evaluate the relation between firm's profitability and growth rate.

Lastly, it is also notable that if we combine all there 3 results, we can build a small economy which is consist of consumers (and investors) and firms with considering stochastic property of each agent. Such description helps us to analyze the actual distribution dynamics of consumer attribute, firm's productivity, etc. as it is, and such approach must be suitable to so called "big data" analysis. I hope our research will be of any help to evolve the macroeconomics more interesting and rigorous.

Part VI

APPENDIX

A

APPENDIX

A.1 DERIVATION OF THE FOKKER-PLANCK EQUATION

For the simplicity, we first assume the consumption bundles evolves with Markov process, and relax this assumption in later section.

Assumption 1. *Any consumption bundles within the household type μ_i evolves with Markov process, i.e.,*

$$\mathcal{P}(\Lambda \cap \Gamma | \mathbf{x_b}(t)) = \mathcal{P}(\Lambda | \mathbf{x_b}(t)) \mathcal{P}(\Gamma | \mathbf{x_b}(t)),$$
$$^{\forall}\Lambda \in \sigma\{\mathbf{x_b}(s),\, s \leq t\},\ ^{\forall}\Gamma \in \sigma\{\mathbf{x_b}(s),\, s \geq t\}$$

where $\mathcal{P}$ is a probability distribution function of the probability process $\mathbf{x_b}(t)$, and $\sigma\{\cdot\}$ is a minimum σ- additive class which makes the probability process written in the bracket measurable.

Hereinafter we write $\mathbf{x_b}(t)$ as $\mathbf{x_b}$ for simplicity. Under Markov process assumption, the dynamics of a probability density function is generally described by its autonomous differential equation as:

Proposition 1.
The probability density function $P_{\mu_i}(\mathbf{x_b},\, t)$ evolves with following time dependent differential equation

$$\frac{\partial}{\partial t} P_{\mu_i}(\mathbf{x_b}, t) = -P_{\mu_i}(\mathbf{x_b}, t)\xi(\mathbf{x_b}) + \int P_{\mu_i}(\mathbf{x'_b}, t) T(\mathbf{x'_b}, \mathbf{x_b}) d\mathbf{x'_b} \tag{A.1}$$

where $\xi(\mathbf{x_b})$ is a probability to transit from $\mathbf{x_b}$, and $T(\mathbf{x'_b}, \mathbf{x_b})$ is a probability to transit from $\mathbf{x_b}'$ to $\mathbf{x_b}$ during unit time scale.

Proof.
First define t' as $t_0 < t' < t$. If we assume the probability process of the $\mathbf{x_b}$ to be a Markov process, the following relation known as a Chapman-Kolmogorov equation is satisfied;

$$P_{\mu_i}(\mathbf{x_b}, t \mid \mathbf{x_{b0}}, t_0) = \int P_{\mu_i}(\mathbf{x_b}, t \mid \mathbf{x'_b}, t')\, P_{\mu_i}(\mathbf{x'_b}, t' \mid \mathbf{x_{b0}}, t_0)\, d\mathbf{x'_b} \tag{A.2}$$

Now we consider transition during an infinitely small amount of time Δt. As definition, a transition probability from $\mathbf{x'_b}$ to $\mathbf{x_b}$ during Δt is calculated as $T(\mathbf{x'_b}, \mathbf{x_b})\Delta t$. On the other hand, a probability to stay at $\mathbf{x_b}$ during Δt is calculated as $\left\{1 - \int T(\mathbf{x_b}, \mathbf{x''_b})\, d\mathbf{x''_b}\right\} \Delta t$. Therefore, a time evolution of the probability density function during Δt becomes:

$$P_{\mu_i}(\mathbf{x_b}, t+\Delta t \mid \mathbf{x_{b0}}, t_0) = \int \left[\left\{ 1 - \int T(\mathbf{x_b}, \mathbf{x''_b})\, d\mathbf{x''_b} \right\} \Delta t\, \delta(\mathbf{x_b} - \mathbf{x'_b}) + T(\mathbf{x'_b}, \mathbf{x_b})\Delta t \right] P_{\mu_i}(\mathbf{x'_b}, t \mid \mathbf{x_{b0}}, t_0)\, d\mathbf{x'_b} \tag{A.3}$$

Taking limit of $\Delta t \to 0$ to obtain

$$\frac{\partial}{\partial t} P_{\mu_i}(\mathbf{x_b}, t \mid \mathbf{x_{b0}}, t_0) = -P_{\mu_i}(\mathbf{x_b}, t \mid \mathbf{x_{b0}}, t_0)\xi(\mathbf{x_b}) + \int P_{\mu_i}(\mathbf{x'_b}, t \mid \mathbf{x_{b0}}, t_0) T(\mathbf{x'_b}, \mathbf{x_b}) d\mathbf{x'_b} \tag{A.4}$$

where $\xi(\mathbf{x_b}) \equiv \int T(\mathbf{x_b}, \mathbf{x''_b})\, d\mathbf{x''_b}$. This is a similar expression as written in the *Proposition 1*. (For simplicity, hereinafter $P_{\mu_i}(\mathbf{x_b}, t \mid \mathbf{x_{b0}}, t_0)$ is written as $P_{\mu_i}(\mathbf{x_b}, t)$.) ■

The intuitive understanding of this equation is very simple. The left hand side of the equation equals to the time differential of the probability to stay at $\mathbf{x_b}$. Meanwhile, the right hand side of the equation equals to the sum of 2 components; (1) transition from $\mathbf{x_b}$, and (2) transition from other points ($\mathbf{x'_b}$) to $\mathbf{x_b}$.

Define $\mathbf{r}$ as $\mathbf{r} = \mathbf{x'_b} - \mathbf{x_b}$ and $\omega(\mathbf{x'_b}, \mathbf{r}) = T(\mathbf{x'_b}, \mathbf{x})$. Substitute this expression into equation (A.4) to obtain

$$\frac{\partial}{\partial t} P_{\mu_i}(\mathbf{x_b}, t) = -\int \omega(\mathbf{x_b}, \mathbf{r})\, d\mathbf{r}\, P_{\mu_i}(\mathbf{x_b}, t) + \int \omega(\mathbf{x_b} - \mathbf{r}, \mathbf{r})\, d\mathbf{r}\, P_{\mu_i}(\mathbf{x_b} - \mathbf{r}, t) \tag{A.5}$$

Here we introduce following formula for the further calculation.

$$\exp\left[-\mathbf{r} \frac{\partial}{\partial \mathbf{x_b}} \right] f(\mathbf{x_b}) = \sum_{n=0}^{\infty} \frac{(-\mathbf{r})^n}{n!} \left(\frac{\partial}{\partial \mathbf{x_b}} \right)^n f(\mathbf{x_b}) = f(\mathbf{x_b} - \mathbf{r}) \tag{A.6}$$

Substitute this formula into the previous equation, and also assume appropriate convergence condition, calculation be proceeded as follows;

$$\begin{aligned} \frac{\partial}{\partial t} P_{\mu_i}(\mathbf{x_b}, t) &= -\int \omega(\mathbf{x_b}, \mathbf{r})\, d\mathbf{r}\, P_{\mu_i}(\mathbf{x_b}, t) \\ &\quad + \int d\mathbf{r} \sum_{n=0}^{\infty} \frac{(-\mathbf{r})^n}{n!} \left(\frac{\partial}{\partial \mathbf{x_b}} \right)^n \omega(\mathbf{x_b}, \mathbf{r})\, P_{\mu_i}(\mathbf{x_b}, t) \qquad \text{(A.7)} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(\frac{\partial}{\partial \mathbf{x_b}} \right)^n \int d\mathbf{r}\, \mathbf{r}^n \omega(\mathbf{x_b}, \mathbf{r})\, P_{\mu_i}(\mathbf{x_b}, t) \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(\frac{\partial}{\partial \mathbf{x_b}} \right)^n \alpha_{\mathbf{n}}(\mathbf{x_b})\, P_{\mu_i}(\mathbf{x_b}, t) \qquad \text{(A.8)} \end{aligned}$$

where $\alpha_n(\mathbf{x_b}) = \int d\mathbf{r}\, \mathbf{r}^n \omega(\mathbf{x_b}, \mathbf{r})$.

If we confine our attention to the dynamics around the equilibrium point, it may be reasonable to ignore higher-order terms of the Taylor expansion. More rigorously, we set the following **Assumption 2** for the better approximation:

Assumption 2. The coefficients $\alpha_n(\mathbf{x_b})$ are finite for every n and $\alpha_n(\mathbf{x_b}) = 0$ for some even n

Proposition 2.
The time dependent partial differential equation on $P_{\mu_i}(\mathbf{x_b}, t)$ could simply be approximated under Assumption 4 as:

$$\frac{\partial}{\partial t} P_{\mu_i} = \left(-\sum_{j=1}^{L} \frac{\partial}{\partial x_j} \alpha_{1j}(\mathbf{x_b}) + \frac{1}{2} \sum_{j=1}^{L} \sum_{k=1}^{L} \frac{\partial^2}{\partial x_j \partial x_k} \alpha_{2jk}(\mathbf{x_b}) \right) P_{\mu_i}(x_1, x_2, ..., x_L, t) \tag{A.9}$$

Proof.

According to the Pawula theorem (Pawula [44]), if the coefficients $\alpha_n(\mathbf{x_b})$ are finite for every n and if $\alpha_n(\mathbf{x_b}) = 0$ for some even n, $\alpha_n(\mathbf{x_b}) = 0$ for all $n \geq 3$. If we employ this assumption, equation (A.8) could be approximated simply as written in the **Proposition 2**. ■

This equation is known as Fokker-Planck equation which describes the dynamics of probability density function under certain potential. In the literature of the physics, the first term of the RHS equals to a first order derivative of the external potential, and the second term equals to the effect of diffusion.

A.2 DERIVATION OF THE KRAMERS EQUATION

Proposition 3.
The dynamics of probability density function under an external potential and a friction could be described in the form (Kramers equation):

$$\frac{\partial}{\partial t} P_{\mu_i}(\mathbf{x_b}, \mathbf{v_b}, t) = \left[-\sum_{j=1}^{L-1} \frac{\partial}{\partial x_{bj}} v_{bj} + \sum_{j=1}^{L-1} \frac{\partial}{\partial v_{bj}} \left(-\frac{\partial}{\partial x_{bj}} u_R(x) + \gamma v_{bj} \right) + D \sum_{j=1}^{L-1} \frac{\partial^2}{\partial v_{bj}^2} \right] P_{\mu_i}(\mathbf{x_b}, \mathbf{v_b}, t) \tag{A.10}$$

Proof.
The derivation of Fokker-Planck equation (A.9) is conducted in general coordinate, and this relation satisfies even if we expand the coordinate to the phase space made of $\mathbf{x_b}$ and $\mathbf{v_b}$ (a time derivative of $\mathbf{x_b}$). Here first define A as $A^t \equiv (\mathbf{x_b}, \mathbf{v_b})$ and rewrite Fokker-Planck equation within this phase space as:

$$\frac{\partial}{\partial t} P_{\mu_i}(\mathbf{x_b}, \mathbf{v_b}, t) = \left(-\sum_{j=1}^{2(L-1)} \frac{\partial}{\partial A_j} \alpha_{1j}(\mathbf{A}) + \frac{1}{2} \sum_{j=1}^{2(L-1)} \sum_{k=1}^{2(L-1)} \frac{\partial^2}{\partial A_j \partial A_k} \alpha_{2jk}(\mathbf{A}) \right) P_{\mu_i}(\mathbf{x_b}, \mathbf{v_b}, t) \tag{A.11}$$

Here $\alpha_{1j}(\mathbf{A}) = \lim_{\Delta t \to 0} \frac{<\Delta x_{bj}>}{\Delta t} = v_{bj}$ for $j \leq L-1$, $\alpha_{1j}(\mathbf{A}) = \lim_{\Delta t \to 0} \frac{<\Delta v_{bj}>}{\Delta t} = \frac{\partial}{\partial x_{bj}} u_R(x) - \gamma v_{bj}$ for $L \leq j \leq 2(L-1)$, $\alpha_{2jk}(\mathbf{A}) = \lim_{\Delta t \to 0} \frac{<\Delta v_j \Delta v_k>}{\Delta t} = 2D\delta_{jk}$, and all other terms vanishes at order of Δt. Thus, the generalized diffusion process satisfies

$$\frac{\partial}{\partial t} P_{\mu_i}(\mathbf{x_b}, \mathbf{v_b}, t) = \left[-\sum_{j=1}^{L-1} \frac{\partial}{\partial x_{bj}} v_{bj} + \sum_{j=1}^{L-1} \frac{\partial}{\partial v_{bj}} \left(-\frac{\partial}{\partial x_{bj}} u_R(x) + \gamma v_{bj} \right) + D \sum_{j=1}^{L-1} \frac{\partial^2}{\partial v_{bj}^2} \right] P_{\mu_i}(\mathbf{x_b}, \mathbf{v_b}, t) \tag{A.12}$$

■

Proposition 4. *The solution of the Kramers equation in equilibrium becomes*

$$P_{\mu_i}^{eq} = \frac{1}{Z} \exp\left[-\frac{H(\mathbf{x_b}, \mathbf{v_b})}{\gamma D} \right], \quad H(\mathbf{x_b}, \mathbf{v_b}) \equiv \sum_{i=1}^{L-1} \frac{v_{bi}}{2} - u_R(\mathbf{x_b})$$

where Z is a normalization factor.

This solution can be easily confirmed by setting LHS of (A.12) as zero and plugging this solution and into the RHS of (A.12).

A.3 DERIVATION OF THE LIOUVILLE EQUATION

First define phase space constructed by $L-1$ generalized coordinate $(x_{1b}, x_{2b}, \cdots, x_{L-1b})$ and its conjugate momentum $(v_{1b}, v_{2b}, \cdots, v_{L-1b})$ (here the mass of the particle is normalized as one), and consider the dynamics in this phase space. Every state realized in this system within the phase space is generally called as representative point. Here define $\rho = \rho(\mathbf{x_b}, \mathbf{v_b}, t)$ as a density of the representative point in $L-1 \times L-1$ dimensional phase space at time t. Then, there are $\rho(\mathbf{x_b}, \mathbf{v_b}, t)\Delta x_{1b}\Delta x_{2b} \cdots \Delta x_{L-1b} \cdot \Delta v_{1b}\Delta v_{2b} \cdots \Delta v_{L-1b}$ representative points in the infinitely small volume as $\Delta x_{1b}\Delta x_{2b} \cdots \Delta x_{L-1b} \cdot \Delta v_{1b}\Delta v_{2b} \cdots \Delta v_{L-1b}$. Now let us consider the dynamics of the representative points in this phase space. The number of representative points which pass through a surface of $x = x_{1b}$ equals to

$$\rho(\mathbf{x_b}, \mathbf{v_b}, t)\dot{x_{b1}}\Delta V_{-x_{1b}} \tag{A.13}$$

where $\Delta V_{-x_{1b}} = \Delta x_{2b} \cdots \Delta x_{L-1b} \cdot \Delta v_{1b}\Delta v_{2b} \cdots \Delta v_{L-1b}$. Similarly, the number of representative points which pass through a surface of $x = x_{1b} + \Delta x_{1b}$ equals to

$$\{\rho(\mathbf{x_b}, \mathbf{v_b}, t)\dot{x_{b1}}\}\,|_{x_{1b}+\Delta x_{1b}} \Delta V_{-x_{1b}} = \left(\rho(\mathbf{x_b}, \mathbf{v_b}, t)\dot{x_{b1}} + \frac{\partial \rho(\mathbf{x_b}, \mathbf{v_b}, t)\dot{x_{b1}}}{\partial x_{b1}} \Delta x_{1b}\right) \Delta V_{-x_{1b}} \tag{A.14}$$

Therefore, the number of representative points in $\Delta x_{1b} \Delta x_{2b} \cdots \Delta x_{L-1b} \cdot \Delta v_{1b} \Delta v_{2b} \cdots \Delta v_{L-1b}$ decreases in every unit time as:

$$\frac{\partial \rho(\mathbf{x_b}, \mathbf{v_b}, t)\dot{x_{b1}}}{\partial x_{b1}} \Delta x_{1b} \Delta x_{2b} \cdots \Delta x_{L-1b} \cdot \Delta v_{1b} \Delta v_{2b} \cdots \Delta v_{L-1b} \tag{A.15}$$

The similar discussion could also be applied to other surfaces and as a result, the time differential equation of the density of representative points could be described as:

$$\frac{\partial}{\partial t}\rho = i\mathscr{L}\rho \equiv \{\rho, H\} = \sum_{j=1}^{L-1} \left(\frac{\partial \rho}{\partial x_{bj}} \frac{\partial H}{\partial v_{bj}} - \frac{\partial \rho}{\partial v_{bj}} \frac{\partial H}{\partial x_{bj}} \right) \tag{A.16}$$

A.4 DERIVATION OF THE GENERALIZED LANGEVIN EQUATION

Proposition 5. *The equation of motion of each consumption bundle could be described in the form of the generalized Langevin equation as follows:*

$$\frac{d}{dt}\mathbf{v_b}(t) = -\int_{-\infty}^{t} \gamma(t-t')\mathbf{v_b}(t')dt' + \nabla u_R(\mathbf{x_b}) + \mathbf{R}(t) \tag{A.17}$$

where $\gamma(t)$ represents a retarded effect of the frictional force at time t, and $\mathbf{R}(t)$ *is a random force.*

Proof.

By expanding the dimension of the phase space to the Hilbert space, the Liouville equation can be written as

$$\frac{d}{dt}A_\mu(t) = i\mathscr{L}A_\mu(t) \tag{A.18}$$

and formally be solved in the form $A_\mu(t) = exp(i\mathscr{L}t)A_\mu(0)$. Next, we define a inner product of dynamical values F ,G with requiring the following restrictions:

$$(F, G) = (G, F)^*, \tag{A.19}$$

$$(G, G) \geq 0, \tag{A.20}$$

$$(\sum_i c_i F, G) = \sum_i c_i (F, G) \tag{A.21}$$

For the simplicity, we assume the orthogonality and normalization for $\{A_\mu(0)\}$ as

$$\big(A_\mu(0), A_\nu(0)\big) = \delta_{\mu\nu} \tag{A.22}$$

where $\delta_{\mu\nu}$ is Dirac's delta function. Here define the projection operator $\mathscr{P}$ which project dynamical value G to the space mapped by $\{A_\mu(0)\}$ as

$$\mathscr{P}G(t) = \sum_\nu (G(t), A_\nu(0)) A_\nu(0) \tag{A.23}$$

The following relation about the projection operator can be easily proved.

$$(\mathscr{P}F, G) = (F, \mathscr{P}G) \tag{A.24}$$

$$\big(\mathscr{P}'F, G\big) = \big(F, \mathscr{P}'G\big) \tag{A.25}$$

$$\mathscr{P}^2 = \mathscr{P},\ \mathscr{P}'^2 = \mathscr{P}',\ \mathscr{P}\mathscr{P}' = \mathscr{P}'\mathscr{P} = 0 \tag{A.26}$$

where $\mathscr{P}' = 1 - \mathscr{P}$. Now we define $\Xi_{\mu\nu}(t)$ as

$$\Xi_{\mu\nu}(t) = \big(A_\mu(t), A_\nu(0)\big) \tag{A.27}$$

Then the projection of $A_\mu(t)$ to A is given by

$$\mathscr{P}A_\mu(t) = \sum_\nu \Xi_{\mu\nu}(t) A_\nu(0) \tag{A.28}$$

On the other hand, we define $A'_\mu(t)$ as

$$A'_\mu(t) = \mathscr{P}' A_\mu(t) \tag{A.29}$$

Using (A.15), (A.16) and the definition of $\mathscr{P}'$, $A_\mu(t)$ can be rewritten in the form

$$A_\mu(t) = \sum_\nu \Xi_{\mu\nu}(t) A_\nu(0) + A'_\mu(t) \tag{A.30}$$

Operating $\mathscr{P}'$ to the Liouville equation from the left and using (A.16) and (A.17) to obtain

$$\frac{d}{dt}A'_\mu(t) = \mathscr{P}' i\mathscr{L} A'_\mu(t) + \sum_\nu \Xi_{\mu\nu}(t)\mathscr{P}' i\mathscr{L} A_\nu(0) \tag{A.31}$$

The solution of (A.18) becomes

$$A'_\mu(t) = \sum_\nu \int_0^t \Xi_{\mu\nu}(s)\Gamma_\nu(t-s)ds \tag{A.32}$$

$$\Gamma_\nu(t) = exp\left[t\mathscr{P}' iL\right] \mathscr{P}' iLA_\nu(0) \tag{A.33}$$

By substituting (A.17) into the LHS of (A.5), we obtain

$$\frac{d}{dt}A_\mu(t) = \sum_\nu \Xi_{\mu\nu}(t) iLA_\nu(0) + iLA'_\mu(t) \tag{A.34}$$

Here we define $i\Omega_{\mu\nu}$ and $M_{\mu\nu}(t)$ as

$$i\Omega_{\mu\nu} \equiv (iLA_\mu(0), A_\nu(0)) \tag{A.35}$$

$$M_{\mu\nu}(t) \equiv -(iL\Gamma_\mu(0), A_\nu(0)) \tag{A.36}$$

Taking an inner product with $A_\nu(0)$ on (A.21) from the right and using (A.19) and (A.20) to obtain a differential equation about $\Xi_{\mu\nu}(t)$ as

$$\frac{d}{dt}\Xi_{\mu\nu}(t) = \sum_\tau \Xi_{\mu\tau}(t) i\Omega_{\tau\nu} - \sum_\tau \int_0^t \Xi_{\mu\tau}(s) M_{\tau\nu}(t-s)ds \tag{A.37}$$

Representing in a form of a matrix,

$$\frac{d}{dt}\Xi(t) = \Xi(t)\cdot i\Omega - \int_0^t \Xi(s)\cdot M(t-s)ds \tag{A.38}$$

The Laplace transformation of (A.25) becomes

$$-\hat{1} + z\Xi(z) = \Xi(z)\cdot i\Omega - \Xi(z)\cdot M(z) \tag{A.39}$$

Therefore

$$\Xi(z) = \frac{\hat{1}}{z - i\Omega + M(z)} \tag{A.40}$$

On the other hand, if we substitute (A.19) into (A.17) and conduct Laplace transformation to obtain

$$A(z) = \Xi(z)\cdot\{A(0) + \Gamma(z)\} \tag{A.41}$$

Using (A.27) and (A.28),

$$\{z - i\Omega + M(z)\} \cdot A(z) = A(0) + \Gamma(z) \tag{A.42}$$

By applying the inverse Laplace transformation on (A.29) and rewrite the equation in the form of generalized coordinate, we obtain the generalized Langevin equation as:

$$\frac{d}{dt}A_\mu(t) = \sum_\nu i\Omega_{\mu\nu}A_\nu(t) - \sum_\nu \int_0^t M_{\mu\nu}A_\nu(t-s)ds + \Gamma_\mu(t) \tag{A.43}$$

■

Here the terms except for $\Gamma_\mu(t)$ are linear with respect to $A_\mu(t)$ and all non-linear effects are re-normalized into a fluctuating term as $\Gamma_\mu(t)$. The second term at the right hand of the equation represents a "memory" of the past movement and the function $M_{\mu\nu}(t)$ are called as memory function. Moreover, the fluctuation dissipation theorem of the second kind holds between memory function and fluctuating force.

A.5 DIFFERENCES IN SEVERAL PROCESSES

In most cases, the random force $\mathbf{R}(t)$ is assumed to be independent and identically distributed. However, our fundamental assumption (**Assumption 2**. in chapter 2) abandoned to distinguish individual households to describe representative households, and setting i.i.d. assumption in addition to this assumption may become too strong to describe real economy. Therefore, we first set assumptions to achieve i.i.d. property for $\mathbf{R}(t)$, and thereafter relax each assumptions in the following sections.

A.5.0.1 *Assumptions for the Normal Diffusion*

For the sake of simplicity and idealization, let us first assume the simple constraints for the property of the error term $\mathbf{R}(t)$ to acquire i.i.d. property as follows:

Assumption ND − 1. There are no auto correlation function in $\mathbf{R}(t)$, i.e.,

$$< \mathbf{R}(t_1)\mathbf{R}(t_2) >= 2\pi\mathbf{G_R}\delta(t_1 - t_2) \tag{A.44}$$

where $\mathbf{G_R}$ is the constant and $\delta(t_1 - t_2)$ is Dirac's delta function

Assumption ND − 2. The process $\mathbf{R}(t)$ is a Gaussian process

Both assumptions are concerning the randomness of the error term. The first assumption is especially concerning the friction term of the generalized Langevin equation. The friction term of the generalized

Langevin equation is $\int_{-\infty}^{t} \gamma(t-t')u(t')dt'$ and the assumption is to set $\gamma(t-t')$ as a Dirac's delta function. The meaning of this assumption is to ignore the "memory effect" of the particle. When we ignore the memory effect of the particle, its dynamics follows Markov processes and no need to preserve "memory" of the previous process as long as we analyze dynamics in $(\mathbf{x_b}, \mathbf{v_b})$ space. In general, the memory less property of the Markov process is described as:

$$\mathcal{P}(T > t + s | T > s) = \mathcal{P}(T > t)$$

The second assumption especially sets its focus on the characteristics of dynamics. Let us visit the fundamental motivation to use this relation into physics. In the model of physics, this Gaussian assumption becomes a better approximation for a particle which has a much larger mass than colliding molecules. This is because the Brownian motion is achieved as a consequence of a great number of collisions which satisfies a condition for the central limit theorem, with many small particles. The Gaussian assumption also sets a restriction on the form of the mean square displacement as:

$$< (\Delta x)^2 > \propto D\Delta t, \quad t \to \infty \tag{A.45}$$

where $< (\Delta x)^2 >$ is the mean square displacement, D is a diffusion constant and Δt is time for the displacement.

As a result of these two assumptions, the generalized Langevin equation could be rewritten simply as:

$$\frac{d}{dt}\mathbf{v_b}(t) = -\gamma \mathbf{v_b}(t) + \nabla u_R(\mathbf{x_b}) + \mathbf{R_w}(t) \tag{A.46}$$

with using standard white noise which satisfy $\mathbf{Assumption\ ND} - \mathbf{1}$. and $\mathbf{Assumption\ ND} - \mathbf{2}$. written as $\mathbf{R_w}(t)$.

A.5.0.2 *Assumption for Non-Markov Process*

As previously indicated, there are two ways to relax approximations employed to derive normal diffusion model.

Here we first relax $\mathbf{Assumption\ ND} - \mathbf{1}$., an assumption of Markov Process, and obtain a description of generalized Langevin equation with white noise term. As the generalized Langevin equation still holds in the Non-Markov process, the generalized Langevin equation with white noise term could be led in the form,

$$\frac{d}{dt}\mathbf{v_b}(t) = -\int_{-\infty}^{t} \gamma(t-t')\mathbf{v_b}(t')dt' + \nabla u_R(\mathbf{x_b}) + \mathbf{R}(t) \tag{A.47}$$

This kind of equation of motion will be realized when we consider some kind of auto regressive (AR) processes as a probability process. If we just assume that the value $\mathbf{v_b}(t)$ follows standard Orshtein-Uhlenbeck process with unit auto regressive term like

$$d\mathbf{v_b}(t) = \mu \mathbf{v_b}(t) \cdot dt + \nabla u_R(\mathbf{x_b}(t)) \cdot dt + \mathbf{R}(t) \tag{A.48}$$

the standard Langevin equation introduced in (A.47) will be derived.

Therefore, if we would like to derive the generalized Langevin equation written in (A.48), we need to assume the auto regressive term to satisfy $\int_{-\infty}^{t} \gamma(t-t')\mathbf{v_b}(t')dt'$ except for the standard AR(1) process as $\mu \mathbf{v_b}(t) \cdot dt$.

When we calculate and discuss this equation, the shape of the retarded effect of the frictional force becomes a problem to be solved. The simplest form for the retarded friction force function is to assume exponential decay for the past memories. In general, the Mittag-Leffler function is used as a generalization of exponential function. The so-called three-parameter Mittag-Leffler function introduced by Prabhakar [45] is described as:

$$E^{\rho}_{\mu,\nu}(t-t') = \sum_{k=0}^{\infty} \frac{(\rho)_k}{\Gamma(\mu k + \nu)} \frac{(t-t')^k}{k!}, \tag{A.49}$$

with $Re(\mu) > 0$, $Re(\nu) > 0$ and $Re(\rho) > 0$ and $z \in \boldsymbol{C}$. When the parameter $\rho = 1$, the equation becomes the two parameter Mittag-Leffler function as described in Agarwal [1]. Further, if we assume $\rho = \nu = 1$, the original Mittag-Leffler function as described in Mittag-Leffler (Mittag-Leffler [40]) is obtained. Also, if we assume $\rho = \nu = \mu = 1$, the standard exponential function is obtained.

The economic meaning of this relation is, as already mentioned, the description of an effect of a relaxation against drastic transformation of the exterior environment. If $\mathbf{Assumption\, ND - 1}$. does not hold in micro data, the retarded effect of the frictional force no longer be the Dirac's delta function, and the relaxation process could no longer be well approximated by an exponential function.

A.5.0.3 *Assumption for Anomalous Diffusion*

Lastly let us now consider relaxation of the assumption on Gaussian process ($\mathbf{Assumption\, ND - 2}$.). The anomalous diffusion is found in many physical systems. The anomalous diffusion shows the non-linear growth of the mean square displacement in the course of time. In the anomalous diffusion case, it is known that the mean square displacement can be described as a function of time as:

Assumption AD − 1. *The auto correlation function of* $\mathbf{R}(t)$ *has the form*

$$< \mathbf{R}(t_1)\mathbf{R}(t_2) >= 2\pi \mathbf{G_R} \delta(t_1 - t_2) \tag{A.50}$$

$$< (\Delta x)^2 > \propto 2D\Delta t^{\alpha}, \quad t \to \infty \tag{A.51}$$

where $\alpha \neq 1$. Firstly, let us consider the meaning of this assumption in terms of the probability process. This assumption will be reasonable if we consider changing the distribution of the step-length. For example, if we assume that the distribution of the step-length follows Pareto distribution, this probability process becomes Lévy flight, and the index $\alpha > 1$ in general.

According to Jesperson et. al. [25],

... sub-diffusive transport ($0 < \alpha < 1$) *is encountered in a diversity of systems, including the charge carrier transport in amorphous semiconductors, NMR diffusometry on percolation structure, and the motion of a bead in a polymer network.*
(snip)
Examples of enhanced diffusion ($\alpha > 1$) *include tracer particles in vortex arrays in a rotating flow, layered velocity field, and Richardson diffusion.*

A.6 PARAMETER SPACE IN LÉVY INDEX

Using the characteristic function of a distribution,

$$\phi(z) \equiv < e^{iXz} > = \int_{-\infty}^{+\infty} e^{iXz} dR(X) \tag{A.52}$$

the relation (A.51) is transferred into

$$\phi^n(z) = \phi(c_n z) \cdot e^{i\epsilon_n z} \tag{A.53}$$

Proposition 6. *(A.52) can be solved analytically and the solution becomes:*

$$\psi(x,t) \equiv log\, \phi(z) = i\mu z - \nu |z|^{\alpha} \left\{ 1 + i\beta \frac{z}{|z|} \omega(z, \alpha) \right\} \tag{A.54}$$

where α, β, μ, ν, *are constants (*μ *is any real number,* $0 < \alpha \leq 2$, $-1 < \beta < 1$, *and* $\nu > 0$ *), and*

$$\omega(z, \alpha) = \left[\begin{array}{ll} tan \frac{\pi\alpha}{2} & if\ \alpha \neq 1 \\ \frac{2}{\pi} log|z| & if\ \alpha = 1 \end{array} \right. \tag{A.55}$$

The proof of this proposition could easily be done by just plugging (A.53) and (A.54) into (A.52).

α is called as a Lévy index. If we assume the Lévy index as $\alpha = 2$, the distribution becomes the Gaussian normal distribution. Also if we assume $\beta = 0$, the distribution becomes symmetric. μ is the parameter which translates the distribution, and ν is a parameter regarding the scaling of X.

Proposition 7. *The equations (A.55) and (A.56) can be calculated as:*

$$\psi(x,t) = -|z|^{\alpha} exp \left\{ i\frac{\pi\beta}{2} sign(z) \right\} \tag{A.56}$$

with the constant parameter β. The range of β is as followings:

$$|\beta| \leq \left[\begin{array}{cc} \alpha & if\ 0 < \alpha \leq 1 \\ 2-\alpha & if\ 1 < \alpha \leq 2 \end{array} \right. \tag{A.57}$$

where $sign(z) \equiv \frac{z}{|z|}$ represents sign of z.

The parameter space for (α, β) is known as the Takayasu Diamond, which is described in Figure A.1. In general, this diamond diagram is based on the (A.57), and therefore as long as the set of parameters (α, β) is on or within the diamond, the distribution becomes stable. In addition, the diamond diagram includes 2 more information. The first is the bold line denoted as "OS" on the diamond. When the set of parameters is on this line, the distribution follows one-sided stabile law. The second is the letters written on the α-axis. When the set of parameters is on 1) "N", the distribution follows the normal or Gaussian law, 2) "H", the distribution follows the Holtsmark distribution 3) "C", the distribution follows the Cauchy or Lorentz distribution and 4) "L", the distribution follows the approximate log-normal distribution.

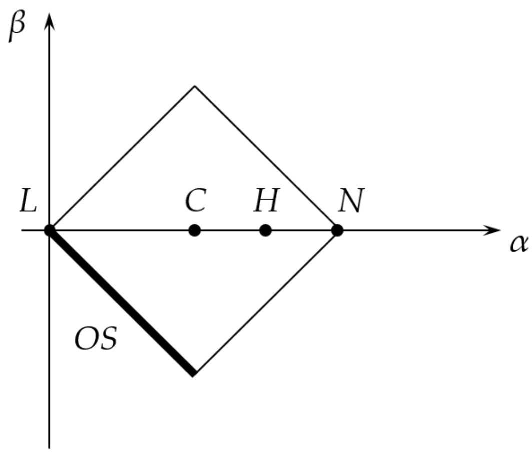

Figure A.1: Parameter Space for Stable Laws

A.7 FUNCTIONAL TYPE OF THE DISTRIBUTION FUNCTION

In the chapter 3, the functional type of the distribution function $\varphi(u_1, u_2)$ was assumed to be:

$$\varphi(u_{A1}, u_{A2}) = \begin{cases} A\,(u_{A1}+d)^{-\alpha}\,(u_{A2}-\delta_2)^{\beta_1} \\ \quad (u_{A1} > \delta_1,\ \delta_2 < u_{A2} < u_{A1}-\delta_{12},\ \alpha > 0,\ \beta_1 > 0) \\ A\,(u_{A2}+d+\delta_{12})^{-\alpha}\,(u_{A1}-\delta_1)^{\beta_2} \\ \quad (u_{A2} > \delta_2,\ \delta_1 < u_{A1} > u_{A2}-\delta_{21},\ \alpha > 0,\ \beta_2 > 0) \end{cases} \tag{A.58}$$

However, this functional type might be assumed as arbitrary and set only to retrieve the CES function. Therefore, in this appendix section, we assume other functional type and assume the robustness of the derivation of the CES.

A.7.1 *Perfectly Symmetric Function*

In the previous distribution function, the term d was included only for the term with αth power (i.e., $(u_{A1}+d)^{-\alpha}$ and $(u_{A2}+d)^{-\alpha}$). The modification of the distribution function into the following form responds the requirement on the asymmetricity:

$$\varphi_2(u_{A1}, u_{A2}) \quad = \begin{cases} A\,(u_{A1}-\delta_1+d)^{-\alpha}\,(u_{A2}-\delta_2+d)^{\beta_1} \\ \quad (u_{A1} > \delta_1,\ \delta_2 < u_{A2} < u_{A1}-\delta_{12},\ \alpha > 0,\ \beta_1 > 0) \\ A\,(u_{A2}-\delta_2+d)^{-\alpha}\,(u_{A1}-\delta_1+d)^{\beta_1} \\ \quad (u_{A2} > \delta_2,\ \delta_1 < u_{A1} > u_{A2}-\delta_{21},\ \alpha > 0,\ \beta_2 > 0) \end{cases} . \tag{A.59}$$

In this distribution function, the demand of good i turns out to be:

$$\begin{aligned} X_i &= A\int_{\delta_i}^{\infty} du_{Ai} \int_{\delta_j}^{u_{Ai}-\delta_{ij}} du_{Aj}\,(u_{Ai}-\delta_i+d)^{-\alpha}\,\left(u_{Aj}-\delta_j+d\right)^{\beta_i} \\ &= \tfrac{A}{1+\beta_i}\int_{\delta_i}^{\infty} du_{Ai}\,(u_{Ai}-\delta_i+d)^{-\alpha+\beta_i+1} \\ &= \tfrac{A}{(1+\beta_i)(-\alpha+\beta_i+2)} d^{-\alpha+\beta_i+2} \end{aligned} . \tag{A.60}$$

The problem of this functional form is that the demand of good i becomes independent from its price, i.e., δ_i. In the standard monopolistic competition literature, the good demand is a decreasing function with respect to its price. Under this relation, it is natural to consider that this assumption on the functional type of the distribution function is not general to retrieve the CES or any other utility functions.

A.8 MATLAB CODE FOR THE CHAPTER 4

The following is a MATLAB code file (example in the case of $\gamma_{min} = 0.50$) for calculating solutions of the Hamilton-Jacobi-Bellman equation.

A.8.1 *Main File*

```
clear all

global gm p al si r mu de th inv v w wi v3 i imax imin almax
    almin gmmax gmmin ra exei itrmax=20;
v1_0=NaN(itrmax,2);
v1_0(1,1)=1000;
v1_0(1,2)=1;

exei=10;
de=0.1007;
r=0.06;
si=0.09;
ra=0.01;
th=1.5;
p=1.05;
mu=0.18;

imax=de+r;
imin=-1;
almax=1;
almin=0.1;
gmmax=1;
gmmin=exei*0.1;

h=1/10000;                   %step size
N=1/h;                       %number of steps
w=zeros(N,1);
w(1)=0;                      %initial value of w
v=zeros(N,2);                %v and v'
v3=zeros(N,1);               %v''
dw=NaN(N,itrmax);
vmax=NaN(N,3,itrmax);
invex=NaN(N,itrmax);         %Policy Function for investment
alex=NaN(N,itrmax);          %Policy Function for alpha
gmex=NaN(N,itrmax);          %Policy Function for gamma

for m=1:itrmax
```

```
v(1,1)=0.9;              %Bolton P.1560 A. Case I.
    Liquidation Para 1, line 2
    v(1,2)=(v1_0(m,1)+v1_0(m,2))/2;
vmax(1,1,m)=v(1,1);
vmax(1,2,m)=v(1,2);
vmax(1,3,m)=0;

for i=1:N-1
    wi=w(i);
    v1=vmax(i,1,m);
    v2=vmax(i,2,m);
    minimization

    if i==1
           for im=1:10
               vmax(1,3,m)=-2*((1+gm*(1-1/p)*(1-al))/(gm*si))
                   ^2*(-(r-inv+de)*vmax(1,1,m)
                                                +(1-gm)/(1+gm
                                                    *(1-1/p)
                                                    *(1-al))*(
                                                    mu+(r-ra)*
                                                    wi-de -th*
                                                    inv^2/2)
                                                +(gm/(1+gm
                                                    *(1-1/p)
                                                    *(1-al))*(
                                                    mu+(r-ra)*
                                                    wi-de-th*
                                                    inv^2/2)
                                                    -(1+wi)*(
                                                    inv-de))*
                                                    vmax(1,2,m
                                                    ));
           minimization
           v3(1)=vmax(1,3,m);
           im=im+1;
       end
    end

    Runge_Kutta_4th_order

    invex(i,m)=inv;
    alex(i,m)=al;
    gmex(i,m)=gm;
    vmax(i+1,1,m)=v(i+1,1);
    vmax(i+1,2,m)=v(i+1,2);
    vmax(i+1,3,m)=-2*((1+gm*(1-1/p)*(1-al))/(gm*si))^2*(-(r-
        inv+de)*v(i+1,1)
                  +(1-gm)/(1+gm*(1-1/p)*(1-al))*(mu+(r-ra)*wi
                      -de -th*inv^2/2)
```

```
                      +(gm/(1+gm*(1-1/p)*(1-al))*(mu+(r-ra)*wi-de
                          -th*inv^2/2)-(1+wi)*(inv-de))*v(i+1,2))
                          ;
        v3(i+1)=vmax(i+1,3,m);
        dw(i,m)=gm/(1+gm*(1-1/p)*(1-al))*(mu+(r-ra)*wi-de-th*inv
            ^2/2)-(1+wi)*(inv-de);
        w(i+1)=i*h;

        if abs(k1(1))>1
            break
        end

        if v(i+1,2)<1
                v1_0(m+1,2)=(v1_0(m,1)+v1_0(m,2))/2;
            v1_0(m+1,1)=v1_0(m,1);
            break
        end

        if vmax(i+1,3,m)>0
            v1_0(m+1,1)=(v1_0(m,1)+v1_0(m,2))/2;
            v1_0(m+1,2)=v1_0(m,2);
            v1_0
            break
        end

        if v(i+1,1)<0
            break
        end

    end

    invex(N,m)=invex(N-1,m);
    alex(N,m)=alex(N-1,m);
    gmex(N,m)=gmex(N-1,m);
    m=m+1;

end
```

A.8.2 *minimization.m*

```
global gm p al si r mu de th inv wi v3 i imax imin almax almin
    gmmax gmmin ra
clear value;

lb(1,1)=imin;
lb(1,2)=almin;
lb(1,3)=gmmin;
hb(1,1)=imax;
hb(1,2)=almax;
hb(1,3)=gmmax;
```

```
is1max=2;
is2max=2;
is3max=2;
value_max=-10000;
vl=zeros(N,1);
value_=NaN(is1max,is2max,is3max);

for is1=1:is1max
    for is2=1:is2max
        for is3=1:is3max
            intl(1) = imin+(imax-imin)*(is1-1)/(is1max-1);
            intl(2) = almin+(almax-almin)*(is2-1)/(is2max-1);
            intl(3) = gmmin+(gmmax-gmmin)*(is3-1)/(is3max-1);

            %x(1)=inv, x(2)=al, x(3)=gm
            fun = @(x) -((x(1)-de)*v1+(1-x(3))/(1+x(3)*(1-1/p)
                *(1-x(2)))*(mu+(r-ra)*wi-de-th*(x(1)-de)^2/2)
                  +(x(3)/(1+x(3)*(1-1/p)*(1-x(2)))*(mu+(r-ra)*wi-
                      de-th*(x(1)-de)^2/2)-(1+wi)*(x(1)-de))*v2
                  +(1/2)*((x(3)*si)/(1+x(3)*(1-1/p)*(1-x(2))))^2*
                      v3(i));
            options = optimoptions('fmincon','Display','off');
            solls = fmincon(fun,intl,[],[],[],[],lb,hb);
            inv_ = solls(1);
            al_ = solls(2);
            gm_ = solls(3);
           value_(is1,is2,is3)=((inv_-de)*v(i,1)+(1-gm_)/(1+gm
               _*(1-1/p)*(1-al_))
                                 *(mu+(r-ra)*wi-de-th*(inv_-de)
                                     ^2/2)
                                +(gm_/(1+gm_*(1-1/p)*(1-al_))
                                 *(mu+(r-ra)*wi-de-th*(inv_-de)
                                     ^2/2)-(1+wi)*(inv_-de))*v(i
                                     ,2)
                                +(1/2)*((gm_*si)/(1+gm_*(1-1/p)
                                    *(1-al_)))^2*v3(i))/r;

            if value_(is1,is2,is3) == 0
                inv_=inv_;
            elseif value_(is1,is2,is3) > value_max
                value_max=value_(is1,is2,is3);
                inv=inv_;
                al=al_;
                gm=gm_;
            end
        end
    end
end
```

A.8.3 *Runge_Kutta_4th_order.m*

```
global v w

k1=h*f(w(i),v(i,1),v(i,2));
k2=h*f(w(i)+h/2, v(i,1)+0.5*k1(1), v(i,2)+0.5*k1(2));
k3=h*f(w(i)+h/2, v(i,1)+0.5*k2(1), v(i,2)+0.5*k2(2));
k4=h*f(w(i)+h, v(i,1)+k3(1), v(i,2)+k3(2));
v(i+1,1)=v(i,1) + (k1(1) + 2*k2(1) + 2*k3(1) + k4(1))/6;
v(i+1,2)=v(i,2) + (k1(2) + 2*k2(2) + 2*k3(2) + k4(2))/6;
```

A.8.4 *f.m*

```
function dv=f(wi,v1,v2)
    global gm p al si r mu de th inv ra
    dv(1)=v2;
    dv(2)=-2*((1+gm*(1-1/p)*(1-al))/(gm*si))^2*(-(r-inv+de)*v1
          +(1-gm)/(1+gm*(1-1/p)*(1-al))*(mu+(r-ra)*wi-de -th*(inv
              -de)^2/2)
          +(gm/(1+gm*(1-1/p)*(1-al))*(mu+(r-ra)*wi-de-th*(inv-de)
              ^2/2)-(1+wi)*(inv-de))*v2);
end
```

BIBLIOGRAPHY

[1] R. P. Agarwal. "A propos d'une note de M. Pierre Humbert," in: *Comptes Rendus de l'Académie des Sciences* 236 (1953), pp. 2031–2032 (cit. on p. 90).

[2] S. P. Anderson, A. D. Palma, and J. F. Thisse. "Demand for Differentiated Products, Discrete Choice Models, and the Characteristics Approach." In: *The Review of Economic Studies* 56.1 (1989), pp. 21–35 (cit. on p. 8).

[3] G. S. Becker. "Irrational Behavior and Economic Theory." In: *The Journal of Political Economy* 70.1 (1962), pp. 1–13 (cit. on pp. 7, 14, 15, 19).

[4] P. Bolton, H. Chen, and N. Wang. "A Unified Theory of Tobin's q, Corporate Investment, Financing, and Risk Management." In: *The Journal of Finance* 66.5 (2011), pp. 1545–1578 (cit. on pp. 8, 47–50, 53, 55, 57, 59, 60, 65–67).

[5] P. Bolton, H. Chen, and N. Wang. "Market Timing, Investment, and Risk Management." In: *The Journal of Financial Economics* 109 (2013), pp. 40–62 (cit. on pp. 47, 48).

[6] A. Deaton and J. Muellbauer. *Economic and Consumer Behavior*. Cambridge University Press, 1980 (cit. on pp. 8, 13).

[7] G. Debreu. "Excess-demand Functions." In: *Journal of Mathematical Economics* 1 (1980), pp. 15–21 (cit. on pp. 6, 13).

[8] J-P. Decamps, T. Mariotti, J-C. Rochet, and S. Villeneuve. "Free Cash Flow, Issuance Costs, and Stock Prices." In: *The Journal of Finance* 66.5 (2011), pp. 1501–1544 (cit. on pp. 47, 50, 51, 55, 60).

[9] P. M. Demarzo, Z. H. Fishman, and N. Wang. "Dynamic Agency and the q Theory of Investment." In: *The Journal of Finance* 67.6 (2012), pp. 2295–2340 (cit. on pp. 50, 57).

[10] P. M. DeMarzo and Y. Sannikov. "Optimal Security Design and Dynamic Capital Structure in a C*ontinuous-Time Agency Model, journal = The Journal of Finance, year = 2007, volume = 61, number = 6, pages = 2681–2724," in: () (cit. on p. 55).

[11] A. K. Dixit and J. E. Stiglitz. "Monopolistic Competition and Optimum Product Diversity." In: *American Economic Review* 67.3 (1977), pp. 297–308 (cit. on pp. 5, 29, 43).

[12] W. Feller. *An Introduction to Probability Theory and Its Applications, vol. 2*. Wiely, 1966 (cit. on p. 24).

[13] S. M. Goldman and H. Uzawa. "A Note on Separability in Demand Analysis." In: *Econometrica* 32.3 (1964), pp. 387–398 (cit. on p. 8).

[14] W. M. Gorman. "Separable Utility and Aggregation." In: *Econometrica* 27 (1959), pp. 469–481 (cit. on p. 8).

[15] W. Hardle, W. Hildenbrand, and M. Jerison. "Empirical Evidence on the Law of Demand." In: *Econometrica* 59.6 (1991), pp. 1525–1549 (cit. on p. 6).

[16] W. Hildenbrand. "On the Law of Demand." In: *Econometrica* 51.4 (1983), pp. 997–1019 (cit. on pp. 6, 7, 14, 15).

[17] W. Hildenbrand. *Market Demand: Theory and Empirical Evidence*. Princeton Press, 1994 (cit. on pp. 6, 7).

[18] W. Hildenbrand and M. Jerison. "The Demand Theory of the Weak Axioms of Revealed Preference." In: *Economic Letters* 29.3 (1989), pp. 209–213 (cit. on p. 6).

[19] W. Hildenbrand and A. Kneip. "Demand Aggregation under Structural Stability." In: *Journal of Mathematical Economics* 31 (1999), pp. 81–109 (cit. on p. 6).

[20] H. S. Houthakker. "The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis." In: *The Review of Economic Studies* 23.1 (1955), pp. 27–31 (cit. on pp. 30, 33, 44).

[21] H. S. Houthakker and L. D. Taylor. *Consumer Demand in the United States: Analyses and Projections, Second Edition*. Harvard UniversityPress, 1970 (cit. on p. 13).

[22] J. Hugonnier, S. Malamud, and E. Morellec. "Credit Market Frictions and Capital Structure Dynamics." In: *Journal of Economic Theory* 157 (2015), pp. 1130–1158 (cit. on p. 9).

[23] Malamud S. Hugonnier J. and E. Morellec. "Capital Supply Uncertainty, Cash Holdings, and Investment." In: *The Review of Financial Studies* 28.2 (2015), pp. 391–445 (cit. on p. 55).

[24] M. Jerison. "Aggregation and Pairwise Aggregation of Demand When the Distribution of Income is Fixed." In: *Journal of Economic Theory* 33 (1984), pp. 1–31 (cit. on p. 25).

[25] S. Jespersen, R. Metzler, and Fogedby H. C. "Lévy Flights in External Force Lelds: Langevin and Fractional Fokker-Planck Equations and Their Solutions," in: *Physical Review E* 59.3 (1999), pp. 2736–2745 (cit. on p. 91).

[26] C. I. Jones. "The Shape of Production Functions and the Direction of Technical Change." In: *The Quartery Journal of Economics* 120.2 (2005), pp. 517–549 (cit. on p. 44).

[27] A. Ya. Khintchine. *Limit Laws for Sums of Independent Random Variables*. ONTI, Moscow-Leningrad, 1938 (cit. on p. 24).

[28] A. P. Kirman. "Whom or What Does the Representative Individual Represent?" In: *Journal of Economic Perspectives* 6.2 (1992), pp. 117–136 (cit. on p. 14).

[29] N. Kiyotaki. "A Perspective on Modern Business Cycle Theory." In: *Economic Quarterly* 97.3 (2011), pp. 195–208 (cit. on p. 14).

[30] R. Kubo. "The Fuctuation Dissipation Theorem." In: *Reports on Progress in Physics* 29 (1966), pp. 255–284 (cit. on p. 22).

[31] F. E. Kydland and E. C. Prescott. "Time to Build and Aggregate Fluctuations." In: *Econometrica* 50.6 (1982), pp. 1345–1370 (cit. on p. 5).

[32] K. J. Lancaster. "A New Approach to Consumer Theory." In: *Journal of Political Economy* 74.2 (1966), pp. 132–157 (cit. on p. 8).

[33] P. Levy. *Theorie de l'addition des variables aleatoires*. Gauthier-Villars, Paris., 1954 (cit. on p. 24).

[34] R. Mantel. "On the Characterization of Aggregate Excess-demand." In: *Journal of Economic Theory* 7 (1974), pp. 348–353 (cit. on pp. 6, 13).

[35] A. Mas-Colell, M. D. Whinston, and J. R. Green. *Microeconomic Theory*. Oxford University Press, 1995 (cit. on pp. 6, 15).

[36] J. A. de Matos. *Theoretical Foundations of Corporate Finance*. Princeton University Press, 2001 (cit. on pp. 9, 48).

[37] D. McFadden. *Econometric Models of Probabilistic Choice In: Manski, C., McFaaded, D. (eds), Structural Analysis of Discrete Data with Eeconometric Applications, pp198-272*. The MIT Press, 1981 (cit. on p. 8).

[38] M. J. Melitz. "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity." In: *Econometrica* 71.6 (2003), pp. 1695–1725 (cit. on pp. 29, 41).

[39] R. Metzler. "The Random Walk's Guide to Anomalous Diffusion: A Fractional Dynamics Approach." In: *Physics Report* 339 (2000), pp. 1–77 (cit. on pp. 21, 24).

[40] G. M. Mittag-Leffler. "Sur la nouvelle fonction E(x)." In: *Comptes Rendus de l'Académie des Sciences Paris* 137 (1903), pp. 554–558 (cit. on p. 90).

[41] F. Modigliani and M. Miller. "The Cost of Capital, Corporation Finance and the Theory of Investment." In: *American Economic Review* 48 (1958), pp. 261–297 (cit. on pp. 8, 54).

[42] H. Mori. "Transport, Collective Motion, and Brownian Motion." In: *Progress of Theoretical Physics* 33.3 (1965), pp. 423–455 (cit. on p. 22).

[43] OECD. "Growth Companies, Access to Capital Markets and Corporate Governance." In: *OECD Report to G20 Finance Ministers and Central Bank Governors* (2015) (cit. on p. 9).

[44] R. F. Pawula. "Approximation of the Linear Boltzman Equation by the Fokker-Planck Equation," in: *Physical Review* 162.1 (1967), pp. 186–188 (cit. on p. 83).

[45] T. R. Prabhakar. "A Singular Integral Equation with Generalized Mittag-Leffler Function in the Kernel." In: *The Yokohama Mathematical Journal* 19 (1971), pp. 7–25 (cit. on p. 90).

[46] B. Smith, Z. Abdoolakhan, and J. Taplin. "Demand and Choice Elasticities for a Separable Product Group." In: *Economic Letters* 108 (2010), pp. 134–136 (cit. on p. 8).

[47] H. Sonnenschein. "Do Walras' Identity and Continuity Characterize the Class of Community Excess Demand Functions?" In: *Journal of Economic Theory* 6 (1973), pp. 345–354 (cit. on pp. 6, 13).

[48] R. H. Stotz. "The Empirical Implications of a Utility Tree." In: *Econometrica* 27 (1957), pp. 482–488 (cit. on p. 8).

[49] H. Takayasu. *Fractals in the Physical Sciences*, Manchester University Press, 1990 (cit. on p. 24).

[50] A. D. Vinales and M. A. Desposito. "Anomalous Diffusion Induced by a Mittag-Leffler Correlated Noise." In: *Physical Review E* 75 (2007), p. 042102 (cit. on p. 25).

[51] M. Weber and L. Schneider. "Marginal Utility Theory and "The Fundamental Law of Psychophysics"." In: *Social Science Quarterly* 56 (1975), pp. 21–36 (cit. on p. 36).