Non-Self-Averaging in Macroeconomic Models: A Criticism of Modern Micro-founded Macroeconomics

Masanao Aoki*
Department of Economics
University of California, Los Angeles
Fax Number 1-310-825-9528, aoki@econ.ucla.edu;
and
Hiroshi Yoshikawa
Faculty of Economics, University of Tokyo
Fax Number 81-3-5841-5521, yoshikawa@e.u-tokyo.ac.jp

June, 2007

Abstract

Using a simple stochastic growth model, this paper demonstrates that the coefficient of variation of aggregate output or GDP does not necessarily go to zero even if the number of sectors or economic agents goes to infinity. This phenomenon known as non-self-averaging implies that even if the number of economic agents is large, dispersion can remain significant, and, therefore, that we can not legitimately focus on the means of aggregate variables. It, in turn, means that the standard microeconomic foundations based on the representative agent has little value for they are expected to provide us with dynamics of the means of aggregate variables.

The paper also shows that non-self-averaging emerges in some representative urn models. It suggests that non-self-averaging is not pathological but quite generic. Thus, contrary to the main stream view, micro-founded macroeconomics such as a dynamic general equilibrium model does not provide solid micro foundations.

Key Words: Non-self averaging phenomena, Power laws, Micro foudnations.

*The first author is grateful for many helps he received from Prof. M. Sibuya.

1 Introduction

The contribution of the literature on endogenous growth ranging from Romer (1986) and Lucas (1988) to Grossman and Helpman (1991), and Aghion and Howitt (1992), has been to *endogenize* the underlying sources of sustained growth in per-capita income. The main analytical exercises in these papers are to explicitly consider the optimization by the representative agent in such activities as education, on-the-job training, basic scientific research, and process and product innovations. This approach is not confined to the study of economic growth, but actually originated in the theory of business cycles. Arguably, the rational expectations model by Lucas (1972, 73) opened the door to modern "micro-founded" macroeconomic theory. In the field of the theory of business cycles, it is now represented by the real business cycle theory (Kydland and Prescott (1982)):

> "Real business cycle models view aggregate economic variables as the outcomes of the decisions made by many individual agents acting to maximize their utility subject to production possibilities and resource constraints. As such, the models have an explicit and firm foundation in microeconomics. (Plosser, 1989, p.53)."

This is the basic tenor which applies not only to the theory of busienss cycles, but also to the endogenous growth literature, or for that matter to the whole macroeconomic thoery. Lucas (1987) declared against the old macroeconomics.

> "The most interesting recent developments in macroeconomic theory seem to me describable as the reincorporation of aggregative problems such as inflation and the business cycle within the general framework of "microeconomic" theory. If these developments succeed, the term 'macroeconomic' will simply disappear from use and the modifier 'micro' will become superfluous. We will simply speak, as did Smith, Ricardo, Marshall and Walras, of *economic* theory (Lucas, (1987; p.107-108))."

In this paper, we argue that this research program which prevails in modern macroeconomics is misguided.

Whether in growth or business cycle models, the fundamental cause for often complex optimization exercises is that they are expected to lead us to our better understanding dynamics of the *means* of *aggregate* variables. The standard model thus begins with the analysis of the optimization of the representative agent, and translates it into the analysis of the economy as a whole.

Economists doing these exercises are, of course, well aware that economic agents differ, and that they are subject to idiosyncratic (or microeconomic)

shocks. As we will observe in the final section, idiosyncratic shocks are indeed the key factor in Lucas (1972, 73)'s theory of business cycles. However, their analyses premise that those microeconomic shocks and differences cancel out each other, and that the behaviors of aggregate variables are represented by their means which, in turn, can be well captured by the analysis based on the representative agent.

The point is best illustrated by the Poisson model which is so widely used in economics ranging from labor search theory to endogenous growth models (*e.g.* Aghion and Howitt (1992)). Suppose that the Poisson parameter is λ which designates the instantaneous probability that an "event" such as technical progress and job arrival occurs. This probability which pertains to one economic agent is assumed to *commonly apply to all the agents and also exogenously gien* —— the crucial assumption! Then, given the same Poisson process with parameter λ for each individual agent, we obtain the Poisson process with the parameter λN for the economy as a whole where there are N economic agents. The mean and the standard deviation of the number of "events" in the macroeconomy are λN and $\sqrt{\lambda N}$, respectively. The coefficient of variation defined as the standard deviation divided by the mean, is, therefore, $\sqrt{\lambda N}/\lambda N = 1/\sqrt{\lambda N}$. Thus, in the Poisson model, when the number of economic agents N becomes large ($N \to \infty$), the coefficient of variation approaches zero. This property known as *self-averaging* provides us with justification for our concentrating on the means of variables in *macro* models; The macroeconomy certainly consists of a large number of economic agents. Now, because the mean depends basically on λ, it is natural to explore how λ is determined in models. Indeed, in standard models, λ is endogenously determined as an outcome of economic agents' optimization and market equilibrium. The Poisson model is just an example. We all know that a considerable part of every main stream macroeconomics paper is now devoted to this kind of micro optimization exercise.

So far, so fine. There is, however, an important point that the standard Poisson model tacitly presumes the representative agent; Economic agents are homogenous in that they face the *same unchanged* instantaneous probability that an "event" occurs to them. Microsoft and small grocery store on the street face "idyosyncratic" or micro shocks which come from the *same* probability distribution! This crucial assumption is also made in the well-known rational expectations model of Lucas (1972, 73), or for that matter in all micro-founded macroeconomic models. When we drop this crucial assumption, we realize that the standard microeconomic foundations for macroeconomics are actually wholly misguided.

Specifically, using simple stochastic models, this paper demonstrates that a tacit and yet the fundamental assumption underlying endogenous growth and real business cycle theories, namely the law of large numbers, is not generally tenable. We show that even if the number of economic agents is large, the behavior of the macroeconomy can not be generally well approximated by the means. The implication is that analyses based on the representative

agent which generate the means of stochastic time paths of aggregate variables, have little value. Put it another way, the standard mico-foundations are not actually true micro-foundations.

Before we proceed to the model, we explain "non-self-averaging," the crucial concept for our purpose. The term "non-self-averaging" is extensively used in the physics literature (see Sornette (2000, p.369)), but is not known in economics. "Non-self-averaging" means that a size-dependent (*i.e.* "extensive" in physics) random variable X of the model has the coefficient of variation that does not converge to zero as model size goes to infinity. The coefficient of variation *(C.V.)* of an extensive random variable, X, defined by

$$C.V.(X) = \frac{\sqrt{variance(X)}}{mean(X)}.$$

is normally expected to converge to zero as model size (e.g. the number of economic agents) goes to infinity. In this case, the model is said to be "self-averaging." We have already shown that the popular Poisson model has this self-averaging property. However, in many models, we are led to non-self-averaging.

The notion of non-self-averaging is important because non-self-averaging models are sample dependent, and some degree of impreciseness or dispersion remains about the time trajectories even when the number of economic agents go to infinity. This implies that focus on the mean path behavior of macroeconomic variables is not justified. It, in turn, means that sophisticated optimization exercises which provide us with information on the means have little value.

In what follows, we first demonstrate this point using the two-parameter Poisson-Dirichlet model. We next show that based on urn models, non-self-averaging is not confined to a particular model which we present in the next section, but is actually quite generic. The final section offers concluding discussion on the implications of non-self-averaging for macroeconomics.

2 Non-self-averaging in a Growth Model

In this section, we present a simple innovation driven growth model in which aggregate output or GDP is non-self-averaging.

The Model

Following the literature on endogenous growth, we assume that the economy grows by innovations. Innovations are shochastic events. There are two kinds of innovations in our model. Namely, an innovation, when it occurs, either raises productivity of one of the existing sectors, or creates a new sector. Thus, the number of sectors is not given, but increases over time.

By the time nth innovation occurs, the total of K_n sectors are formed in the economy wherein the i-th sector has experienced n_i innovations ($i = 1, 2, \ldots, K_n$). By definition, the following equality holds:

$$n_1 + n_2 + \cdots + n_k = n \tag{1}$$

when $K_n = k$. If n-th innovation creates a new sector (secotor k), then $n_k = 1$.

The aggregate output or GDP when n innovations have occured is denoted by Y_n. Y_n is simply the sum of outputs in all the sectors, y_i.

$$Y_n = \sum_{i}^{K_n} y_i. \tag{2}$$

Output in secotr i grows thanks to innovations which stochastically occur in that sector. Specifically, we assume

$$y_i = \eta\gamma^{n_i}. \qquad (\eta > 0, \gamma > 1) \tag{3}$$

For our purpose, it is convenient to rewrite equation (1) as follows.

$$n = \sum_{j}^{n} j a_j(n) \tag{4}$$

In equation (4), $a_j(n)$ is the number of sectors where j innovations have occurred. The vector $a(n)$ consisting of $a_j(n)$, is called *partition vector*[1]. With this partition vector, $a(n)$, K_n can be expressed as

$$K_n = \sum_{j}^{n} a_j(n). \tag{5}$$

Using the following approximation

$$\gamma^{n_i} = exp(n_i \ln \gamma)) \approx 1 + \ln(\gamma) n_i,$$

we can rewrite equation (3) as

$$y_i = \eta + \eta \ln(\gamma) n_i. \tag{6}$$

Thus, from equations (1), (2), (4), (5) and (6), we obtain

$$Y_n \approx K_n + \beta \sum_{j}^{n} j a_j(n). \tag{7}$$

where $\beta = \ln(\gamma) > 0$. Here, without loss of generality, we assume that η is one. Obviously, the behavior of the aggregate output, Y_n depends on how innovations occur.

[1]See chapter 2 of Aoki and Yoshikawa (2006) for partition vector.

The Poisson-Dirichlet Distribution of Innovations

We now describe how innovations stochastically occur in the model. An innovation follows *the two-parameter Poisson-Dirichlet (PD) distribution.*[2]

Given the two-parameter *PD* (α, θ) distribution, when there are k clusters of sizes n_i, $(i = 1, 2, \ldots, k)$, and $n = n_1 + n_2 + \cdots + n_k$, an innovation occurs in one of the existing sectors of "size" n_i with probability rate p_i:

$$p_i = \frac{n_i - \alpha}{n + \theta}. \tag{8}$$

The "size" of sector i, n_i is equal to the number of innovations that have already occurred in sector i. The two parameters α and θ satisfy the following conditions:

$$\theta + \alpha > 0, \quad \text{and} \quad 0 < \alpha < 1.$$

With $\alpha = 0$ there is a single parameter θ, and the distribution boils down to the one-parameter *PD* distribution, $PD(\theta)$.

p_i is the probability that an innovation occurs in one of the existing sectors. Now, a new sector emerges with probability rate[3] p:

$$p = 1 - \sum_{1}^{k} \frac{n_i - \alpha}{n + \theta} = \frac{\theta + k\alpha}{n + \theta}. \tag{9}$$

It is important to note that in this model, sectors are *not* homogeneous with respect to the probability that an innovation occurs. The larger sector i is, the greater the probability that an innovation occurs in sector i becomes. Moreover, there probabilities *change endogenously* as n_i changes over time.

In the two-parameter $PD(\alpha, \theta)$ distribution, the probabilitiy that the number of sectors increases by one in $n + 1$ conditional on $K_n = k$, is given

[2]Kingman invented the one-parameter Poisson-Dirichlet distribution to describe random partitions of populations of heterogeneous agents into distinct clusters. The one-parameter Poisson-Dirichlet model is also known as Ewens model, (Ewens (1972)); See Aoki (2000a, 2000b) for further explanation. The one-parameter model was then extended to the two-parameter Poisson-Dirichlet distributions by Pitman; See Kingman (1993), Carlton (1999), Feng and Hoppe(1998), Pitman (1999, 2002), and Pitman and Yor (1996), among others. Aoki (2006) has shown that the two-parameter Poisson-Dirichlet models are qualitatively different from the one-parameter version because the former is not self-averaging while the latter is. These models are therefore not exponential growth models familiar to economists but they belong to a broader class of models without steady state constant exponential growth rate. None of the previous works, however, have comparatively examined the asymptotic behavior of the coefficient of variation of these two classes of models.

[3]Probabilities of new types entering Ewens model, are discussed in Aoki (2002, Sec.10.8, App. A.5).

by[4]

$$\Pr(K_{n+1} = k+1 | K_1, \ldots, K_n = k) = p = \frac{\theta + k\alpha}{n+\theta}. \tag{10}$$

On the other hand, the corresponditing probability that the number of sectors remains unchanged is

$$\Pr(K_{n+1} = k | K_1, \ldots, K_n = k) = \sum_i p_i = \frac{n - k\alpha}{n+\theta}. \tag{11}$$

We show that this two-parameter *PD* model is non-self averaging. It is interesting to observe that the one parameter *PD* model ($\alpha = 0$) is self-averaging. Before se proceed, it may be helpful to say a few words why the two-parameter *PD* model is non-self averaging. The answer lies in (10) and (11). In this model, innovations occur in one of the two different types of sectors, one, the new type and the other, known or pre-existing types. The probability that an innovation generates a new sector is $(\theta + K_n\alpha)/(n+\theta)$, and the probability that an innovation occurs in one of the existing sectors is $(n - K_n\alpha)/(n+\theta)$, where K_n is the number of types of sectors in the model by the time n innovations occurred. These probabilities and their ratio vary **endogenously**, depending on the histories of how innovations occured. In other words, the mix of old and new sectors evolve endogenously, and is path-dependent. This is the reason why non-self averaging emerges in the two parameter *PD* model. We note that in one parameter *PD* model in which $\alpha = 0$, two probabilities (10) and (11) become independent of K_n, and that the model becomes self-averaging.

Now, the standard endogenous growth literature focuses on profit motives for innovations. The name "endogenous growth" comes from explicit analysis of innovations as outcomes of profit-seeking activities. There is no denying that innovations are at least partly outcomes of intentional profit-seeking activities. However, we skip such analysis in the present analysis. The basic reason is that aggregate output, Y_n is non-self averaging. To explain this point in detail is, in fact, the purpose of this paper.

GDP is Non-self Averaging

Given the model, we are interested in the behavior of GDP, namely Y_n. Specifically, we would like to see whether or not Y_n is self-averaging. Toward this goal, we first normalize Y_n by n^α. Then, from equation (7), we obtain

[4]Because the following inequality holds:

$$\frac{\theta + k\alpha}{n+\theta} > \frac{\theta}{n+\theta},$$

we observe that the probability that a new sector emerges is higher in the two-parameter *PD* model than in the one-parameter *PD* model.

$$\frac{Y_n}{n^\alpha} = \frac{K_n}{n^\alpha} + \beta \Sigma_j^n \frac{a_j(n)}{n^\alpha}. \tag{12}$$

In what follows, we show that Y_n is non-self-averaging. Toward this goal, we first define partial sums of K_n and Y_n up to $l(< n)$, $K_n(1,l)$ and $Y_n(1,l)$, as follows:

$$K_n(1,l) = \Sigma_{j=1}^{l} a_j(n) \tag{13}$$

and

$$Y_n(1,l) = K_n(1,l) + \beta \hat{Y}_n(1,l) \tag{14}$$

where

$$\hat{Y}_n(1,l) = \Sigma_{j=1}^{l} j a_j(n). \tag{15}$$

Yamato and Sibuya (2000; p.7 their prop. 4.1 and 4.2) showed that given l, $K_n(1,l)/n^\alpha$ and $\hat{Y}_n(1,l)/n^\alpha$ converge in distribution as n approaches infinite ($\longrightarrow d$) as follows:

$$\frac{K_n(1,l)}{n^\alpha} \longrightarrow^d C_1(l) L \tag{16}$$

and

$$\frac{\hat{Y}_n(1,l)}{n^\alpha} \longrightarrow^d C_2(l) L \tag{17}$$

where

$$C_1(l) = 1 - \frac{(1-\alpha)^{[l]}}{l!}$$

$$C_2(l) = \frac{(2-\alpha)^{[l-1]}}{(l-1)!}.$$

Here, $[j]$ in $C_1(l)$ and $C_2(l)$ denotes an ascending factorial:

$$x^{[j]} = x(x+1)\dots(x+j-1).$$

The random variable L in (16) and (17) has the probability density function $g_{\alpha,\theta}(x)$:

$$g_{\alpha\theta}(x) = \frac{\Gamma(\theta+1)}{\Gamma(\theta/\alpha+1)} x^{\frac{\theta}{\alpha}} g_\alpha(x) \tag{18}$$

where g_α is the density of the Mittag-Leffler distribution[5] with parameter α. Pitman (1999) also showd the a.s.convergence. See Yamato and Sibuya (2000, p.8).

It is shown by Yamato and Sibuya and by Pitman that

$$\frac{K_n}{n^\alpha} \longrightarrow^d L, \tag{19}$$

$$\frac{K_n}{n^\alpha} \longrightarrow L \qquad a.s. \tag{20}$$

and

$$C.V.\left(\frac{K_n}{n^\alpha}\right) \longrightarrow C.V.(L) > 0. \tag{21}$$

Because $C_1(l)$ and $C_2(l)$ are constant in (16) and (17), for each fixed l, and $\alpha > 0$, we obtain

$$C.V.\left(\frac{K_n(1,l)}{n_\alpha}\right) \longrightarrow C.V.(L) > 0, \tag{22}$$

and

$$C.V.\left(\frac{(\hat{Y}_n(1,l)}{n^\alpha}\right) \longrightarrow C.V.(L) > 0. \tag{23}$$

Therefore, given (14), we obtain

$$C.V.\left(\frac{(Y_n(1,l)}{n^\alpha}\right) \longrightarrow C.V.(L) > 0 \tag{24}$$

Thus, for sufficiently large l,

$$C.V.\left(\frac{Y_n}{n^\alpha}\right) \longrightarrow C.V.(L). \tag{25}$$

Mittag-Leffler function $g_\alpha(x)$ has the property that its p th moment is given by

$$\int_0^\infty x^p g_\alpha(x)dx = \frac{\Gamma(p+1)}{\Gamma(p\alpha+1)} \qquad (p > -1). \tag{26}$$

Thus, using (18) and (26), we can obtain the first and second moments of L, $E_{\alpha,\theta}(L)$ and $E_{\alpha,\theta}(L^2)$ as follows:

$$E_{\alpha,\theta}(L) = \frac{\Gamma(\theta+1)}{\alpha\Gamma(\theta+\alpha)}, \tag{27}$$

[5]See Blumenfeld and Mandelbrot (1997), Erdely, A., W.Magnus, F Oberhettinger, F.G. Tricomi (1953-1954), or Pitman (1999) on Mittag-Leffler function.

and

$$E_{\alpha,\theta}(L^2) = \frac{(\theta+\alpha)\Gamma(\theta+1)}{\alpha^2\Gamma(\theta+2\alpha)}. \tag{28}$$

The variance of $L, var(L)$ is, therefore,

$$var(L) = E_{\alpha,\theta}(L^2) - [E_{\alpha,\theta}(L)]^2 = \gamma_{\alpha,\theta}\frac{\Gamma(\theta+1)}{\alpha^2}, \tag{29}$$

where

$$\gamma_{\alpha,\theta} := \frac{\theta+\alpha}{\Gamma(\theta+2\alpha)} - \frac{\Gamma(\theta+1)}{[\Gamma(\theta+\alpha)]^2}. \tag{30}$$

The coefficient of variation of L is given by

$$C.V.(L) = \frac{\sqrt{var(L)}}{E_{\alpha,\theta}(L)} = \sqrt{\frac{\gamma_{\alpha,\theta}}{\Gamma(\theta+1)}}\Gamma(\theta+\alpha). \tag{31}$$

Note that $\gamma_{\alpha,\theta}$ defined by (30) is zero when $\alpha = 0$, but that it is positive when $\alpha > 0$. Therefore, $C.V.(L)$ is zero in the one-parameter PD model ($\alpha = 0$), but is positive in the two-parameter PD model ($\alpha > 0$).

Now, we have shown above that $C.V.(Y_n/n^\alpha)$ converges to $C.V.(L)$. Thus, thanks to (25) and (31), we finally obtain

$$C.V.\left(\frac{Y_n}{n^\alpha}\right) \to \sqrt{\frac{\gamma_{\alpha,\theta}}{\Gamma(\theta+1)}}\Gamma(\theta+\alpha). \tag{32}$$

The right-hand side of (32) does not approach zero even if n goes to infinity in the two-parameter PD model ($\alpha > 0$). Thus, we have established the following proposition.

Proposition
In the two-parameter Poisson-Dirichlet model, the aggregate output Y_n is non-self averaging.

3 Non-self-averaging in Triangular Urn Models

In the previous section, we considered a simple innovation-driven growth model in which GDP is non-self-averaging. Stochastic events are not confined to innovations, of course. To name but a few, job offer, any profit opportunity, and discovery of new resources are all stochastic. No wonder, modern macroeconomics —— rational expectations models, real business cycle theory, labor search theory, and endogenous growth theory —— explicitly

takes into account stochastic "shocks" in respective models. Most models can be interpreted as a variety of stochastic processes.

Now, many stochastic processes can be interpreted as urn models. For example, a random walk model on a finite set of states whereby the boundary states are reflecting is equivalent to the famous Ehrenfest urn model. More generally, by drawing balls not uniformly but at random times governed by exponential distribution, urn models can be reinterpreted as stochastic processes as shown by Athrea and Karlin (1968).

An important characteristic of urn models is that such processes are *path dependent*. Feller (1968; p.119) calls path-dependence "aftereffect." Namely, he says that "it is conceivable that each accident has an *aftereffect* in that it either increases or decreases the chance of new accidents." An obvious example would be contagious diseases. Indeed, in an classical paper by Eggenberger and Polya (1923), an urn model was introduced to describe contagious diseases; In Polya's urn scheme, the drawing of either color increases the probability of the same color at the next drawing, and we are led to such path-dependence as seen in contagious diseases. We can easily conceive path-dependent phenomena in economics. They can be described by urn models. For example, Winter, Kaniovski and Dosi (2000) analyze industrial dynamics with innovative entrants by urn models. In this section, we show that a class of urn models lead to non-self averaging. They are meant to demonstrate that non-self-averaging is not pathological but is generic.

Balanced Triangular Urn Models

Using the scheme of Flajolet Gabarro and Pkari (2005), we describe urns with two types of balls, black and white. The balls may be interpreted as sectors or innovations. The color of balls represent different kinds. The interpretation is quite flexible.

We can describe this urn model by *the replacement matrix* M. Specifically, we use a 2×2 triangular matrix M, with elements $m_{1,1} = a > 0$, $m_{1,2} = b - a$, $b > a$, $m_{2,1} = 0$, and $m_{2,2} = b$. This matrix M specifies that if a black ball (ball 1) is drawn, it is returned to the urn together with a additional black balls, and $b - a$ white balls. If a white ball (ball 2) is drawn, then it is returned to the urn together with b white balls; No white ball is added in this case, and therefore, the replacement matrix M is triangular. The urn is called balanced because the two row sums of M are equal (both equal to b). It means that the total number of balls in the urn is the same regardless of the color of a ball drawn.

In what follows, we show that the stochastic process described by this urn model is non-self-averaging. Non-self-averaging is caused by the fact that the generating mechanism, that is the mix of balls of two types is path-dependent for the same reason as the *PD* model in the preceeding section. Note that in this model, the ratio of black and white balls is path-dependent, and varies endogenously; The number of balls of each types being put into

the urn clearly depends on the way two types of balls have been drawn in the past.

Suppose that there are r black balls and s white balls after n draws. This composition of the types of balls is represented by a monomial as $u^r v^s$. Then, there are altogether $a_0 + b_0 + b \times n$ balls in the urn, where a_0 and b_0 are the numbers of initial black and white balls; Recall that the urn is balanced. Now, with r black balls in the urn, there are r ways of picking a black ball, each such draw results in a additional black balls and $b - a$ additional white balls. Therefore, after a draw of a black ball, the monomial $u^r v^s$ is transformed into $ru^{r+a}v^{s+b-a}$. Likewise, a draw of a white ball changes the monomial into $su^r v^{s+b-a}$. This evolution is represented by the following operator Γ:

$$\Gamma = u^{a+1}v^{b-a}\frac{\partial}{\partial u} + v^{b+1}\frac{\partial}{\partial v}. \tag{33}$$

All possible compositions of this urn at time n is represented by a polynomial in u and v, $f_n(u, v)$. Using the operator Γ defined by (33), we have

$$f_{n+1}(u, v) = \Gamma f_n(u, v). \tag{34}$$

By defining the exponential generating function

$$H(z, u, v) = \sum_{n\geq 0} f_n(u, v)z^n/n!, \tag{35}$$

we obtain its first-order partial differential equation

$$u^{a+1}v^{b-a}\frac{\partial H}{\partial u} + v^{b+1}\frac{\partial H}{\partial v} = \frac{\partial H}{\partial z}. \tag{36}$$

This equation can be solved by the method of characteristics, see Aoki (2002, A.1), for example. The partial differential equation (36) is converted into a set of ordinary differential equations.

$$du/dt = u^{a+1}v^{b-a},$$

$$dv/dt = v^{b+1},$$

and

$$dz/dt = -1.$$

Eliminating dt from the above, we obtain

$$\frac{dv}{v^{b+1}} = -dz \tag{37}$$

and

$$\frac{du}{u^{a+1}} = \frac{dv}{v^{a+1}} \tag{38}$$

The equation for v can be integrated directly. Then the other equation is integrated yielding two constants of integration. The general solution is a function of these two constants of integration. To be concrete, suppose that $a = 1$ and $b = 2$. Then, we obtain the first integral as follows:

$$\phi_1(z, u, v) = z - \frac{1}{2v^2} \tag{39}$$

and

$$\phi_2(z, u, v) - \frac{1}{v} - \frac{1}{u}. \tag{40}$$

Hence,

$$H(z, u, v) = h\left(z - \frac{1}{2v^2}, \frac{1}{v} - \frac{1}{u}\right) \tag{41}$$

where

$$H(0, u, v) = un. \tag{42}$$

With this generating function, $H(z, u, v)$, we can obtain the probability distribution of X_n, the number of black balls at time n. Note that because the urn is balanced, and the total number of balls at n is not random, once we know the number of black balls, we automatically know the number of white balls as well. Puyhaubert (2005) establishes the following results:

$$E(X_n) = \frac{a\Gamma\left(\frac{a_o+a}{a}\right)\Gamma\left(\frac{t_o}{b}\right)}{\Gamma\left(\frac{a_o}{a}\right)\Gamma\left(\frac{t_o+a}{b}\right)} n^{\frac{a}{b}} + 0(1) \tag{43}$$

and

$$E(X_n^2) = \frac{a^2\Gamma\left(\frac{a_o+2a}{a}\right)\Gamma\left(\frac{t_o}{b}\right)}{\Gamma\left(\frac{a_o}{a}\right)\Gamma\left(\frac{t_o+2a}{b}\right)} n^{\frac{2a}{b}} + 0(n^{\frac{a}{b}}). \tag{44}$$

As we have seen it in the previous section, we can show that $C.V.(X_n)$ remains positive even if n approaches infinite. Thus, we have established the following proposition.

Proposition: The number of black balls in the balanced triangular urn model is non-self-averaging.

Non-Balanced Triangular Urn Models

Janson (2006) examins triangular urns which are not balanced. Specifically, Janson (2006; Theorem 1.3) derives that when the replacement matrix M consists of $m_{1,1} = a = c + d$, $m_{1,2} = c$, $m_{2,1} = 0$, and $m_{2,2} = d$, we obtain

$$n^{-d/a} X_n \to^d W, \tag{45}$$

where X_n is the number of black balls, and n is the number of drawings. The variable W has a generalized Mittag-Leffler distribution with moments

$$E(W/d)^p = \frac{\Gamma((a_0+b_0)/a)\Gamma(a_0/d+p)}{\Gamma(b_0/d)\Gamma(a_0+b_0+d/a)}. \qquad (p=1,2,\ldots.) \qquad (46)$$

By identifying two parameters θ and α in the PD model in the following way:

$$\frac{a_0+b_0}{a} = \theta + 1, \qquad (47)$$

and

$$\frac{b_0}{d} = \frac{\theta}{\alpha}, \qquad (48)$$

we can observe that these two Mittag-Leffler moment expressions are the same as L in the two-parameter PD model presented in the previous secdtion. This fact means that two distributions are identical because the moments of Mittag-Leffler distributions uniquely determine the distribution (Bingham, Goldie, Teugels (1987, 391)). Janson (2006; Theorem 1.3) shows that depending on parameters of the replacement matrix, namely a, c, and d, X_n, the number of black balls, becomes non-self-averaging[6].

We can summarize this analysis as follows.

Proposition: In non-balanced triangular urn models, depending on the values of parameters, non-self-averaging emerges. Non-self-averaging is generic in the sense that a set of parameters for which non-self-averaging emerges is not of measure zero.

4 Concluding Discussion

Almost all the economic opportunities such as job offer, discoveries of new technology, market, and resources are stochastic. Modern micro-founded macroeconomics —— Lucas' rational expectations model, real business cycle theory, labor search theory, and endogenous growth theory —— rightly takes into account stochastic events. However, in these micro-founded models, it is taken for granted that as the number of agents goes to infinity, any

[6]Janson (2005; Theorem 1.3) shows that the behavior of the model is determined by the relataive sizes of the eigenvalues of the replacement matrix M. In the triangular urn, $\lambda_1 = max(a,d)$ and $\lambda_2 = min(a,d)$. With $\lambda_2 \leq \lambda_1/2$, the composition of the urn is effectively determined by the outcome of the large number of later draws, each having a negligible effect. On the other hand, if $\lambda_2 \geq \lambda_1/2$, the imbalance caused by the first random draw magnifies at a sufficient rate to remain important for large number draws, n. Because the element (1,2) is zero, type 1 balls do not affect type 2 balls and no smoothing effects on the type 2 balls are caused by the large number of type 1 draws. This explains why no normal distributions emerge for d less than $a/2$.

micro or "idyosyncratic" fluctuations vanish, and that well defined deterministic macroeconomic relations prevail. That is, self-averaging is tacitly presumed. However, if model is non-self-averaging, dispersion remains even if the number of economic agents become infinite. It means that we can not focus on the means of macro variables. This, in turn, means that sophisticated optimization and market equilibrium exercises which provide us with dynamics of the means of macro variables have, in fact, little value.

Self-averaging emerges when every agent is assumed to face the *same unchanged* "well-behaved" probability distribution such as the normal and the Poisson distributions. This is indeed the standard assumption in microfounded macroeconomics. Lucas (1972, 73)'s famous model of business cycles is a primary example. It is instructive to trace his model in detail. Lucas begins to model a supplier's behavior in each individual market as follows;

> "Quantity supplied in each market will be viewed as the product of a normal (or secular) component common to all markets and a cyclical component which varies from market to market. Letting z index markets, and using y_{nt} and y_{ct} to denote the *logs* of these components, supply in market z is:
>
> (1) $\quad y_t(z) = y_{nt} + y_{et}(z)$
>
> ..... The cyclical component varies with perceived, *relative* prices and with its own lagged value:
>
> (3) $\quad y_{ct}(z) = \gamma\left[P_t(z) - E(P_t \mid I_t(z))\right] + \lambda y_{c,t-1}(z)$
>
> Where $P_t(z)$ is the actual price in z at t and $E(P_t \mid I_t(z))$ is the mean current, general price level, conditioned on information available in z at t, $I_t(z)$.

Given this framework, he goes on to the information structure of the economy.

> "The information available to suppliers in z at t comes from two sources. First, traders enter period t with knowledge of the past course of demand shifts, of normal supply y_{nt}, and of past deviations $y_{c,t-1}, y_{c,t-2}, \ldots$. While this information does not permit exact inference of the *log* of the current general price level, P_t, it does determine a "prior" distribution on P_t, common to traders in all markets. We assume that this distribution is knwon to be normal, with meand $\bar{P}_t$ (depending in a known way on the above history) and a constant variance σ^2. Second, we suppose that the actual price deviates form the (geometric) economy-wide average by an amount which is distributed independently of P_t. Specifically, let the percentage deviation of the price in z from the average P_t be denoted by z (so that markets are indexed by their

price deviations from average) where z is normally distributed, independent of P_t, with mean zero and variance τ^2. Then the observed price in $z, P_t(z)$ (in *logs*) is the sum of independent, normal variates

$$(4) \qquad P_t(z) = P_t + z$$

The information $I_t(z)$ relevant for estimation of the unobserved (by suppliers in z at t), P_t consists then of the observed price $P_t(z)$ and the history summarized in $\bar{P}_t$ (Lucas(1973; p.328))".

The assumption of rational expectations then permits suppliers in individual markets to make efficient inferences on the relative prices. This leads to micro supply functions. Given micro supply functions, the aggregate supply function, is trivially derived.

"To utilize this information, suppliers use (4) to calculate the distribution of P_t, conditional on $P_t(z)$ and $\bar{P}_t$. This distribution is (by straightforward calculation) normal with mean:

$$(5) \qquad E(P_t \mid I_t(z)) = E(P_t \mid P_t(z), \bar{P}_t) = (1-\theta)P_t(z) + \theta\bar{P}_t$$

where $\theta = \gamma^2/(\sigma^2 + \gamma^2)$, and variance $\theta\sigma^2$. Combining (1), (3), and (5) yields the supply function for market z:

$$(6) \qquad y_t(z) = y_{nt} + \theta\gamma[P_t(z) - \bar{P}_t] + \lambda y_{e,t-1}(z)$$

Averaging over markets (integrating with respect to the distribution of z) gives the aggregate supply function:

$$(7) \qquad y_t = y_{nt} + \theta\gamma(P_t - \bar{P}_t) + \lambda[y_{t-1} - y_{n,t-1}]$$

(Lucas(1973; p.328))"

As is well known, the aggregate supply function is the core of Lucas' rational expectations model of business cycles. In this model, the crucial assumption is his equation (4) above. More specifically, the random variable z is assumed to be normally distribution with mean zero and variance τ^2. That is, each supplier faces the same probability distribution of micro shock although a realization of such a shock, of course, differs across suppliers.

This assumption is taken by most economists as innocuous. However, it actually means that Microsoft and small grocery store on the street face micro shocks drawn from the *same unchanged* probability distribution! It presumes *homogeneity with respect to the probability distribution of micro shocks*, and extremely unrealistic. Lucas' model emphasizes the role of micro shocks which by definition differ across sectors or agents. In this sense, it rejects the representative agent. However, like other micro-founded macro models, it is built on the crucial premise that every agent faces the same

unchanged probability distribution of micro shocks. This assumption entails self-averaging. Specifically, in his model, one can easily obtain the aggregate supply function by "averaging over markets (integrating with respect to the distribution of z.)" Note that his aggregate output y_t (his equation (7)) is nothing but the mean of stochastic aggregate output.

We must discard the assumption that micro agent or sector faces the same unchanged probability distribution of micro shocks. Under the more realistic assumption that each agent or sector faces a different probability distribution of micro shocks, and that such a probability distribution endogenously changes over time, in general, non-self-averaging can emerge. Examples we present in this paper are extremely simple. Despite their simplicity, each sector or agent is assumed to be subject to a different probability distribution of micro shocks, and such a probability distribution endogenously changes over time. In these examples, non-self-averaging emerges.

In fact, income distribution, firms size distribution, distribution of stock price change, and many other distributions of important economic variables are now known to obey power-laws that entail non-self-averaging (See Mantegna and Stanley (2000)). Thus, shouldn't we be ready to expect non-self-averaging in the economy?

Non-self-averaging deprives us of a justification for our focusing on means. It, in turn, means that such sophisticated microeconomic analyses as infinite horizon stochastic dynamic programming which are common in macroeconomic models, and are expected to give us the exact mean time paths of aggregate variables, have, in fact, little value. Those analyses provide us with no foundations for macroeconomic analyses because time paths of macro variables are sample dependent in any way.

Summing up, macroeconomics must seek different microeconomic foundations from the standard optimization of the representative agent (See Aoki and Yoshikawa (2006)). Contrary to the motto of modern micro-founded macroeconomics, it is actually useful to separate macroeconomics from sophisticated optimization exercises. Solow(2000), for example, suggests that we might reasonably separate macroeconomic growth theory from microeconomic analysis of technical progress.

> "It may be too strong a statement, but only a little too strong, to suggest that growth theory "proper" is the study of the long-run behavior of an economy conditional on $A(t)$. But then there is a separate, though closely related, field of study that is concerned with $A(t)$ itself, or more generally with the understanding of the process of technological change. It goes without saying that the results of this second branch of economics will be of central importance to growth theory. One of the advantages of this distinction is that the economics of technical change will certainly involve considerations — about industrial organization, management, practices, and other such things — that have little

> in common with the macroeconomics of growth, but are essential to the economics of technology. (Solow (2000), p.101)"

Non-self-averaging leads us to the same conclusion. The point is actually not confined to the theory of growth, but more widely applies to macroeconomics. Macroeconomics is better freed from too much of optimization exercises. This is the fundamental implication of non-self-averaging for macroeconomics. In this paper, we have demonstrated that once we drop the dubious assumption that agents or sectors face the same probability distribution of micro shocks, non-self-average can emerge.

References

[1] Aghion, P., and P. Howitt, (1992), "A model of growth through creative destruction" *Econometrica*, **60**, 323–351.

[2] Aoki, M. (2006), "Thermodynamic Limits of Macroeconomic or Financial Models: One- and Two-Parameter Poisson-Dirichelt Models", forthcoming *J. Econ.Dyn. Control*.

[3] ——, (2003), "Models with Random Exchangeable Structures and Coexistence of Several Types of Agents in the Long-Run: New Implementations of Schumpeter's Dynamics" Tech. Report, Research Initiative and Development, Chuo Univ. Tokyo, Dec. 2003.

[4] ——, (2002), *Modeling Aggregate Behavior and Fluctuations in Economics: Stochastic Views of Interacting Agents* , New York: Cambridge Univ. Press.

[5] ——-, (2000a), "Open models of share markets with two dominant types of participants, ", *J.Econ. Behav. Org.* **49** 199-216.

[6] —-, (2000b), "Cluster size distributions of economic agents of many types in a market", *J. Math Anal. Appl*, **249**,32-52.

[7] —— (1996), *New Approaches to Macroeconomic Modeling: Evolutionary Stochastic Dynamics, Multiple Equilibria, and Externalities as Field Effects* , New York: Cambridge Univ. Press.

[8] ——, and H. Yoshikawa (2006), *Reconstructing Macroeconomics: A Perspective from Statistical Physics and Combinatorial Stochastic Processes* , New York: Cambridge University Press.

[9] Bingham, N.H., C. M. Goldie, and J. L. Teugels (1987), *Regular variation*, Cambridge, UK: Cambridge Univ. Press.

[10] Blumenfeld, R., and B. B. Mandelbrot, (1997), "Lévy dusts, Mittag-Leffler statistics, mass fractal lacunarity, and perceived dimension," *Phy. Rev. E* **56** 112-118

[11] Charambides, A. C., *Enumerative Combinatorics* , Chapman Hall/CRC, 2002, london

[12] Carlton, M. A. (1999), *Applications of the Two-Parameter Poisson-Dirichlet Distribution* Ph.D. thesis, Dept. Math. Univ. California, Los Angeles

[13] Devroye, L., (1993), "A triptych of discrete distributions related to the stable law," *Statist. Probab. Lett* **18**, 349–351

[14] Eggenberger, F. and G. Polya, (1923), *Über die Statistik verketteter Vorg änge,* Zeitschrift für Angewandte Mathematik and Mechanik, Vol.3.

[15] Erdélyi, A. (ed.) (1955) *Higher Transcendental Functions, vol 3.*, New York: McGraw-Hill.

[16] Ewens, W. J. (1972), "The sampling theory of selectively neutral alleles," *Theor. Pop. Biol.*, **3**, 87-112.

[17] Feller, W. (1968), *An Introduction Probability Theory and Its Applications*, Vol.I, 3rd Edition, New York: John Wiley & Sons.

[18] Feng, S., and F.M. Hoppe, (1998), " Large deviation principles for some random combinatorial structures in population genetics and brownian motion," *Ann. Appl. Probab.* **8**, 975–994.

[19] Grossman, G.M., and E. Helpman, (1991), *Innovation and Growth in the Global Economy,* Cambridge, MA: MIT Press.

[20] Janson, S. (2006), *Limit theorems for triangular urn schemes,* Probability Theory and Related Fields **134**, 417-452.

[21] Karlin, S., and J. McGregor, (1967), "The number of mutant froms maintained in a population," Proc. 5th Berkeley Symp. on Math. Stat. and Probab. Vol. 4.

[22] Kingman, J.F.C. (1978), "The representtion of partition structure", *J. London Math. Soc.*,**18**, 374–380.

[23] ———, (1993), *Poisson Processes*, Oxford UK: Clarendon Press.

[24] Kydland, F. and E. Prescott, (1982), "Time to Build and Aggregate Flucutaions," *Econometrica,* November.

[25] Lucas, R.E., (1972), "Expectation and the Neutrality of Money," *Journal of Economic Theory*, Vol.4, Issue 2, 103-124.

[26] ———, (1973), "Some International Evidence on Output-Inflation Tradeoffs," *The American Economic Review,* Vol. 63, No. 3, 326–334.

[27] ———, (1987), *Models of Business Cycles*, New York: Blackwell.

[28] ———, (1988), "On the Mechanics of Economic Development," *Journal of Monetary Economics,* July.

[29] Mantegna, R.N. and H.E. Stanley (2000), *An Introduction to Econophysics: Correlations and Complexity in Finance,* Cambridge UK: Cambridge University Press.

[30] Pitman, J., (2002), *Lecture Notes of the summer school on probability* , St. Flour, France, forthcoming from Springer-Verlag.

[31] Plosser, C., (1989), "Understanding Real Business Cycles," *Journal of Economic Perspectives,* Summer.

[32] Puyhaubert, V. (2003), "Analytic Urns of Triangular Form," *Algorithms Seminar*, 2002-2004, F. Chyzk (ed.), INRIA, (2005), p.61-64.

[33] Romer, P.M.,(1986), "Increasing Returns and Long Run Growth," *Journal of Political Economy.*

[34] Solow, R. M. (2000), *Growth Theory 2nd ed.*, New York: Oxford University Press.

[35] Sutton, J., (2002), "The variance of firm growth rates: the "scaling puzzle"", *Physica A*, **312**, 577-590.

[36] Winter, S.G., and Y.M. Kaniovski, G. Dosi, (2000), "Modeling Industrial Dynamics with Innovative Entrants," *Structural Change and Economic Dynamics*, Vol.11, 255-293.

[37] Yamato, H., and M. Sibuya, (2000), "Moments of some statistics of Pitman sampling formula", *Bull. Inform. Cybernetics*, **6**, 463-488.

[38] ———(2003), "Some topics on Pitman's probabilistic partition" *Statistical Math.***51**, 351–372. (in japanese)