RIETI Discussion Paper Series 14-E-038

Representative Agent in a Form of Probability Distribution

INOSE Junya
University of Tokyo

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 14-E-038
July 2014

Representative Agent in a Form of Probability Distribution

INOSE Junya
University of Tokyo

Abstract

This paper provides a methodology for demand aggregation with rigorous micro foundation by using the probability distribution function. In the model setting, a set of agents defined by a 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 a single representative agent as a whole economy is provided for the following three cases: i) normal diffusion with Markov process, ii) normal diffusion with non-Markov process, and iii) anomalous diffusion. In the case of normal diffusion, no restriction condition for the shape of the probability density function of the consumption bundle is required because the Gaussian distribution has a property of stability. On the other hand, in the case of anomalous diffusion, some restricting conditions on the parameter of the probability process are required to maintain stability.

Keywords: Representative agent, Heterogeneity, Probability distribution, Statistical dynamics
JEL classification: E21

This study is conducted as a part of the Project "Issues Faced by Japan's Economy and Economic Policy Part III: Heterogeneity among economic agents" undertaken at Research Institute of Economy, Trade and Industry (RIETI)

Representative Agent in a Form of Probability Distribution

Junya Inose

Department of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033

1 Introduction

A struggle on describing heterogeneity and its demand aggregation based on rigorous micro foundation has a long history. The standard aggregation theory has shown its major development until the early 80s (as a reference, see for example, Deaton and Muellbauer (1980), Houthakker and Taylor (1970), etc.). The basic standard aggregation theory confine its interest to the market transaction under privately consumption only (e.g., no public goods nor externalities), and provide two core approach for 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 progress exerted influence on the development of recent 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 (1973), Mantel (1974) and Debreu (1974), and now known as the Debreu-Sonnenschein-Mantel Theorem. They raised a problem on whether the structure of the individual demand, which is based on utility maximization and generates a lot of structure, is preserved by demand aggregation. 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 (1983). In a series of papers referring result in 1983, 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 (1989), Härdle et al. (1991), and Hildenbrand and Kneip (1999)).

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 (1994)). 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, in stead 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 exist 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 (1994)) and household attributes, such as age and employment status or household size, are employed for the empirical calculation.

In this paper, we extend traditional and recent works on demand aggregation under various household attributes by formulating consumer's utility in probability density function. Here it should be notified that, in the literature of Hildenbrand, a demand function of each household is a starting point, and hope nor to resort to sophisticated technical conditions on preferences or utility functions is abandoned. However, in our literature, we both employ the preference, or utility functions, and the core concept of μ model, to divide household's purchasing based on their attributes

Moreover, the introduction of this framework also enable us 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 allocation that maximizes 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 (1992)). Following these critics, some research papers with counterarguments are also published (see, for example, Kiyotaki (2011)). These discussions are, as a matter of cause, based on the market completeness. In this paper, we address a restricting condition for the approximation of representative agent when the micro utility functions are defined with a shape of probability density functions. The contribution of this paper in this field is to provide the restriction condition for the representative agent both in case of Markov and non-Markov process.

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 demand part of the general equilibrium model. Therefore we do not set any focus on the profit maximization of firms, market clearing conditions, etc. in this paper, but confine our interest only into the demand aggregation to provide better approximated description.

The reminder of the paper 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 Micro Foundation and Problem Setup

2.1 Definition of the Micro Foundation

The concept of Pareto optimality is rigorously defined in microeconomics (see, for example, Green (1996)). Consider an economy consisting of I consumers and L goods. 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 "$\succeq_i$". The total amount of each good $l = 1, ..., L$ initially available in the economy, called the total endowment of good l, is denoted by $\omega_l \geq 0$ for $l = 1, ..., L$.

It is often of interest to ask whether an economic system is producing an "optimal" economic outcome. An essential requirement for any optimal economic allocation is that it possess the property

of Pareto optimality. Here, a feasible allocation is Pareto optimal (or Pareto efficient) if there is no other feasible allocation such that

$$^{\forall}i,\ x_i' \succeq_i x_i \ and \ ^{\exists}i \ s.t. \ x_i' \succ_i x_i. \tag{2.1}$$

Moreover, the first fundamental theorem of welfare economics shows that the condition of a competitive equilibrium, or Walras equilibrium, and Pareto optimal becomes rigorously equal under the complete market assumption in price taking situation. Therefore, the definition of Pareto optimality (2.1) and definition of Walras equilibrium (2.2) becomes rigorously equal:

$$^{\forall}i,\ x_i^* \in B_i(p^*) \ and \ x_i^* \ \succeq_i \ x_i', \ \ where \ B_i(p^*) \equiv \{x_i \in R_+^L \,|\, p^* \cdot x_i \leq p^* \cdot \omega_i\} \tag{2.2}$$

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.

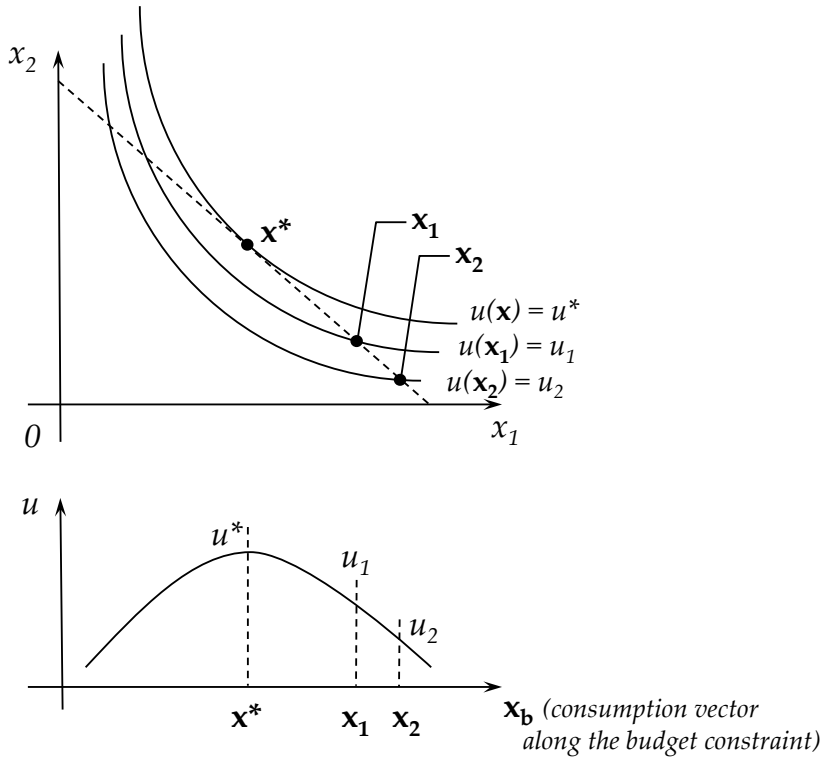

Figure 2.1: Utility along the Budget Constraint Plane

The similar discussion could be extended to the L goods situation. In the L goods situation, the budget constraint becomes a hyper plane 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 of the consumption bundle with μ character has some kind of distribution written in Figure 2.2, and each consumption bundle of household i, $i \in I(\mu)$ satisfies the Pareto optimal condition according to each utility function $u_i(x)$.

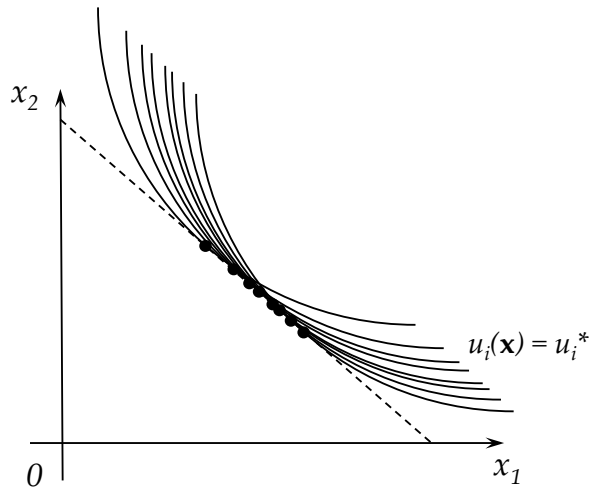

Figure 2.2: Observable Cloud of Consumption Bundles under the Same Household Characteristics

2.2 Describing Probability Density Function of the Observable Cloud of Consumption Bundles

One standard way to analyze the distribution of the observable clouds of consumption bundles is to introduce the concept of probability distribution into the system. 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. Here we consider the probability process of the consumption bundle $\mathbf{x_b} = \{\mathbf{x_b}(t)\}_{t>0}$ under price taking situation. Let $\mathbf{x_b}$ be a non-trivial Lévy process in R_+^{L-1} which follows the stochastic differential equation 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 hyperplane*

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 economical meaning of these 2 assumptions are to assume representativeness for a group of similar household attributes as an approximation. If we closely divide adequate amount of consumers by its attributes, the assumptions allows us to assume 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 households, 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 counterargument 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, there are no need to discuss the validity of this approximation in theoretical field only. 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 more close 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 understandings for this approximation is described in the Figure 2.3. Most theories with rigorous micro foundation assumes that all of the household's utility function can be different and also can move with errors. However, every consumption bundles 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 equilibrium 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. The major approximation assumed here is i) the distribution of any individual household's Walras equilibrium point could be approximated within single utility function, and ii) the dynamics of each household's Walras equilibrium point could be approximated by well-defined process (like Brownian motion, Lévy flights, etc.).

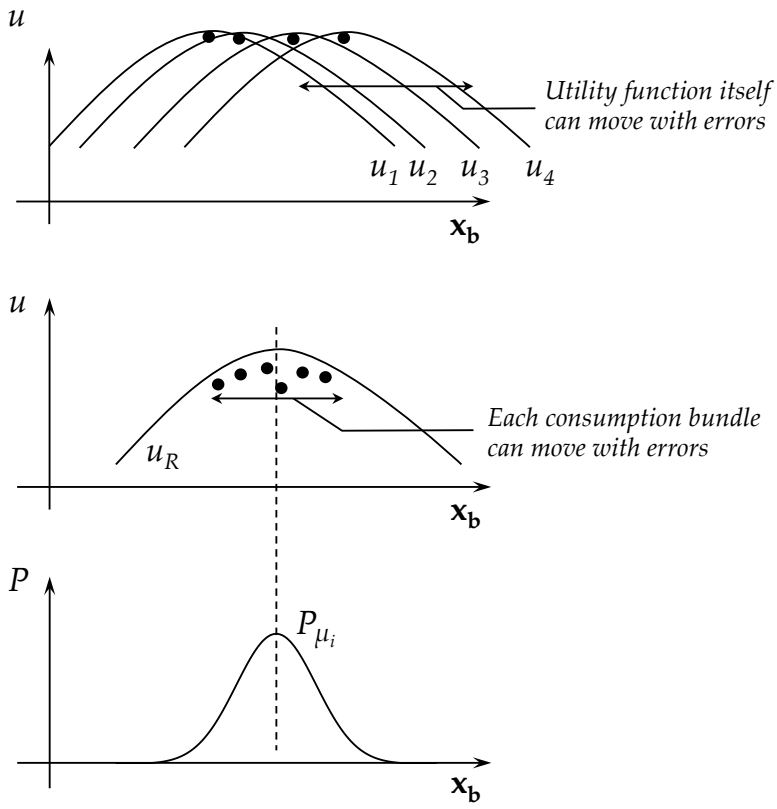

Figure 2.3: Aggregation of Heterogeneous Utility Functions and Probability Density Function of the Consumption Bundles

2.2.1 Describing Dynamics of Observable Clouds

To describe dynamic stochastic path toward equilibrium, an equation to describe dynamics of the observable clouds is required. For the simplicity, we first assume the consumption bundles evolves with Markov process, and relax this assumption in later section.

Assumption 3. *Any consumption bundles within the household type μ_i evolves with Markov process, i.e.,*

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

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*

$$\frac{\partial}{\partial t}P_{\mu_i}(\mathbf{x_b},t) = -P_{\mu_i}(\mathbf{x_b},t)\xi(\mathbf{x_b}) + \int P_{\mu_i}(\mathbf{x'_b},t)T(\mathbf{x'_b},\mathbf{x_b})d\mathbf{x'_b} \tag{2.3}$$

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;

$$P_{\mu_i}(\mathbf{x_b},t \mid \mathbf{x_{b0}},t_0) = \int P_{\mu_i}(\mathbf{x_b},t \mid \mathbf{x'_b},t')\, P_{\mu_i}(\mathbf{x'_b},t' \mid \mathbf{x_{b0}},t_0)\, d\mathbf{x'_b} \tag{2.4}$$

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:

$$P_{\mu_i}(\mathbf{x_b},t+\Delta t|\mathbf{x_{b0}},t_0) = \int \left[\left\{1 - \int T(\mathbf{x_b},\mathbf{x''_b})\, d\mathbf{x''_b}\right\}\, \Delta t\, \delta(\mathbf{x_b}-\mathbf{x'_b}) + T(\mathbf{x'_b},\mathbf{x_b})\Delta t\right] P_{\mu_i}(\mathbf{x'_b},t|\mathbf{x_{b0}},t_0)\, d\mathbf{x'_b} \tag{2.5}$$

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

$$\frac{\partial}{\partial t}P_{\mu_i}(\mathbf{x_b},t \mid \mathbf{x_{b0}},t_0) = -P_{\mu_i}(\mathbf{x_b},t \mid \mathbf{x_{b0}},t_0)\xi(\mathbf{x_b}) + \int P_{\mu_i}(\mathbf{x'_b},t \mid \mathbf{x_{b0}},t_0)T(\mathbf{x'_b},\mathbf{x_b})d\mathbf{x'_b} \tag{2.6}$$

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 $\mathbf{x_b}$.

2.2.2 Deriving Fokker Plank Equation

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

$$\frac{\partial}{\partial t}P_{\mu_i}(x_1,x_2,...,x_L,t) = \left(-\sum_{j=1}^{L}\frac{\partial}{\partial x_j}\alpha_{1j}(\mathbf{x_b}) + \frac{1}{2}\sum_{j=1}^{L}\sum_{k=1}^{L}\frac{\partial^2}{\partial x_j \partial x_k}\alpha_{2jk}(\mathbf{x_b})\right) P_{\mu_i}(x_1,x_2,...,x_L,t) \tag{2.7}$$

Proof. Define $\mathbf{r}$ as $\mathbf{r} = \mathbf{x}'_{\mathbf{b}} - \mathbf{x}_{\mathbf{b}}$ and $\omega(\mathbf{x}'_{\mathbf{b}}, \mathbf{r}) = T(\mathbf{x}'_{\mathbf{b}}, \mathbf{x})$. Substitute this expression into equation (2.6) to obtain

$$\frac{\partial}{\partial t} P_{\mu_i}(\mathbf{x_b}, t) = -\int \omega(\mathbf{x_b}, \mathbf{r})\, d\mathbf{r}\, P_{\mu_i}(\mathbf{x_b}, t) + \int \omega(\mathbf{x_b} - \mathbf{r}, \mathbf{r})\, d\mathbf{r}\, P_{\mu_i}(\mathbf{x_b} - \mathbf{r}, t) \tag{2.8}$$

Here we introduce following formula for the further calculation.

$$\exp\left[-\mathbf{r}\frac{\partial}{\partial \mathbf{x_b}}\right] f(\mathbf{x_b}) = \sum_{n=0}^{\infty} \frac{(-\mathbf{r})^n}{n!} \left(\frac{\partial}{\partial \mathbf{x_b}}\right)^n f(\mathbf{x_b}) = f(\mathbf{x_b} - \mathbf{r}) \tag{2.9}$$

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

$$\begin{aligned}
\frac{\partial}{\partial t} P_{\mu_i}(\mathbf{x_b}, t) &= -\int \omega(\mathbf{x_b}, \mathbf{r})\, d\mathbf{r}\, P_{\mu_i}(\mathbf{x_b}, t) + \int d\mathbf{r} \sum_{n=0}^{\infty} \frac{(-\mathbf{r})^n}{n!} \left(\frac{\partial}{\partial \mathbf{x_b}}\right)^n \omega(\mathbf{x_b}, \mathbf{r})\, P_{\mu_i}(\mathbf{x_b}, t) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(\frac{\partial}{\partial \mathbf{x_b}}\right)^n \int d\mathbf{r}\, \mathbf{r}^n \omega(\mathbf{x_b}, \mathbf{r})\, P_{\mu_i}(\mathbf{x_b}, t) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(\frac{\partial}{\partial \mathbf{x_b}}\right)^n \alpha_{\mathbf{n}}(\mathbf{x_b})\, P_{\mu_i}(\mathbf{x_b}, t)
\end{aligned} \tag{2.10}$$

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 4** for the better approximation:

Assumption 4. *The coefficients $\alpha_n(\mathbf{x_b})$ are finite for every n and $\alpha_n(\mathbf{x_b}) = 0$ for some even n*

According to the Pawula theorem (Pawula (1967)), 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 (2.9) could be approximated simply as follows.

$$\frac{\partial}{\partial t} P_{\mu_i}(x_{b1}, x_{b2}, ..., x_{bL-1}, t) = \left(-\sum_{j=1}^{L-1} \frac{\partial}{\partial x_j} \alpha_{1j}(\mathbf{x_b}) + \frac{1}{2} \sum_{j=1}^{L-1} \sum_{k=1}^{L-1} \frac{\partial^2}{\partial x_j \partial x_k} \alpha_{2jk}(\mathbf{x_b}) \right) P_{\mu_i}(x_{b1}, x_{b2}, ..., x_{bL-1}, t) \tag{2.11}$$

□

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.

Now let us move our focus back to our assumptions. As previously described, we assumed this process as the Markov process, and therefore the effect of friction or any other term which violate a Markov process assumption is ignored. On the other hand, an effect of friction plays important role in economics because this term generally governs a speed of relaxation after drastic change of external environment (e.g., tax revisions, monetary policy change, etc.). Therefore, we introduce an expansion of this equation to include the effect of friction and derive Kramers equation in the next subsection.

2.2.3 Deriving 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):*

$$\frac{\partial}{\partial t}P_{\mu_i}(\mathbf{x_b},\ \mathbf{v_b},t)=\left[-\sum_{j=1}^{L-1}\frac{\partial}{\partial x_{bj}}v_{bj}+\sum_{j=1}^{L-1}\frac{\partial}{\partial v_{bj}}\left(-\frac{\partial}{\partial x_{bj}}u_R(x)+\gamma v_{bj}\right)+D\sum_{j=1}^{L-1}\frac{\partial^2}{\partial v_{bj}^2}\right]P_{\mu_i}(\mathbf{x_b},\ \mathbf{v_b},t) \tag{2.12}$$

Proof. The derivation of Fokker-Planck equation (2.12) 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:

$$\frac{\partial}{\partial t}P_{\mu_i}(\mathbf{x_b},\ \mathbf{v_b},t)=\left(-\sum_{j=1}^{2(L-1)}\frac{\partial}{\partial A_j}\alpha_{1j}(\mathbf{A})+\frac{1}{2}\sum_{j=1}^{2(L-1)}\sum_{k=1}^{2(L-1)}\frac{\partial^2}{\partial A_j\partial A_k}\alpha_{2jk}(\mathbf{A})\right)P_{\mu_i}(\mathbf{x_b},\ \mathbf{v_b},t) \tag{2.13}$$

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

$$\frac{\partial}{\partial t}P_{\mu_i}(\mathbf{x_b},\ \mathbf{v_b},t)=\left[-\sum_{j=1}^{L-1}\frac{\partial}{\partial x_{bj}}v_{bj}+\sum_{j=1}^{L-1}\frac{\partial}{\partial v_{bj}}\left(-\frac{\partial}{\partial x_{bj}}u_R(x)+\gamma v_{bj}\right)+D\sum_{j=1}^{L-1}\frac{\partial^2}{\partial v_{bj}^2}\right]P_{\mu_i}(\mathbf{x_b},\ \mathbf{v_b},t) \tag{2.14}$$

□

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 (2.15) as zero and plugging this solution and into the RHS of (2.15).

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 force to stay close to the consumption bundle of the typical household. This effect could be seen when the crowd with the same attribute conform some trend and most of households follow similar consumption attribute. 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 motives for abandoning the ideal of optimization. For example, facts like i) no exact optimizing strategy is known in many real-world situations or, ii) it may demand unrealistic amounts of knowledge about alternatives and consequences even when an optimizing strategy exists, could be motives for abandoning ideal of optimization.

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.

3 Demand Aggregation

3.1 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 economical phenomenon.

3.1.1 Derivation of the generalized Langevin equation

Proposition 5. *Under the assumption of the ergodic property, following Liouville equation is derived*

$$\frac{\partial}{\partial t}\rho = i\mathscr{L}\rho \equiv \{\rho, H\} = \sum_{j=1}^{L-1}\left(\frac{\partial \rho}{\partial x_{bj}}\frac{\partial H}{\partial v_{bj}} - \frac{\partial \rho}{\partial v_{bj}}\frac{\partial H}{\partial x_{bj}}\right) \tag{3.1}$$

where $\mathscr{L}$ is a Liouville operator, $\rho = \rho(\mathbf{x_b}, \mathbf{v_b}, t)$ is a density of representative points $(\mathbf{x_b}, \mathbf{v_b})$ at time t and H is a Hamiltonian which satisfies following condition:

$$\dot{x_{ib}} = \frac{\partial H}{\partial v_{bj}},\ \dot{v_{ib}} = \frac{\partial H}{\partial x_{bj}}$$

The detailed proof of this proposition is left in the Appendix A-1.

Proposition 6. *The equation of motion of each consumption bundle could be described in the form of the generalized Langevin equation as follows:*

$$\frac{d}{dt}\mathbf{v_b}(t) = -\int_{-\infty}^{t}\gamma(t-t')\mathbf{v_b}(t')dt' + \nabla u_R(\mathbf{x_b}) + \mathbf{R}(t) \tag{3.2}$$

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

The detailed proof of this proposition is left in the Appendix A-2. This concept is firstly developed in the field of statistical physics (see, for example, Mori (1965) and Kubo (1966)), and here let us spare several lines to consider economical 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 equals to the friction toward sudden change of the consumption bundle. This term 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 equilibrium, 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.

In most cases, the random force $\mathbf{R}(t)$ is assumed to be independent and identically distributed. However, our fundamental assumption (**Assumption 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.

3.1.2 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.,*

$$< \mathbf{R}(t_1)\mathbf{R}(t_2) >= 2\pi \mathbf{G_R} \delta(t_1 - t_2) \tag{3.3}$$

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 reasonable for a Brownian particle having a mass much larger than the colliding molecules, because the motion of Brownian particle is a result of a great number of successive collisions, which is a condition for the central limit theorem. The Gaussian assumption also sets restriction on the form of the mean square displacement as:

$$< (\Delta x)^2 > \propto D\Delta t, \;\; t \to \infty \tag{3.4}$$

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:

$$\frac{d}{dt}\mathbf{v_b}(t) = -\gamma \mathbf{v_b}(t) + \nabla u_R(\mathbf{x_b}) + \mathbf{R_w}(t) \tag{3.5}$$

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

3.1.3 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 **Assumption ND − 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,

$$\frac{d}{dt}\mathbf{v_b}(t) = -\int_{-\infty}^{t} \gamma(t-t')\mathbf{v_b}(t')dt' + \nabla u_R(\mathbf{x_b}) + \mathbf{R}(t) \tag{3.6}$$

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 Oshtein-Uhlenbeck process with unit auto regressive term like

$$d\mathbf{v_b}(t) = \mu\mathbf{v_b}(t)\cdot dt + \nabla u_R(\mathbf{x_b}(t))\cdot dt + \mathbf{R}(t) \tag{3.7}$$

the standard Langevin equation introduced in (3.5) will be derived. Therefore, if we would like to derive the generalized Langevin equation written in (3.6), 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 most 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 (1971) is described as:

$$E^{\rho}_{\mu,\nu}(t-t') = \sum_{k=0}^{\infty}\frac{(\rho)_k}{\Gamma(\mu k+\nu)}\frac{(t-t')^k}{k!}, \tag{3.8}$$

with $Re(\mu) > 0$, $Re(\nu) > 0$ and $Re(\rho) > 0$ and $z \in \boldsymbol{C}$. When the parameter $\rho = 1$, we get the two parameter Mittag-Leffler function as was introduced by Agarwal, and with $\rho = \nu = 1$, we recover the Mittag-Leffler function originally introduced by Mittag-Leffler. The case $\rho = \nu = \mu = 1$ reduces to the exponential function.

The economical meaning of this relation is, as already mentioned, the description of an effect of a relaxation against drastic transformation of the exterior environment. If **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.

3.1.4 Assumption for Anomalous Diffusion

Lastly let us now consider relaxation of the assumption on Gaussian process (**Assumption ND − 2.**). In a wide diversity of physical systems, anomalous diffusion is found in several cases. 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 relation between the mean square displacement and time changes as:

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

$$< \mathbf{R}(t_1)\mathbf{R}(t_2) >= 2\pi\mathbf{G_R}\delta(t_1 - t_2) \tag{3.9}$$

$$< (\Delta x)^2 > \propto 2D\Delta t^{\alpha}, \ \ t \to \infty \tag{3.10}$$

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 to change 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 Jespersen et. al. (1999), in the field of physics, 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. On the other hand, the enhanced diffusion ($\alpha > 1$) include tracer particles in vortex arrays in a rotating flow, layered velocity field, and Richardson 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 economical activity in this paper because of the lack of adequate data set of consumption bundles, however, it is technically possible to evaluate α if we could obtain adequate panel data.

3.2 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) and standard assumption for the error term, like having an i.i.d. property, no longer holds. This leads to the failure of the assumption of the representative agent because if the aggregated variable of the representative agent does not have normal error term, the usual calibration methodology can not be applied. 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.

3.2.1 Case of Normal Diffusion under Markov Process

In the case of normal diffusion under Markov process (case of no previous memory), 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.

3.2.2 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 (3.5) 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. The detailed formulation of this effect is described in the section 4.

3.2.3 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-Plank equation, see Metzler and Klafter (2000))

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 (1990)), whose basic concept was originally derived by Lévy (1937) and Khintchine (1938). The definition of the stable distribution is given by Feller (1966) 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*

$$Y_n \stackrel{d}{=} c_n X + \epsilon_n \tag{3.11}$$

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

In general, the sum of random variables with a common distribution becomes a random variable with a distribution of different form. However, for random variables with a stable distribution, an appropriate linear transformation makes the sum of random variables obey the same distribution.

Using the characteristic function of a distribution,

$$\phi(z) \equiv < e^{iXz} > = \int_{-\infty}^{+\infty} e^{iXz} dR(X) \tag{3.12}$$

the relation (3.10) is transferred into

$$\phi^n(z) = \phi(c_n z) \cdot e^{i\epsilon_n z} \tag{3.13}$$

Proposition 8. *(3.12) can be solved completely and the solution becomes:*

$$\psi(x,t) \equiv log\, \phi(z) = i\mu z - \nu|z|^\alpha \left\{ 1 + i\beta \frac{z}{|z|} \omega(z,\alpha) \right\} \tag{3.14}$$

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

$$\omega(z,\alpha) = \begin{cases} tan\frac{\pi\alpha}{2} & if\ \alpha \neq 1 \\ \frac{2}{\pi} log|z| & if\ \alpha = 1 \end{cases} \tag{3.15}$$

α is called the Lévy index or characteristic exponent. The limiting case $\alpha = 2$ corresponds to the Gaussian normal distribution governed by the central limit theorem. For $\beta = 0$, the distribution is symmetric. μ is the parameter which translates the distribution, and ν is a scaling factor for X. The proof of this proposition could easily be done by just plugging (3.13) and (3.14) into (3.12).

Proposition 9. *The equations (3.13) and (3.14) can be rewritten more simply in the form:*

$$\psi(x,t) = -|z|^{\alpha} exp\left\{ i\frac{\pi\beta}{2} sign(z) \right\} \tag{3.16}$$

with the new centering constant β which is restricted in the following region:

$$|\beta| \leq \begin{bmatrix} \alpha & if\ 0 < \alpha \leq 1 \\ 2-\alpha & if\ 1 < \alpha \leq 2 \end{bmatrix} \tag{3.17}$$

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

The resulting allowed parameter space (Takayasu Diamond) is portrayed in Figure 3.1. All pairs of indices inside and on the edge of the diamond shape refer to proper stable laws. The double line denotes one-sided stable laws (OS). The letters represent the normal or Gaussian law (N), the Holtsmark distribution (H), the Cauchy or Lorentz distribution (C), and the approximate log-normal distribution (L) close to $\alpha \approx 0$.

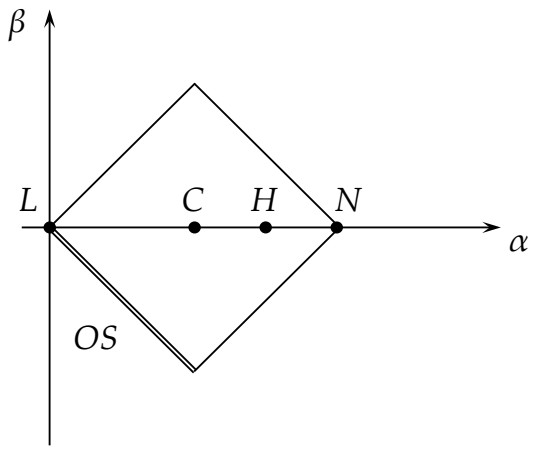

Figure 3.1: Parameter Space for Stable Laws

4 Toward an Empirical Test

4.1 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 if we set an assumption that the consumer panel data is acquired under equilibrium. 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.

4.2 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 Viñales and Despósito (2007), mean square displacement for times $t \gg \tau$ under generalized Langevin equation with Mittag-Leffler function is calculated as:

$$< X^2(t) > \approx 2\gamma D t^2 E^1_{2-\mu,3}[-(\omega_\mu t)^{2-\mu}] + 2\gamma D \frac{\tau^\mu}{\gamma_\mu} \left\{1 - E^1_{2-\mu,1}[-(\omega_\mu t)^{2-\mu}]\right\}. \tag{4.1}$$

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.

4.3 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 Jespersen et. al. (1999) and the result shows asymptotic power law behavior:

$$P^{eq}_{\mu_i}(\mathbf{x_b}) \approx \frac{D\gamma}{\mu\lambda|\mathbf{x_b}|^{1+\mu}} \tag{4.2}$$

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.

5 Conclusions

In this paper we established new methodology for approximating huge amount of consumers. The approximation of this model is 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 unintended payment) as i.i.d. and treat sum of these two effects as single error. If this assumption holds generally 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 follows non-Markov process, or anomalous diffusion, the assumption of the representative agent still holds within limited parameter space.

We would like to point out that our approach allows us to provide new methodology to verify the rationality of representative agent, and also hope that this theoretical framework encourages future applications in empirical works.

A Mathematical Miscellany

A.1 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

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

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

$$\{\rho(\mathbf{x_b}, \mathbf{v_b}, t)\dot{x_{b1}}\}\,|_{x_{1b}+\Delta x_{1b}}\Delta V_{-x_{1b}} = \left(\rho(\mathbf{x_b}, \mathbf{v_b}, t)\dot{x_{b1}} + \frac{\partial\rho(\mathbf{x_b}, \mathbf{v_b}, t)\dot{x_{b1}}}{\partial x_{b1}}\Delta x_{1b}\right)\Delta V_{-x_{1b}} \tag{A.2}$$

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:

$$\frac{\partial\rho(\mathbf{x_b}, \mathbf{v_b}, t)\dot{x_{b1}}}{\partial x_{b1}}\Delta x_{1b}\Delta x_{2b}\cdots\Delta x_{L-1b}\cdot\Delta v_{1b}\Delta v_{2b}\cdots\Delta v_{L-1b} \tag{A.3}$$

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:

$$\frac{\partial}{\partial t}\rho = i\mathscr{L}\rho \equiv \{\rho, H\} = \sum_{j=1}^{L-1}\left(\frac{\partial\rho}{\partial x_{bj}}\frac{\partial H}{\partial v_{bj}} - \frac{\partial\rho}{\partial v_{bj}}\frac{\partial H}{\partial x_{bj}}\right) \tag{A.4}$$

A.2 Derivation of Generalized Langevin Equation

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

$$\frac{d}{dt}A_\mu(t) = i\mathscr{L}A_\mu(t) \tag{A.5}$$

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:

$$(F, G) = (G, F)^*, \tag{A.6}$$

$$(G, G) \geq 0, \tag{A.7}$$

$$(\sum_i c_i F, G) = \sum_i c_i (F, G) \tag{A.8}$$

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

$$(A_\mu(0), A_\nu(0)) = \delta_{\mu\nu} \tag{A.9}$$

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

$$\mathscr{P}G(t) = \sum_\nu (G(t), A_\nu(0)) A_\nu(0) \tag{A.10}$$

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

$$(\mathscr{P}F, G) = (F, \mathscr{P}G) \tag{A.11}$$

$$(\mathscr{P}'F, G) = (F, \mathscr{P}'G) \tag{A.12}$$

$$\mathscr{P}^2 = \mathscr{P},\ \mathscr{P}'^2 = \mathscr{P}',\ \mathscr{P}\mathscr{P}' = \mathscr{P}'\mathscr{P} = 0 \tag{A.13}$$

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

$$\Xi_{\mu\nu}(t) = (A_\mu(t), A_\nu(0)) \tag{A.14}$$

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

$$\mathscr{P}A_\mu(t) = \sum_\nu \Xi_{\mu\nu}(t) A_\nu(0) \tag{A.15}$$

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

$$A'_\mu(t) = \mathscr{P}'A_\mu(t) \tag{A.16}$$

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

$$A_\mu(t) = \sum_\nu \Xi_{\mu\nu}(t) A_\nu(0) + A'_\mu(t) \tag{A.17}$$

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

$$\frac{d}{dt}A'_\mu(t) = \mathscr{P}' i\mathscr{L} A'_\mu(t) + \sum_\nu \Xi_{\mu\nu}(t) \mathscr{P}' i\mathscr{L} A_\nu(0) \tag{A.18}$$

The solution of (A.18) becomes

$$A'_\mu(t) = \sum_\nu \int_0^t \Xi_{\mu\nu}(s) \Gamma_\nu(t-s) ds \tag{A.19}$$

$$\Gamma_\nu(t) = exp\left[t\mathscr{P}'iL\right]\mathscr{P}'iLA_\nu(0) \tag{A.20}$$

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

$$\frac{d}{dt}A_\mu(t) = \sum_\nu \Xi_{\mu\nu}(t)iLA_\nu(0) + iLA'_\mu(t) \tag{A.21}$$

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

$$i\Omega_{\mu\nu} \equiv (iLA_\mu(0), A_\nu(0)) \tag{A.22}$$

$$M_{\mu\nu}(t) \equiv -(iL\Gamma_\mu(0), A_\nu(0)) \tag{A.23}$$

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

$$\frac{d}{dt}\Xi_{\mu\nu}(t) = \sum_\tau \Xi_{\mu\tau}(t)i\Omega_{\tau\nu} - \sum_\tau \int_0^t \Xi_{\mu\tau}(s)M_{\tau\nu}(t-s)ds \tag{A.24}$$

Representing in a form of a matrix,

$$\frac{d}{dt}\Xi(t) = \Xi(t)\cdot i\Omega - \int_0^t \Xi(s)\cdot M(t-s)ds \tag{A.25}$$

The Laplace transformation of (A.25) becomes

$$-\hat{1} + z\Xi(z) = \Xi(z)\cdot i\Omega - \Xi(z)\cdot M(z) \tag{A.26}$$

Therefore

$$\Xi(z) = \frac{\hat{1}}{z - i\Omega + M(z)} \tag{A.27}$$

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

$$A(z) = \Xi(z)\cdot\{A(0) + \Gamma(z)\} \tag{A.28}$$

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

$$\{z - i\Omega + M(z)\}\cdot A(z) = A(0) + \Gamma(z) \tag{A.29}$$

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:

$$\frac{d}{dt}A_\mu(t) = \sum_\nu i\Omega_{\mu\nu}A_\nu(t) - \sum_\nu \int_0^t M_{\mu\nu}A_\nu(t-s)ds + \Gamma_\mu(t) \tag{A.30}$$

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.

References

Andreu Mas-Colell, Michael D. Whinston and Jerry R. Green, Microeconomic Theory, Oxford University Press, 1995.

Deaton, A. and Muellbauer J., Economics and Consumer Behavior, Cambridge University Press, 1980.

Debreu, G., Excess-demand functions, J. Mathematical Economics, 1, 15–21, 1974.

Feller, W., An introduction to probability theory and its applications, Wiely, vol. 2, 1966.

Härdle W., Hildenbrand W. and Jerison M., Empirical Evidence on the Law of Demand. Econometrica, 59, No.6, 1525-1549, 1991.

Hildenbrand, W., On the "Law of Demand", Econometrica, 51, No.4, 997-1019, 1983.

Hildenbrand, W. and Jerison M., The demand theory of the weak axioms of revealed preference, Economics Letters, 29(3), 209-213, 1989.

Hildenbrand, W. and Kneip, Family expenditure data, heteroscedasticity and the Law of Demand, Ricerche Economiche 47, 137-165, 1993.

Hildenbrand W., Market Demand: Theory and Empirical Evidence, Princeton Press, 1994.

Hildenbrand, W. and Kneip, A., Demand Aggregation under Structural Stability, J. Mathematical Economics, 31, 81-109, 1999.

Hoderlein S., How many consumers are rational?, J. Econometrics, 164, 294-309, 2011.

Houthakker, H. S. and Taylor, L. D., Consumer Demand in the United States: Analyses and Projections, Second Edition, Harvard University Press, 1970.

Jerison M., Aggregation and Pairwise Aggregation of Demand When the Distribution of Income is Fixed, J. Economic Theory, 33, 1-31, 1984.

Jespersen, S., Metzler, R., and Fogedby H. C., Lévy flights in external force fields: Langevin and fractional Fokker-Planck equations and their solutions, Phys. Rev. E., 59 (3), 2736-2745, 1999.

Khintchine, A. Ya., Limit Laws for Sums of Independent Random Variables, ONTI, Moscow-Leningrad (Russian), 1938.

Kirman, A. P., Whom or What Does the Representative Individual Represent?, he Journal of Economic Perspectives, 6 (2) , 117-136, 1992.

A Perspective on Modern Business Cycle Theory, Kiyotaki N., Economic Quarterly, 97 (3), 195–208, 2011

Kubo R., The fluctuation dissipation theorem, Rep. Prog. Phys., 29, 255-284, 1966.

Lévy, P., Théorie de l'addition des variables aléatoires, Gauthier-Villars, Paris., 1937, 1954.

Mantel, R., On the characterization of aggregate excess-demand, J. Economic Theory, 7, 348–353, 1974.

Metzler, R. and Klafter, J., The random walk's guide to anomalous diffusion: a fractional dynamics approach, Physics Reports, 339, 1-77, 2000.

Mori, H, Transport, Collective Motion, and Brownian Motion, Prog. Theor. Phys., 33 (3), 423-455, (1965)

Pawula, R. F., Approximation of the Linear Boltzman Equation by the Fokker-Planck Equation, Phys. Rev., 162 (1) 186-188, 1967.

Prabhakar T. R., A singular integral equation with generalized Mittag-Leffler function in the kernel., Yokohama Math. J., 19, 7-25, 1971.

Sonnenschein, H., Do Walras' Identity and Continuity Characterize the Class of Community Excess Demand Functions?, J. Economic Theory, 6, 345–54, 1973.

Takayasu H., Fractals in the Physical Sciences, Manchester University Press, 123-128, 1990.

Viñales A. D. and Despósito M. A., Anomalous diffusion induced by a Mittag-Leffler correlated noise, Phys. Rev. E, 75, 042102, 2007