Mittag-Leffler Distributions and Long-Run Behavior of Some Macroeconomic Models

Masanao Aoki*
Department of Economics
University of California, Los Angeles
Fax Number 310-825-9528, e-mail aoki@econ.ucla.edu

May,2006

Abstract

This paper discusses long-run behavior of economic models with many interacting heterogeneous agents, and point out the connection with the class of Mittag-Leffler distributions.

In the process, the paper summarizes some known asymptotic properties of a class of one- and two-parameter Poisson-Dirichlet distribution models, and those of the model discussed by Feng and Hoppe. These models have also known long-run behavior after some suitably normalized numbers of partitions and the components of partition vectors, such as non-vanishing variances of cluster sizes as the number of agents becomes large. Some differences in the long-run behavior between the class of one-parameter models and that with two-parameters are pointed out. Convergence behavior is expressed in terms of generalized Mittag-Leffler distributions in the statistics literature. We exhibit power laws when they exist as well.

Second, a numerical example of a model which is outside the framework of one- and two-parameter Poisson Dirichlet models mentioned above. This model has more than two parameters but is a simple model composed of two types of agents, innovators and immitators. This model has non-self averaging variaces and the covariance of the sizes of the two sectors, that is the variances and the covariance do not vanish as the number of agents approach infinity.

Key Words:Two-parameter Poisson-Dirichlet distributions; Mittag-Leffler distributions; nom-self averaging phenomena, Power laws.

*The author is grateful for many helps he received from M. Sibuya. Ted Theodosopoulos and an anonymous referee's comments are useful in revising the original draft. I thank them all. The simulation was run by G. Yoshida and K. Ono, Department of Physics, Chuo University, Tokyo.

Introduction

In old industrial organization literature, several tests and measures of degree of industrial concentration have been used to to decide if a given industry is monopolistic or not. See for example Scherer (1980) for various case studies. One such test is called Herfindahl, or Herfindahl-Hirschman index of concentration. It is defined as the sum of squares of fractions of shares, i.e.,

$$H = \sum_i x_i^2,$$

where x_i is the fraction of "share" of markets or sales by sector i or firm i. By definition x_i is positive, and sum to one, $\sum_i x_i = 1$. As we discuss shortly, this literasture used a rudimentary version of the size-biased sampling scheme as a test for oligopoly. This meassure of concentration is used in both domestic and foregin trade context. It is sometimes (mistakenly) called Gini-index.[1]

The question of concentration is that of distribution of fractions of the numbers of clusters, and the numbers of agents by types. A simple application of shares of market by two types of agents, using one-parameter Poisson-Dirichlet distribution (also called Ewens distribution, Ewens (1972, 1979, 1990)) has been made by Aoki (2000a, 2000b).

This paper develops further the original ideas in these papers by applying some of the results from two-parameter Poisson-Dirichlet distributions in the recent combinatorial stochastic process literature, by Kingman (1993), Carlton (1999), Holst (2001), Pitman (1999, 2002), and his associates.

In physics literature, Mekjian and Chase (1997) have used two-parameter models. They refer to the work by Pitman (1996). There are other works in the physics literature, in particular the papers by Derrida-Flyvbjerg mentioned in footnote 1, and Derrida (1994a, 1997).[2] There are other papers in the physics literature that deal with random partitions. Higgs (1995) have noted the similarities of some physical distributions and power laws, and mention population genetics papers by Ewens in particular. Frontera, Goicoechea, Rafols, and Vivies (1995), and Krapivsky, Grosse, and Ben Nadin (2002) discuss partitions and fragmentations, that is, stick-breading version of the residual allocation processes explicitly. They have not touch on connections with the two-parameter Poisson-Dirichlet distributions.

In macroeconomic and finance modelings, agents of different characteristics or strategies are of different types and form separate clusters and affect aggregate behavior. In this paper, we therfore explore more broadly economic implications of long-run relations that may exist among non-self averaging economic or financial variables.

[1]Sometimes it is called Gini-Simpson index of divesity. See Hirschman (1960) about the origin and mis-attribution of this notion to Herfindahl. In the population genetics literature H is called homozygosity. See Ewens (1972). Interestingly, the same measure has been used by Derrida-Flyvbjerk (1989) in discussing relative sizes of basins of attractions of Kaufman random maps and ramdom dynamics in statistics and physics. These, however, involve a sigle parameter θ in their statistical description. See also Aldous (1985).

[2]Derrida (1994b) has added some material on residual allocation models.

In the first part of this paper we introduce the reader to some basic notions on random partitions from the literature of combinatorial stochastic processes, in particular the works by statisticians, J. Pitman (1996, 2002) and Yamato and Sibuya (2000). Size-biased permutation, residual allocation models, notions of frequency spectrum and structure distribution, Mittag-Leffler probability density and power-laws are introduced in the process of describint long-run behavior of models.

We show, among other things, that components of partition vectors in $PD(\alpha, \theta)$ with positive α have non-vanishing variances (non-self averaging in the physics terminology), while in $PD(\theta)$ they do not.

Invariance under Size-biased Permutation

We introduce the notion of invariance under size-biazed sampling or permutation in the statistics literature as a proper concepts of distribution of sizes of types in statisstical equilibrium.

Heuristically this notion may arise in the following way: Suppose that fractions of "shares" are arranged in decreasing order, $x_1 > x_2 > \cdots$. We may be interested in the question of how large is the share of the second type, excluding the presence of the first, that is the largest type. This is the fraction $x_2/(1 - x_1)$. Analogously, we may be interested in the i-th largest type excluding or correcting for the effects of the first through the $(i-1)$th shares, given by $x_i/(1 - x_1 - \cdots x_{i-1})$. Actually, this is one of the ways industrial organizaation economists measured the concentration of industries, even though they did not know of the notion of the size-biased sampling or permutation. This is precisely what is involved in size-biased sampling.

More formally, we consider the set of all possible fractions $(p_1, p_2, \ldots)$ where p_i, the fraction of type i agents, is posistive, and the fractions sum to 1, $\sum_i p_i = 1$. Suppose that one agent is sampled. The probability that the first sampled agent is of type j is

$$\Pr(\hat{p}_1 = p_j | p_1, p_2, \ldots, p_n) = p_j, : j = 1, 2, \ldots n.$$

This first pick is called the size-biased pick, because types of agents with larger fraction are most likely to be sampled. This equation says that the sample is taken in proportion to the sizes of various types. More generally, having picked $\hat{p}_1, \ldots, \hat{p}_k$, the next sampled agent is of type n with probability given by

$$\Pr(\hat{p}_{k+1} = p_n | \hat{p}_i, i = 1, 2, \ldots, k; p_1, p_2, \ldots) = \frac{p_n}{1 - \hat{p}_1 - \hat{p}_2 - \cdots - \hat{p}_k},$$

provided that $p_n \neq \hat{p}_i, i = 1, 2, \ldots, k$. The collection, $\{\hat{p}_j\}$, is called size-biased sampling or permutation abbreviated as SBP.

Since distributions of agents by types are more useful when they are in statistical equilibrium, we define the set of fractions is invariant under size biased permutation (abbreviated as ISBP) when

$$\{\hat{p}_n\} =^d \{p_n\},$$

where $=^d$ means equality in distribution.

Pitman (1996) considered $\{p_n\}, p_n > 0, a.s.$, for all n, $\sum_n p_n = 1$, such that $\{p_n\}$ are distributed as RAM (residual allocation model) for independent random variables W_i, $i = 1, 2, \ldots$, that is ps are generated by the following formula

$$p_1 = W_1, p_2 = W_2(1 - W_1), \cdots, p_n = W_n(1 - W_1)(1 - W_2)\cdots(1 - W_{n-1}).$$

Note that $p_1 = W_1, p_2/(1-p_1) = W_2, \cdots, p_n/(1 - p_1 - \cdots - p_{n-1}) = W_n$ are independent.

Let α and θ be such that $0 \le \alpha < 1$, and $\theta > -\alpha > 0$. Let W_i be Beta distriuted random variable, $Be(1-\alpha, \theta + i\alpha)$, where random variable X has density $Be(a, b)$ when the density is given by

$$f_X(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1},$$

for $0 < x < 1$, where $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$.

Then $\{p_n\}$ is said to have a $GEM(\alpha, \theta)$ distribution.[3] With $\alpha = 0$, the above reduces to the one-parameter Poisson-Dirichlet distribution, due to Kingman (1978). See Perman, Pitman, and Yor (1992), and Pitman and Yor (1997) on earlier works.

Then he showed that $\{p_n\}$ is invariant under size-biased permutation, abbreviated as ISBP, if and only if $\{p_n\}$ is distributed as $GEM(\alpha, \theta)$.

Next, arrange samples by order statistics, i.e., we reorder $\hat{p}_i$, $i = 1, 2, \ldots$ as

$$p_{(1)} > p_{(2)} > \cdots.$$

When $\{p_n\}$ is distributed as $GEM(\alpha, \theta)$, then the ranked sequence $\{p_{(n)}\}$ is said to have the two-parameter Poisson Dirichlet distribution, $PD(\alpha, \theta)$.

To summarize, if fractions of agents of type n is given by $\{p_n\}$, $p_n > 0$, a.s., and $\sum_n p_n = 1$, the size-biased permutation of $PD(\alpha, \theta)$ is a $GEM(\alpha, \theta)$, and the ranked sequence of a $GEM(\alpha, \theta)$ is a $PD(\alpha, \theta)$. Furthermore, $GEM(\alpha, \theta)$ is ISBP. See Carlton (1999), for example.

With $\alpha = 0$, $PD(\alpha, \theta)$ reduces to the Ewens distribution, denoted from now on by $PD(\theta)$.

Structural Distribution and Frequency Spectrum

The structural distribution, F, of $\{p_n\}$, is defined by Engen to be the distribution on $(0, 1]$ of the first size-biased pick, that is the first term of a size-biased permutation of the distribution of agents by type, $\{p_n\}$, that is $\hat{p}_1$. The importance of this first pick is demonstrated by the lemma below of Pitman and Yor (1997).

When $\{p_n\}$ is distriubted as a two parameter Poisson-Dirichlet distribution $PD(\alpha, \theta)$, let W_1 be distributed as $Be(1-\alpha, \theta+\alpha)$ (Beta distribution). We drop subscript 1 from W_1 from now on. The first size-biased pick is

[3] The name GEM was given by Ewens to honor the pioneers, Griffiths, Engen, and McCloskey.

$\hat{p_1} = W$ as we have shown above. The structural distribution is important because it shows that $\hat{p}_1$ summarizes the distribution of $\{p_n\}$ as shown next.

Lemma: For any positive measureable function $g(t) \sim O(t)$ as t goes to zero,

$$E[g(W)/W] = E[g(\hat{p_1}/\hat{p_1}]$$

$$= E\{E\sum_i \frac{g(p_n)}{p_n} Pr(\hat{p_1} = p_n | p_1, p_2, \ldots)\}$$

$$= E(\sum \frac{g(p_n)}{p_n} p_n) = E[\sum g(p_n)].$$

Pitman (1996) pointed out that $v^{-1}F(dv)$ is the frequency spectrum. By the above lemma, the expected value of any positive measurable function g is expressible in terms of the structural distribution as

$$E(\sum_n g(p_n)) = \int_0^1 \frac{g(v)}{v} F(dv).$$

If one takes g to be $I(a < v < b)$, this expression gives the average number of n such that $a < p_n < b$, hence $v^{-1}F(dv)$ is the same as the frequency spectrum in population genetics literature. In that literature, there is a measure of cluster size distribution called frequency spectrum. See Ewens (1979). Aoki (2002, p.173, 2002a) has some elementary economic applications of this notion. In words, the frequency spectrum is the expected number of types with fraction in the interval $(x, x + dx)$.

Given order statistics of cluster sizes governed by $PD(\theta)$, $x_1 > x_2 > \cdots$, the largest size x_1 has the density

$$f(x_1) = \theta x_1^{-1}(1 - x_1)^{\theta-1},$$

for x_1 in the range $1/2 < x_1 < 1$, that is when the largest cluster is more than $1/2$ of the whole.[4] This density behaves like x_1^{-1} for small x_1. This indicates that there are many types with small fractions and $f(x)$ is not normalizable. However, $g(x) = xf(x) = \theta(1 - x)^{\theta-1}$ is normalizable. This function is interpreted as the probability that a randomly selected sample is of the type with fraction in $(x, x + dx)$.

The two largest fractions, x_1 and x_2 have the joint density

$$f(x_1, x_2) = \theta^2 (x_1 x_2)^{-1}(1 - x_1 - x_2)^{\theta-1},$$

when the two sizes are such that $0 < x_1 + x_2 < 1$, and more importantly when

$$\frac{x_2}{1 - x_1} > \frac{1}{2}.$$

Note that similar inequalities arise in size-biased permutation. See Aoki (2002, Sec. 10.6) for heuristic derivations based on Watterson and Guess (1977). [5]

[4] The expression is more complicated when x_1 is less than 1/2. See Watterson and Guess (1977).

[5] Karlin (1967) focussed on the situation with many types of small probabilities such that $\beta(x) = x^{-\gamma}L(x)$, with $0 < \gamma < 1$, and where $\beta(x) = \sum_i^\infty I(p_n \geq x)$, and where $L(.)$ is some slowly varying function.

In economic applications we are more interested in a few types with large shares, such as the ones discussed in Aoki (2000a).

For the one-parameter Poisson-Dirichlet process, the expected sizes of the three largest clusters are shown in the next table (see Griffiths (2005))

θ	largest	second	third
0.1	0.935	.059	.005
0.5	.758	.171	.049
1.0	.624	.210	.088

For example, with $\theta = 0.1$, the expected size of the largest and the second largest clusters sum to 99 per cent of the whole agents. With $\theta = 1/2$, the sum is about 93 per cent.

Number of Clusters in two-parameter Poisson-Dirichlet Distributions

The probabilities of new types entering models in $PD(\theta)$, and the number of clusters have been applied for example in Aoki (2002, p.176, App. A.5). In the two-parameter Poisson-Dirichlet distribution the conditional probabilities for the number of clusters in a sample of size n, K_n is given by

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

and

$$\Pr(K_{n+1} = k | K_1, \ldots, K_n = k) = \frac{n - k\alpha}{n + \theta}. \tag{2}$$

In other words, the random variable K_n is the number of different types of agents present in a sample of size n. Eq.(1) means that the $(n+1)$th entrant is a new type. Eq.(2) means that it is one of the previously existing types. Hence the number of clusters does not change.

From (1) and (2) the probability for $K_n = k$, $q(n,k)$, can be recursively computed using the conditional probability equation above

$$q(n+1,k) = \frac{(n-k\alpha)}{(n+\theta)} q(n,k) + \frac{\theta + (k-1)\alpha}{n+\theta} q(n,k-1), \tag{3}$$

for $1 \le k \le n$, given the boundary formula

$$q(n,1) = \frac{(1-\alpha)(2-\alpha)\cdots(n-\alpha)}{(\theta+1)(\theta+2)\cdots(\theta+n)},$$

and

$$q(n,n) = \frac{(\theta+\alpha)(\theta+2\alpha)\cdots(\theta+n\alpha)}{(\theta+1)(\theta+1+\alpha)\cdots(\theta+1+\alpha(n-1))}.$$

These expressions generalize the recurrence relation for the one-parameter $PD(\theta)$. In the one-parameter case, $\theta/(\theta+n)$ is a probability that the (n+1)th agent that enter the model is a new type, and $n/(\theta+n)$ is the probability that the next agent is one of the types already in the model.

In the one-parameter case, $q_{n,k} := P(K_n = k)$ is governed by the recurrence relation

$$q_{n+1,k} = \frac{n}{n+\theta} q_{n,k} + \frac{\theta}{\theta+n} q_{n,k-1}.$$

The solution of this recurrence equation is expressible as

$$q_{n,k} = \frac{c(n,k)\theta^k}{\theta^{[n]}},$$

where $\theta^{[n]} := \theta(\theta+1)\cdots(\theta+n-1) = \frac{\Gamma(\theta+n)}{\Gamma(\theta)}$, and $c(n,i)$ is the unsigned (signless) Stirling number of the first kind. It satisfies the recursion

$$c(n+1,k) = nc(n,k) + c(n,k-1).$$

Since $q_{n,k}$ sums to one with respect to k we have

$$\theta^{[n]} = \sum_{k=1}^{n} c(n,k)\theta^k. \tag{4}$$

See Aoki (2002, p.208) for example on the Stirling numbers, and their combinatorial interpretations. In the two-parameter $PD(\alpha,\theta)$ case, the probability of the number of clusters is given by

$$P_{\alpha,\theta}(K_n = k) = \frac{\theta^{[k,\alpha]}}{\alpha^k \theta^{[n]}} c(n,k;\alpha), \tag{5}$$

where

$$\theta^{[k,\alpha]} := \theta(\theta+\alpha)(\theta+2\alpha)\cdots(\theta+(k-1)\alpha),$$

and the expression $c(n,k;\alpha)$ generalizes the signless Stirling number of the first kind of one-parameter situation.

Let $S_\alpha(n,k) := \frac{1}{\alpha^k} c(n,k;\alpha)$. It satisfies the recursion

$$S_\alpha(n+1,k) = (n-k\alpha)S_\alpha(n,k) + S_\alpha(n,k-1).$$

This is called generalized Stirling number of the first kind. See Charalambides (2002). Instead of (4) we have

$$\theta^{[n]} = \sum_{k=1}^{n} S_\alpha(n,k)\theta^{[k,\theta]}. \tag{6}$$

Pitman (1999) obtained its asymptotic expression as

$$S_\alpha \sim \frac{\Gamma(n)}{\Gamma(k)} n^{-\alpha} \alpha^{1-k} g_\alpha(x),$$

where $k \sim xn^\alpha$. Here, g_α is the Mittag-Leffler function. This function is discussed in the next section.

Asymptotic Behavior of Cluster Sizes

We collect here some known facts about cluster sizes as $n \to \infty$.

The number of clusters K_n

Yamato and Sibuya (2000) obained

$$EK_n = \frac{\theta}{\alpha}[\frac{(\theta+\alpha)^{[n]}}{\theta^{[n]}} - 1],$$

where we note that

$$\frac{(\theta+\alpha)^{[n]}}{\theta^{[n]}} = \frac{\Gamma(\theta)}{\Gamma(\theta+\alpha)}\frac{\Gamma(\theta+\alpha+n)}{\Gamma(\theta+n)}.$$

Applying the asymptotic expression for the Gamma function for large n

$$\frac{\Gamma(n+a)}{\Gamma(n)} \sim n^a,$$

to the above expression, we have an asymptotic expression,

$$E(\frac{K_n}{n^\alpha}) \sim \frac{\Gamma(\theta+1)}{\alpha\Gamma(\theta+\alpha)}. \tag{7}$$

They also calculate the asymptotic value of the variance of K_n/n^α,

$$var(K_n/n^\alpha) \sim \frac{\Gamma(\theta+1)}{\alpha^2}\gamma_{\alpha,\theta} \geq 0, \tag{8}$$

where

$$\gamma_{\alpha,\theta} := \frac{\theta+\alpha}{\Gamma(\theta+2\alpha)} - \frac{\Gamma(\theta+1)}{[\Gamma(\theta+\alpha)]^2}. \tag{9}$$

Note that

Fact: $\gamma_{0,\theta} = 0$.

This fact is important in the long-run behavior of components of the partition vectors, to be discussed in the next subsection.

Actually they calculate more generally

$$limE(\frac{K_n}{n^\alpha})^r = \mu'_r,$$

where μ'_r is the $r-th$ moment of the generalized Mittag-Leffler distribution with density

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

where $\theta/\alpha > -1$, and where $g_\alpha(x)$ is the Mittag-Leffler (α) density function. It is known that this function is uniquely determined by the moment conditions

$$\int_0^\infty x^p g_\alpha(x)dx = \frac{\Gamma(p+1)}{\Gamma(p\alpha+1)},$$

for all $p > -1$. Note that the integral of $g_{\alpha,\theta}$ over the interval from zero to infinity is 1, as it should be.

See also Blumenfeld and Mandelbrot (1997) who credit Feller (1949) as the original source.

Mittag-Leffler distributions

As we discuss more fully later, Pitman (2002, Sec. 3) has stronger result:

$$K_n/n^\alpha \to \mathcal{L}, a.s.,$$

where the expression $\mathcal{L}$ has the density

$$\frac{d}{ds}P_{\alpha,\theta}(\mathcal{L} \in ds) = g_{\alpha,\theta}$$

where letting $\eta = \frac{\theta}{\alpha}$ we define

$$g_{\alpha,\theta}(s) := \frac{\Gamma(\theta+1)}{\Gamma(\eta+1)}s^\gamma g_\alpha(s),$$

where $s > 0$, and where $g_\alpha = g_{\alpha,0}$ is the Mittag-Leffler density

$$g_\alpha(s) = \frac{1}{\pi}\sum_{k=1}^{\infty}[\frac{\Gamma(k\alpha)}{\Gamma(k)}sin(k\pi\alpha)(-s)^{k-1}].$$

We note that

$$\mu_1' = E_{\alpha,\theta}(\mathcal{L}) = \Gamma(\theta+1)/\alpha\Gamma(\theta+\alpha),$$

and

$$\mu_2' = E_{\alpha,\theta}(\mathcal{L}^2) = \Gamma(\theta+1)(\theta+\alpha)/\alpha^2\Gamma(\theta+2\alpha).$$

Hence variance of $\mathcal{L}$ is given as $\mu_2' - (\mu_1')^2$.

For the record we have

$$var_{\alpha,\theta}\{\frac{K_n}{n^\alpha}\} = var_{\alpha,\theta}\mathcal{L}.$$

The partition vector a

Denote the partition vector by $\mathbf{a} = (a_1, a_2, \ldots)$, where we recall that a_i is the number of distinct clusters of size i, hence $\sum_i a_i = K_n$, and $\sum_i ia_i = n$. Yamato and Sibuya obtain the limit of the first component, a_1

$$limE[\frac{a_1}{n^\alpha}] = \frac{\Gamma(\theta+1)}{\Gamma(\theta+\alpha)},$$

and

$$lim\, var(\frac{a_1}{n^\alpha}) = \Gamma(\theta+1)\gamma_{\alpha,\theta} \geq 0.$$

In fact a_j/n^α are all non-self averaging, as well as ja_j/n^α, where ja_j is the total number of agents in the clusters of size j. Note that their variances are all zero with $\alpha = 0$, that is the asymptotic variance of a_j/n^α are all zero in $PD(\theta)$ models.

Fact

The expression a_i/n^α, $i \geq 1$ are all non-self averaging.

Sibuya (2005) used Formula 6.1.41 in Abramovitz and Stegun (1965) to obtain asymptotic expression

$$E(\frac{a_j}{n^\alpha}) \approx \frac{(1-\alpha)^{[j-1]}}{j!}\frac{\Gamma(\theta+1}{\Gamma(\theta+\alpha)} + O(n^{-1}).$$

We state this as

Proposition: As in (9)

$$lim\, var_{\alpha,\theta}(K_n/n^\alpha) = var_{\alpha,\theta}(\mathcal{L}),$$

and

$$lim\, var_{\alpha,\theta}(a_j/n^\alpha) = \alpha^2 var_{\alpha,\theta}(\mathcal{L}).$$

They also show that covariances of components of the partition vectors are non-self averaging with positive α values:

$$limCov(\frac{a_i}{n^\alpha}, \frac{a_j}{n^\alpha}) = \Gamma(\theta+1)\gamma_{\alpha,\theta} \times \frac{(1-\alpha)^{[i-1]}}{i!}\frac{(1-\alpha)^{[j-1]}}{j!} > 0, : \alpha > 0.$$

It is also known that

$$\frac{j!\Gamma(1-\alpha)}{\alpha\Gamma(j-\alpha)} \to^d \mathcal{L}. \tag{10}$$

We have

$$E(\frac{a_j}{n^\alpha}|k_n = k) \sim \frac{(1-\alpha)^{[j-1]}}{j!}(1-j/n)^{-(1+\alpha)} \times \xi,$$

where ξ depends on $g(_{\alpha,\theta})$.

The number of clusters, K_n is spread among the components of the partition vector, $a_i, i = 1, 2, \ldots, n$ at the proportion $\alpha(1-\alpha)^{[j-1]}/j!$, $0 < \alpha < 1$. Devroye (1993) calls this Sibuya distribution.

We also note that

$$Lim\frac{E(a_i)}{E(K_n} = \frac{\alpha^2}{\Gamma(\theta+\alpha)\gamma_{\alpha,\theta}}.$$

We note that a_j/K_n is self-averaging for all $j = 1, \ldots, n$. Yamato and Sibuya also examined the clusters of size k or less

$$K[1,k] := a_1 + a_2 + \cdots + a_k,$$

and the number of agents in $K[1,k]$, denoted by $N[1,k]$ and obtained their limiting expressions as

$$\frac{K[1,k]}{n^\alpha} \to^d \{1 - \frac{(1-\alpha)^{[k]}}{k!}\}\mathcal{L},$$

and

$$\frac{N[1,k]}{n^\alpha} \to^d \alpha\frac{(2-\alpha)^{[k-1]}}{(k-1)!}\mathcal{L},$$

Sibuya also notes that

$$\{\frac{a_1}{n^\alpha}, \frac{2a_2}{n^\alpha} \cdots \frac{.ka_k}{n^\alpha}\}$$

converges in distribution to a sequence of random variables depending on $\mathcal{L}$ as

$$\{1, \frac{(1-\alpha)}{1!}, \dots, \frac{(1-\alpha)^{[k-1]}}{(k-1)!}\}.$$

In $PD(\alpha)$ it is known that

$$\frac{K_n - \theta lnn}{\sqrt{\theta lnn}} \to N(0,1).$$

Hence (K_n/lnn) is self-averaging.

Almost sure convergence

Denote by $a_j(n)$ the number of clusters of size j when there are n agents in the model. We noted earlier that $\sum_{j=1}^n ja_j(n) = n$, and $K_n := \sum_j^n a_j(n)$ is the total number of clusters formed by the total of n agents.

By Rouault (1976, 1978)

$$\frac{a_j(n)}{K_n} \to \frac{\alpha\Gamma(j-\alpha)}{\Gamma(1-\alpha)j!}, a.s.$$

Recallint that $K_n/n^\alpha \to \mathcal{L}$,a.s., we have

$$a_j(n)/n^\alpha \to \frac{\alpha\Gamma(j-\alpha)}{\Gamma(1-\alpha)j!}\mathcal{L}, a.s.$$

wbere

$$\frac{a_j(n)}{K_n} \to \frac{\alpha}{j!}P_{\alpha,j},$$

where

$$P_{\alpha j} = \frac{\Gamma(j-\alpha)}{\Gamma(1-\alpha)},$$

for every $j = 1, 2, \dots$ a.s. as n goes to infinity, and that $a_j(n) \sim P_{\alpha,j}\mathcal{L}n^\alpha$ in a two-parameter Poisson-Dirichlet case.

Local Limit Theorem

Suppose N independent positive random variables X_i, $i = 1, 2, \dots N$ are normalized by their sum $S_N = X_1 + \cdots + X_N$

$$x_i = X_i/S_N, i = 1, \dots N,$$

so that

$$Y_1 := \sum_i x_i = 1.$$

Suppose that the probability density of X_i is such that it has a power-law tail,

$$\rho(x) \sim Ax^{-1-\mu},$$

with $0 < \mu < 1$. Then, $S_N/N^{1/\mu}$ has a stable distribution (called Lévy distribution).

Pitman's formula for the probability of $K_n = k$, with $k \sim sn^\alpha$ indicates that the power law n^α which is $2\alpha < 2$ or $2\alpha = 1 + \mu$ with $0 < \mu < 1$, the case in Derrida.

With the 2-parameter PD distribution satisfying the power law condition, Derrida's conclusion that the Hs are non-self averaging applies to this case as well.

Estimating the Parameters

Carlton (1999) and Sibuya (2005)are the only systematic source on estimating the parameters of two-parameter Poisson-Dirichlet distributions.

With $\alpha = 0$, Ewens had shown that K_n is the sufficient statistics for θ. Carlton discusses the case where α is known and θ unknown. He derives the asymptotic distribution of the maximum likelihood estimate of θ, given n samples.

Lemma

Given α in (0,1), the maximum-likelihood estimate of θ, $\hat{\theta}_n$ is given by

$$\psi(1 + \hat{\theta}_n/\alpha) - \alpha\psi(1 + \hat{\theta}_n) \to logS, as.$$

Here ψ is the digamma function.

With θ known, and α unknown, Carlton proves

Lemma

Let $\{A_1, \ldots, A_n\}$ is distributed according to the two-parameter Ewens distribution of size n. (His Eq. (4.2) on page 55.) Then,

$$\hat{\alpha}_n = \frac{logK_n}{logn} \to \alpha \ a.s.$$

Sibuya uses the conditional probability distribution of the partition vector components, given that $\sum_i a_i = k$, and expresses the distribution

$$P(\mathbf{a}|\sum a_j = k) = \frac{1}{S_\alpha(n,k)} \frac{n!}{\prod a_j!} \prod_j \{\frac{(1-\alpha)^{[j-1]}}{j!}\}^{a_j}$$

which is proportional to

$$exp\{-\sum \frac{j}{2(j-2)!} a_j\}\alpha + O(\alpha^2)$$

and test the hypothesis $\alpha = 0$, against the alternative hypothesis $\alpha < 0$.

He proposes the rejection region

$$\sum \frac{j}{2(j-2)!} - a_j > const.k.$$

When both parameters are unknown, the estimation problem is apparently unsolved.

Some Potential Applications

In physics literature, Derrida (1994 a, b)sketched a derivation that the expected values of $Y_k = \sum_i x_i^k, k = 2, 3, \ldots$ can be calculated for mean field spin glass models using the Parisi replica approach, and remarkably the formula is the same as the GEM model described above.

In the rest of this section we focus on economic examples.

Example 1 Scaling of GDP growth rates was considered by Canning, Amaral, Lee, Meyer, and Stanley (1998). They showed that the standard deviation of the GDP growth rate may sclae as $Y^{-\beta}$, with β about 0.15. Here, we heuristically explain how their finding may be explained using a random partition framework.

We modify the model of Huang and Solomon (2001) and apply the same procedures to estimate the growth rate of real GDP.[6] View the real economy as composed of K sectors of various sizes. Stochastically one or more of the sectors experience what we call elementary events, the aggregate of which yields the real growth of the economy, leading to its random growth rates. To be simple one may assume that the individual elementary growth of sectors is random $\lambda = 1+g$, where $g = \pm\gamma$ randomly with some positive γ. Further, we adopt the mechanism of Huang and Solomon that a random number τ of this type of elementary events are experienced in a unit of calendar time. The random growth rate is the composite effects of these random elementary events.

We refer the detail of the mechanism to their paper, and mention only that the growth rate will be exponential only if the number of changes τ is less than some critical value τ_c, and change in GDP has a power law density with index $-(1+\alpha)$.

The value of α is defined to be the ratio of minimum and average real consumption in the model $q = c_{min}/c_{average}$, and is tied to α by

$$\alpha \approx 1/(1-q),$$

when K is sufficiently larger that $e^{1/q}$, due to inherent normalization conditions of densities involved.

For example, setting $q = 0.25$ leads to $\alpha = 1.33$, and K must be such that $K >> e^4 > 55$. The value of τ_c is defined by $(N/2q)^\alpha$. With τ less than τ_c, the growth rate r can be shown to have the density

$$p(r) = Cexp(-a|r - r_m|),$$

for $r > r_m$, with a different constant for the case $r < r_m$.

The deviation of r is then related to variability of K and τ, among others. From this one can deduce that the average deviation in the growth rates is basically determined by percentage changes of the size of the largest cluster which can be related to the GDP when the productivity is assumed not to vary too much, and the conclusion follows that the standard deviation of

[6] Their focus is on financial sector, not real sector. See Aoki and Yoshikawa (2006 a, b).

the growth rate is $Y^{-\mu}$ with μ less than 1. See Aoki and Yoshikawa (2006a, b) for detail.

Example 2: Long-run effects of innovation and imitation

This example is based on a two-sector model discussed in Aoki (2002, Sec. 7.4), Aoki, Nakano, Yoshida (2004), and Aoki, Nakano, and Ono (2006). There are two types of firms, innovators and imitators.

Our model has two sectors; one technically advanced sector and the other less so. By a suitable choice of units we denote the sizes of the two sectors by a vector (n_1, n_2). We may think of them as the number of firms in some suitably chosen standard units. Firms in sector one succeed in creating innovative firms at rate f which is, for simplicity, exogenously fixed in this model.[7]

Firms' stochastic behavior is described by a continuous time Markov chain which is uniquely determined by a set of transition rates. We write the transition rate from state a to b by $w(a, b)$. This means that the probability that the system moves from state a to b in some small time interval is given by the time interval times the transition rates up to o(time interval size). They are specified as follows: The first two describe entry (growth) rates

$$w\{(n_1, n_2), (n_1 + 1, n_2)\} = c_1 n_1 + f,$$

$$w\{(n_1, n_2), (n_1, n_2 + 1)\} = c_2 n_2.$$

Here c_i is the rate of growth of type i firm size, $i = 1, 2$.

The next two specify exit rates from the model

$$w\{(n_1, n_2), (n_1 - 1, n_2)\} = d_1 n_1,$$

$$w\{(n_1, n_2), (n_1, n_2 - 1)\} = d_2 n_2.$$

Here d_i is the exit (death) rate of type i firms from the economy, $i = 1, 2$.

The last set of two transition rates describes how firms change their types

$$w\{(n_1, n_2), (n_1 + 1, n_2 - 1)\} = \mu g_1 n_2 (n_1 + h),$$

with $g_2 = c_2/d_2$, and $h = f/c_1$, and

$$w\{(n_1, n_2), (n_1 - 1, n_2 + 1)\} = \mu g_2 n_1 n_2,$$

with $g_i = c_i/d_i$, $i = 1, 2$, and $\mu = \lambda d_1 d_2$. This parameter λ is the coefficient in the transition rates of type changes by firms in the two sectors. The first of the two shows the rate at which one of type 1 firm becomes technologically obsolete and join the cluster made up of type 2 firms. The second equation specifies how firms of type 2 successfully imitate firms of type 1 and join their cluster. for example.

With these transition rates, we write the master equation. We compute the probability generating function, and then convert it into the cumulant generating function, since we are interested in calculating only the first and second order moments, $k_1, k_2, k_{1,1}, k_{1,2}$, and $k_{2,2}$, and verify that the $2\times$ 2

[7] It will be interesting to endogenize this rate in a way that is not equivalent to increasing the birth rate c_1 in the model of this section.

covariance matrix is positive definite in steady state. Fortunately, this model is specified in such a way that the equations for the moment are closed at the second moments, that is no higher order moments appear in the equations for the first and second moments. We derive a coupled ordinary differential equations for these moments. With the help of Mathematica we calculate the stationary state values of these moments for varioous parameter values, and verify the positive definiteness of the second moment matrix.

To the knowledge of the author this is the first example of Schumpeterian dynamics with innovations and immitation effects for which the first two moments have been analytically derived and numerically evaluated. The model allows us to examine parametrically the relative importance of net death rate and innovation rate, and draw conclusions about qualitative behavior of interacting two sectors. The model shows that both sectors co-exist in the long-run. We show also that the means of stationary locally stable equilibria scale with parameters of the innovation rate, and death rate.

The stochastic dynamic equation is easy to state. It is a backward Chapman-Kolmogorov equation, also known as the master equation. (We use the latter name as it is short, and implies that everything you need to know about stochastic behavior is implicit in the master equation.)

$$\frac{\partial P(n_1, n_2; t)}{\partial t} = I(n_1, n_2; t) - O(n_1, n_2; t), \tag{11}$$

where the first term collects all inflows of probability flux into state (n_1, n_2), and the second term collects all outflows of probability fluxes out of this state. There are six distinct flows. In detail we have

$$I(n_1, n_2; t) = P(n_1 + 1, n_2; t)d_1(n_1 + 1) + P(n_1, n_2 + 1; t)d_2(n_2 + 1)$$

$$+P(n_1 - 1, n_2; t)c_1(n_1 - 1 + h) + P(n_1, n_2 - 1)c_2(n_2 - 1)$$

$$+P(n_1+1, n_2-1; t)\mu g_2(n_1+1)(n_2-1)+P(n_1-1, n_2+1; t)\mu g_1(n_1-1+h)(n_2+1).$$

The second term in (1) is given by

$$O(n_1, n_2; t) = P(n_1, n_2; t)\{c_1+n_1+f+c_2n_2+d_1n_1+d_2n_2+\mu g_1n_2(n_1+h)+\mu g_2n_1n_2)\}.$$

To solve the master equation, we first convert it into the probability generating function

$$G(z_1, z_2; t) = \sum_{n_1,n_2} P(n_1, n_2; t)z_1^{n_1} z_2^{n_2}.$$

We obtain a partial differential equation for $G(z_1, z_2; t)$. It is given in Appendix. This partial differential equation is rather intractable, and for that reason we convert it into the cumulant generating function and solve for the expected values of first and second moments.[8]

Cumulant generating functions are related to the probability generating functions by

$$K(\theta_1, \theta_2; t) = \ln G(e^{-\theta_1}, e^{-\theta_2}),$$

[8] In some cases the resulting ordinary differential equations for the moments turn out to be an infinite set of coupled ordinary differential equation. Fortunately, the differential equations for the first and second cumulants are self-contained in this model.

where we change variables from z_1, and z_2 into θ_1 and θ_2.

It is known that the cumulant generating function has a Taylor series expansion of the form

$$K(\theta_1, \theta_2; t) = k_1\theta_1 + k_2\theta_2 + \frac{1}{2}(\theta_1, \theta_2)\Theta(\theta_1, \theta_2)' + \cdots,$$

where $k_1 = E(n_1)$, and $k_2 = E(n_2)$, that is, they are the expected sizes of the two types, and where Θ is a covariance matrix made up of the variances and covariances of the two sizes,

$$\Theta = \begin{pmatrix} k_{1,1} & k_{1,2} \\ k_{1,2} & k_{2,2} \end{pmatrix}.$$

See Aoki (2002, Chapt. 7) for further information on these generating functions, and some simple examples.

From the cumulant generating functions we derive a set of five ordinary differential equations for $k_1, k_2, k_{1,1}, k_{1,2}$, and $k_{2,2}$.

Appendix gives the explicit expressions.

Stationary means and variances

The equations for the two means are:

$$\frac{dk_1}{dt} = f - d_1(1 - g_1)k_1 + \lambda f d_2 k_2 + 2\mu\lambda A_0, \tag{12}$$

and

$$\frac{dk_2}{dt} = -d_2(1 - g_2 + \lambda f)k_2 - 2\mu\lambda A_0, \tag{13}$$

where $\lambda = \mu/d_1 d_2$, where $g_i = c_i/d_i$, $A_0 = k_{1,2} + k_1 k_2$, and $2\mu = d_1 d_2(g_1 - g_2)$. Note that $A_0 =< n_1 n_2 >\geq 0$.

Since A_0 depends on $k_{1,2}$ we need solve for it as well.

Stationary means are described by setting the left-hand sides of (2) and (3) to zero:

$$f - d_1(1 - g_1)k_1 + \lambda f d_2 k_2 + 2\mu\lambda A_0 = 0, \tag{14}$$

$$-d_2(1 - g_2 + \lambda f)k_2 - 2\mu\lambda A_0 = 0. \tag{15}$$

By adding (14) and (15) to express an important relation between f, k_1 and k_2

$$f = d_1(1 - g_1)k_1 + d_2(1 - g_2)k_2. \tag{16}$$

This equation clearly shows that a fraction of innovation flow accounts for the new firms in sector 1, and the rest accounts for the net exit flow of firms from sector 2. We later show that the expected value of the stationary values of the size of sector 1 scales with $\gamma_1 := \lambda d_1$, and that of sector 2 scales with $\gamma_2 := \lambda d_2$,

Recalling the definition $d_i(1 - g_i) = d_i - c_i$, we see that the rate of innovation f equals the sum of the expected exit rates of firms of both sectors, since $(d_i - c_i)k_i$ is the net exit rate of firms of sector i, $i = 1, 2$.

In this model it is necessary that both g_1 and g_2 are not greater than one. There are two cases: $g_1 > g_2$ and $g_2 > g_1$. In the former $\mu > 0$ and $\mu < 0$ in the latter. Eqs. (14) and (16) exclude the case $g_1 > g_2$. In the other cases, we could have $g_2 > 1 > g_1$ or $1 > g_2 > g_1$.

To reflect these consideration we introduce two parameters m and n[9] by

$$g_1 = 1 - n\gamma,$$

and

$$g_2 = 1 - m\gamma,$$

where $\gamma = \lambda f$. The constants are bounded by

$$\frac{1}{\gamma} > n > 0,$$

and

$$n > m > -1.$$

The inequality $m > -1$ is derived from (15).

Solving (4) and (5), we obtain the means of the sizes of the two sectors

$$k_1 = \frac{1 - v}{n\gamma_1,} \tag{17}$$

with

$$v := \frac{m(n-m)}{1+m} d_1 d_2 \lambda^2 A_0,$$

and

$$k_2 = \frac{v}{m\gamma_2}. \tag{18}$$

The stationary variances $k_{1,1}$ and $k_{2,2}$ are derived in Appendix; the results are

$$\begin{aligned} k_{1,1} \quad &= \quad \frac{1}{2\gamma_1 \left(n + \frac{n-m}{m} v\right)} \left\{ \frac{2}{n\gamma} + \frac{2}{\gamma\gamma_1 m n^2 (n-m)} \Big[\gamma_1 n (m^2 + n) \right. \\ &\left. -\gamma m \Big(\gamma_1 n^2 (1+m) - m(1+n) \Big) \Big] v + \frac{2(2+n)}{\gamma_1 n^2} v^2 + \frac{2(n-m)}{\gamma_1 m n^2} v^3 \right\}, \end{aligned}$$

and

$$\begin{aligned} k_{2,2} \quad &= \quad \frac{1}{2\gamma_2 \left(1 + m - \frac{n-m}{n}(1-v)\right)} \\ &\times \quad \left\{ \frac{2(1-m\gamma)(1+n)}{m(n-m)\gamma} v + \frac{2(1+n)}{mn\gamma_2} v^2 + \frac{2(n-m)}{m^2 n \gamma_2} v^3 \right\}. \end{aligned}$$

The covariance $k_{1,2}$ is expressed in terms of θ through the definition $k_{1,2} = A_0 - k_1 k_2$ as

$$k_{1,2} = \frac{1}{\gamma_1 \gamma_2} \frac{1}{nm(n-m)} v \left[(1+n)m + (n-m)v \right].$$

[9] not to be confused with n_1 or n_2

What remains is to determine v. The self-consistent equation for v is derived in Appendix. Although the equation is a fifth order equation of v because of the five unknown quantities, the highest term vanishes so that

$$vF(v) = 0 \tag{19}$$

where

$$F(\theta) = r_0 + r_1 v + r_2 v^2 + r_3 v^3. \tag{20}$$

The forms of r_i are given in Appendix.

The root $\theta = 0$ is of interest because this value of θ yields a stationary state in which sector 2 vanishes, $k_2 = 0$, $k_{2,2} = 0$, and $k_{12} = 0$.

Hence we have the solutions for (19); solutions of $F(v) = 0$ in addition to $v = 0$. For $v = 0$ one has

$$k_1 = \frac{1}{n\gamma_1}, \quad k_{11} = \frac{1}{n^2\gamma\gamma_1}, \quad k_2 = k_{22} = k_{12} = 0,$$

which corresponds to the situation that only sector 1 survives. From $F(v) = 0$ we obtain three values of v The roots must be such that v is real and the obtained values of $k_1, k_2, k_{1,1}$ and $k_{2,2}$ are positive; $k_{1,2}$ is not necessarily positive. Although the analytic solutions may be obtained for special set of parameters, such solutions are not possible in general.

Mathematica, however, enables us to numerically solve $F(v) = 0$. In order that those solutions exist in reality, the solutions must be the stable fixed points.

As an example we describe in detail the case where $m = .01$; $n = 2$, $\gamma = \gamma_1$, and $\gamma_2 = \gamma_1 + \epsilon$, with a small positive ϵ. In this case there is only one root for which the dynamics are locally stable. It is given by $v = 0.472$.

The stability of the stationary states is examined in the following way. The starting equations are (12), (13) and (21), (22) and (23) in Appendix. By setting the left hand sides of those equations we have the stationary values, which are confirmed to numerically coincide with the solutions from $vF(v) = 0$. Then the linearized equations for deviations $\delta k_1, \delta k_2, \delta k_{11}, \delta k_{22}$ and δk_{12} from the stationary values are derived. The eigenvalues of those equations are numerically calculated with a help of Mathematica. If real parts of all five eigenvalues associated with a stationary point are negative, the stationary point is stable.

The value of $v = .472$ corresponds to a locally stable solution. This leads to

$$k_1 = \frac{.264}{\gamma_1}, \;\; k_2 = \frac{47.2}{\gamma_2}.$$

From (5) we obtain

$$A_0 = \frac{v}{\gamma_1\gamma_2}\frac{1+m}{m(n-m)} = \frac{23.956}{\gamma_1\gamma_2}.$$

From this we derive

$$k_{1,2} = A_0 - k_1 k_2 = \frac{11.495}{\gamma_1\gamma_2}.$$

From (9) through (11) we can obtain approximate order of magnitude values for the second moments $k_{1,1}$ and $k_{2,2}$ as follows.

$$k_{1,1} = \frac{C_{1,1}}{\gamma_1^2},$$

with

$$C_{1,1} \approx \frac{1}{n^2}(1+v^2) \approx .31,$$

which is close to .309 obtained in the numerical example below, and

$$k_{2,2} = \frac{C_{2,2}}{\gamma_2^2},$$

with

$$C_{2,2} \approx \frac{1}{m^2}[v^2 + \frac{m(1+n)}{n}v] = 2300.$$

We also have an approximate expression for k_{12} as $C_{12}/\gamma_1\gamma_2$ with

$$C_{12} = \frac{\theta^2}{mn} \approx 11.15,$$

which is in good agreement with the value obtained above as 11.495.

Numerical Examples

We focus on a stationary solutions. Since there are five parameters, we have many solutions.

To keep the sizes of the two sectors at reasonable values, we examine cases with the death rates close to the birth rates. Namely, we choose g_i to be close to unity. Previously we have indicated that g_2 can be either larger than one or smaller than one, while g_1 is always less than one. First, we consider the case that the death rate d_i is slightly larger than the birth rate c_i, so that $n\gamma$, $m\gamma \ll 1$. Although the death rate of sector 2 is considered to be larger than that of sector 1, we assume that both are almost the same. We focus on the following parameters; $\gamma = \gamma_1 = \gamma_2 = 0.01, n = 2.0, m = 0.01$. Then we have three types of solutions; (1) $k_1 = 50, k_{11} = 2500, k_2 = k_{22} = k_{12} = 0$, (2) $k_1 = 49.97, k_2 = 4.77, k_{11} = 2501, k_{22} = 46505, k_{12} = 3.71$ and (3) $k_1 = 26.4, k_2 = 4719, k_{11} = 3093, k_{22} = 2.37 \times 10^7, k_{12} = 114918$. The stable solution is only the first type; only sector 1 survives. The second and third types are not stable. If we increase γ_2 slightly to $\gamma_2 = 0.011$, a remarkable change occurs in the type 3 solution. The numbers for (1) are the same as the previous case. On the other hand, (2) $k_1 = 49.977, k_2 = 4.122, k_{1,1} = 2501, k_{2,2} = 40232, k_{1,2} = 3.2$ and (3) $k_1 = 23.94, k_2 = 4738, k_{1,1} = 3216.7, k_{2,2} = 2.378 \times 10^7, k_{1,2} = 127042$. The second solution is not stable, but the third solution turns out to be stable in this case.

We vary a value of γ_2 with other parameters fixed. We found that the stable fixed point exists in a narrow range such that $0.02 \gtrsim \gamma_2 \gtrsim 0.0102$.

What parameters are chosen to increase the number of companies? For that purpose we should decrease $\gamma, \gamma_1, \gamma_2$. When $n = 2.0$ and $m = 0.01$

are fixed, we employ $\gamma = \gamma_1 = 0.001, \gamma_2 = 0.0011$. Then we have the stable third solution $k_1 = 264, k_2 = 47193$ with the correlation coefficient $k_{1,2}/\sqrt{k_{1,1}k_{2,2}} = 0.42$. In other sets of parameters with n and m fixed at the above values, $\gamma = \gamma_1$ and γ_2 being slightly larger than γ, we have the following scaling relation

$$k_1 = \frac{0.264}{\gamma_1}, \quad k_2 = \frac{47.2}{\gamma_2}, \quad k_{1,1} = \frac{0.309}{\gamma^2}, \quad k_{2,2} = \frac{2374}{\gamma^2}, \quad k_{1,2} = \frac{11.5}{\gamma^2}.$$

The correlation coefficient is 0.42.

The coefficient of variations are 2.11 and 1.03 for the two sectors respectively. We also note that with γ_2 nearly the same as γ_1 only 0.6 percent of the total sizes of the capital resides in sector 1.

Next we examine negative values of m. Take $m = -0.01$ while keeping the values of the other parameters the same as before. The numerical calculation gives a negative value of k_2. Although we have not done an extensive study, a negative value of m, i.e., g_2 is larger than one, may not yield stable stationary situations.

Aoki, Nakano, and Ono (2006) has more extensive simulations and verify non-self averaging property for the stationary sizes of the two sectors.

Example3: Disequilibrium theory of long run profits. Iwai's model has more than two sectors with different productivity coefficients. His paper is too long and involved to give a thumb-nail sketch here. Instead we offer three quotes from his paper to explain what he does.

> ...while both the differential growth rates among different efficiency firms and the diffusion of better technologies through imitations push the state of technology towards uniformity, the punctuated appearance of technological innovations disrupts this equilibrating tendency.

> ... over a long passsage of time these conflicting microscopic forces will balance each other in a statistical sense and give rise to a long-run distribution of relative efficiencies across firms. This long-run distribution will in turn allow us to deduce an *upward-sloping* long-run supply curves...

> This paper has challenged this long-held tradition in economics. It has introduced a simple evolutionary model which is capable of analyzing the development of the industry's state of technology as a dynamic interplay among many a firm's growth, imitation and innovation activities. And it has demonstrated that what the industry will approach over a long passage of time is not a classical or neoclassical equilibrium of uniform technology but a statistical equilibrium of technological disequilibria which maintains a relative dispersion of efficiencies in a statistically balanced form. Positive profits willl never disappear from the economy nomatter how long it is run. 'Disequilibrium' theory of 'long-run profits' is by no means a condtradition in terms.

We see that our random partiton framework along the line of Aoki, Nakano, and Yoshida (2004) can be applied to at least three types of firms, and their tail distribution may satisfy power laws to substantiate Iwai's claim by using long-run in time rather than the thermodynamic limits.

Concluding Remarks

In physics non-self-averaging phenomena abound. In traditional microeconomic foundations of economics, one deals almost exclusively with well-posed optimization problems for the representative agents with well defined peaks and valleys of the cost functions. It is also taken for granted that as the number of agents goes to infinity, any unpleasant fluctuations vanish and well defined deterministic macroeconomic relations prevail. In other words, non-self-averaging phenomena are not in the mental pictures of average macro- or microeconomists.

However, we know that as we go to problems which require agents to solve some combinatorial optimization problems, this nice picture may disappear. In the limit of the number of agents going to infinity some results are sample-dependent and deterministic results will not follow. Some of this type of phenomena have been reported in Aoki (1996, Sec. 7.1.7) and also in Aoki (1996, p. 225) where Derrida's random energy model was introduced to the economic audience. Unfortunately it did not catch the attention of the economic audiences. See Mertens (2000) for a simple example, or Krpisvsky et al (2000). This paper is another attempt at exposing non-self-averaging phenomena in economics, in particular in problems involving combinatorial optimization. We also have mentioned a possibility of extending the phrase to cover existence of non-degenerate distributions with time going to infinity. What are the implications if some economic models have non-self averaging property? For one thing, it means that we cannot blindly try for larger size samples in the hope that we obtain better estimates.

The example above is just an indication of the potential of this approach of using exchangeable random partition methods. It is the opinion of this author that subjects such as in the papers by Fabritiis, Pammolli, and Riccaboni (2003), or by Amaral et al (1998) could be re-examined from the random combinatorial partition approach with profit. Another example is Sutton (2002). He modeled independent business in which the business sizes vary by partitions of integers to discuss the dependence of variances of firm growth rates. He assumed each partition is equally likely, however. Use of random partitions discussed in this paper may provide more realistic or flexible framework for the question he examined. It would be an interesting application of the random partition theory and see if non-self-averaging phenomena exist in the sense of physics literture in this area.

References

Aldous, D.J., (1985), "Exchangeability and related topics" in Lecture notes in mathemaatics, No.1117, Springer-Verlarg, Berlin

Amaral, Luis A. Nunes, S. V. Buldyrev, S. Havlin, M.A. Salinger and H.E. Stanley,(1998) "Power law scaling for a system of interacting units with complex internal structure, *Phys. Rev. Lett*,**80**, 1385–1388.

Aoki, M., and H. Yoshikawa (2002). "Demand saturation-creation and economic growth " *J. Econ. Behav. Org.*,**48**, 127-154. Aoki, M., (2000a), "Open models of share markets with two dominant types of participants, ", *J.Econ. Behav. Org.***49** 199-216.

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

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

Aoki, M., T. Nakano, and G. Yoshida, (2004), "Two sector Schumpeterian model of Industry" Mimeo, Dept. Physics, Chuo University, Tokyo.

—–, ——, and K. Ono, (2006), "Simulation Results of a Two-sector Model of Innovation and Immitation"

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

—– , and—-, (2006b),"Stock prices and real economy: Exponential and Power-Law Distributions", forthcoming invited paper *Journal of Interaction and Coordination of Heterogeneous Agents*, No.1, Volume 1. Springer-Verlag, New York.

Arratia, R., and S. Tavaré (1992) "The cycle structure of random permutation", *Ann. prob.* **20**,1567-1591.

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

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

Canning, D., L.A.N.Amaral, Y. Lee, M. Meyer, H. E. Stanley, (1998) " Scaling the volatility of GDP growth rates", Econ. Lett., **60**, 335-341.

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

Derrida, B.,(1981) "Random energy model", *Phys. Rev. B*, **24**, 2613-2626.

Derrida, B., (1994a), "From Random Walks to Spin Glasses",*Physica D*, **107**, 166-198.

——, (1994b) " Non-self-averaging effects in sums of random variables, spin glasses, random maps and random walks" in *On Three Levels Micro-Meso- and Macro-Approaches in Physics*, M. Fannes, C. Maies, and A. Verberre (eds), Plenum Press, New York

—–, and H. Flyvbjerg (1987), "The random map model: a disordered model with deterministic dynamics", *J. Physique*, **48**,971-978.

Devroye, (1993), " A triptych of discrete distributions related to the stable law", *Probab. Letters* **18**, 349–351.

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

—–, (1979), *Mathematical Population Genetics* , Springer-Verlag, Berlin.

—–, (1990) "Population genetics theory —The past and the future", in *Mathematical and statistical problems in evolution* ed. by S. Lessard, Kluwer Academic Pulbishers, Boston.

Fabritiis, G.de, F.Pammolli, and M. Riccaboni, (2003) "On size and growth of business firms," *Physica A***324**, 38–44.

Feller, W., (1949) "Fluctuation theory of recurrent events" *Trans. Am. Math. Soc*,**67**, 98-119.

Griffiths, R., (2005), " Poisson Dirichlet Process" Version o.1 Mimeo.

Higgs, P. (1995), "Frequency distributions in population genetics parallel those in statistical physics", *Phy. Rev. E* **51**, 95-101.

Hirschman, A.O. (1960)"The paternity of an index", *Amer. Econ. Review* bf 54 761.

Huang, Z-F, and S.Solomon (2001), " Power, Lévy, Exponential and Gaussian Regimes in Autocatalytic Financial Systems," *Euro.Phys. Jou. B*, **20**, 601–607.

Iwai, K.(1997)."A contribution to the evolutionary theory of innovation, imitation and growth,",*J. Econ. Behav. Org.*, **43**, 167-198

——- (2001), "Schumpeterian dynamics: A disequilibrium theory of long run profits" in L.Punzo (ed) Cycles, Growth and Structural Change: Theories and empirical evidence, Routledge, London and New York.

Karlin, S. (1967), "Central limit theorem for certain infinite urn schemes", *J. Math. Mech.*, **17**, 373-401.

Kingman, J.F.C. (1978), "The representtion of partition structure", *J. London Math. Soc.*,**18**, 374–380. Krapivsky, P.L., I. Grosse, and E. Ben-Nadin, (2000). "Scale invariance and lack of self-averaging in fragmentation", *Phy. Rev. E*, **61**, R993-R996.

Mekjian, A.Z.,and K.C. Chase (1997), *Phy. Lett A* **229**,340-346.

Mertens, S. (2000) "Random costs in combinatorial optimization", *Phy. Rev. Lett.* **84**, 1347-1350.

Pitman, D (2002) "Sequential construction of random partitions" Lecture notes, St.Flour Summer Institute.

——, (1996) "Random discrete distributions invariant under size-biased permutation", *Adv. Appl. Prob.* **28**, 525–539.

—-, and M.Yor (1997), " The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator," *Ann. prob.* **25**, 855-900

Sherer, F. M. (1980) *Industrial Market Structure and Economic Performance* , 2nd ed. Houghton Mifflin Co. Boston

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

Yamoato H., M. Sibuya, (2000), "Moments of some statistics of Pitman sampling formula",*Bull. Inform. Cybernet.* **32**, 1-10.

Watterson, G. A., (1974), "The Sampling Theory of Selectively Neutral Alleles," *Adv. Appl. Probab.* **6**, 463–488.

—–, and H. A. Guess (1977), "Is the most frequent allele the oldest ? " *Theor. Pop. Biol.* **11**, 141-160.

Appendices

Markov Chains

We can construct Markov chains using the transition probabilities of (1) and (2). Some special cases of these equations for the case $\alpha = 0$ have been simulated by Aoki (2002, Sec. 8.6). We give some details later in this paper as **Example 2**. More extensive examples are to be found in the forthcoming book by Aoki and Yoshikawa (2006).

The probability generating function

With only one state variable n, the probability generating function is defined by $G(z,t) = \sum_k z^n P(n,t)$. Its partial differential equation is obtained by noting that that

$$\sum_k z^k P_{k-1}(t) = zG(z,t),$$

$$\sum_k = 1^\infty (k+1) z^k P_{k+1}(t) = \partial G(z,t)/\partial z,$$

$$\sum_{k=1}^\infty k z^k P_k(t) = z\partial G(z,t)/\partial z,$$

and

$$\sum_{k=1}^\infty (k-1) z^k P_{k-1}(t) = z^2 \partial G/\partial z.$$

With two state variables n_1 and n_2, similar relations. The result is

$$\frac{\partial G}{\partial t} = [d_1(1-z_1) + c_1 z_1(z_1-1) + \mu g_2(z_2-z_1)]\frac{\partial G}{\partial z_1}$$

$$+[d_2(1-z_2) + c_2 z_2(z_2-1) + \mu g_1 h(z_1-z_2)]\frac{\partial G}{\partial z_2}$$

$$+[\mu g_1 z_1(z_1-z_2) + \mu g_2 z_2(z_2-z_1)]\frac{\partial^2 G}{\partial z_1 \partial z_2} + f(z_1-1)G.$$

The cumulant generating function

Noting that

$$\frac{\partial G}{\partial t} = G\frac{\partial K}{\partial t},$$

$$\frac{\partial G}{\partial z_i} = -Ge^{\theta_i}\frac{\partial K}{\partial \theta_i},$$

$i = 1, 2$, and

$$\frac{\partial^2 G}{\partial z_1 \partial z_2} = Ge^{\theta_1+\theta_2}H$$

with

$$H = \frac{\partial K}{\partial \theta_1}\frac{\partial K}{\partial \theta_2} + \frac{\partial^2 K}{\partial \theta_1 \partial \theta_2},$$

we convert the partial differential equation for G into that for K

$$\frac{\partial K}{\partial t} = \frac{1}{G}\frac{\partial G}{\partial t} = -\sum_{i=1}^2 [d_i(e^{\theta_i}-1) + c_i(e^{-\theta_i}-1)\frac{\partial K}{\partial \theta_i} + f(e^{-\theta_1}-1)$$

$$+\mu[g_1(e^{(\theta_2-\theta_1}-1) + g_2(e^{(\theta_1-\theta_2)}-1)]H.$$

Extract coefficients of θ_i and equate them to dk_i/dt, $i = 1, 2$, and those of θ_1^2, θ_2^2 with the derivatives $dk_{1,1}/dt$ and $dk_{2,2}/dt$, and the coefficient of $\theta_1\theta_2$ with the derivative $dk_{1,2}/dt$.

In this way we generate a set of five differential equations for $k_1, k_2, k_{1,1}, k_{2,2}$, and $k_{1,2}$.

Calculations of the variances and covariance

The equations for the variance and covariance are derived as follows:

$$\begin{aligned}
\dot{k}_{11} &= f - 2d_1(1-g_1)k_{11} + d_1(1+g_1)k_1 \\
&\quad +\lambda d_2 f(2k_{12}+k_2) + 4\alpha\lambda(k_1k_{12}+k_2k_{11}) + 2\beta\lambda A_0, \qquad (21)\\
\dot{k}_{22} &= -2d_2(1-g_2+\lambda f)k_{22} + d_2(1+g_2+\lambda f)k_2 \\
&\quad -4\alpha\lambda(k_1k_{22}+k_2k_{12}) + 2\beta\lambda A_0, \qquad (22)\\
\dot{k}_{12} &= -[d_1(1-g_1)+d_2(1-g_2+\lambda f)]k_{12} + \lambda d_2 f(k_{22}-k_2) \\
&\quad -2\alpha\lambda(k_1k_{12}+k_2k_{11}-k_1k_{22}-k_2k_{12}) - 2\beta\lambda A_0, \qquad (23)
\end{aligned}$$

where $\beta = d_1d_2(g_1+g_2)/2$. Stationary values of variances k_{11} and k_{22} are obtained by setting the left hand sides of (21) and (22) equal to zero:

$$\begin{aligned}
k_{11} &= \frac{1}{2\gamma_1\left(n+\frac{n-m}{m}\theta\right)}\left\{\frac{2}{n\gamma} + \frac{2}{\gamma\gamma_1 mn^2(n-m)}\Big[\gamma_1 n(m^2+n)\right. \\
&\left. -\gamma m\left(\gamma_1 n^2(1+m) - m(1+n)\right)\Big]\theta + \frac{2(2+n)}{\gamma_1 n^2}\theta^2 + \frac{2(n-m)}{\gamma_1 mn^2}\theta^3\right\}, \qquad (24)
\end{aligned}$$

and

$$\begin{aligned}
k_{22} &= \frac{1}{2\gamma_2\left(1+m-\frac{n-m}{n}(1-\theta)\right)} \\
&\quad \times\left\{\frac{2(1-m\gamma)(1+n)}{m(n-m)\gamma}\theta + \frac{2(1+n)}{mn\gamma_2}\theta^2 + \frac{2(n-m)}{m^2n\gamma_2}\theta^3\right\}. \qquad (25)
\end{aligned}$$

Self consistent values of v

Substituting (24) for k_{11}, (25) for k_{22} and

$$k_{12} = A_0 - k_1k_2 = \frac{1}{\gamma_1\gamma_2}\frac{1+m}{m(n-m)}\theta\left[1 - \frac{n-m}{n(1+m)}(1-\theta)\right]$$

into the equation which is derived by setting the left hand side of (23) equal to zero yields the fifth order equation for θ. Luckily, however, the highest term vanishes, so that the equation becomes quartic;

$$\theta F(\theta) = 0 \qquad (26)$$

where

$$F(\theta) = r_0 + r_1\theta + r_2\theta^2 + r_3\theta^3. \qquad (27)$$

Here

$$r_0 = \frac{m(1+n)}{\gamma(n-m)}\Big\{\gamma_1\gamma_2\Big[-m+2n(1+n)\Big]+\gamma\Big[\gamma_1 n^2(1+n)$$
$$+\gamma_2\Big(-\gamma_1 n^2(1+n)+m[1+2n+(1-\gamma_1)n^2]\Big)\Big]\Big\}, \quad (28)$$

$$r_1 = -\frac{1}{\gamma m}\Big\{\gamma_1\gamma_2\Big[n^3-4mn(1+n)+m^2(2+n)\Big]+\gamma m\Big[-\gamma_1 n(1+5n+4n^2)$$
$$+\gamma_2\Big(\gamma_1 n^2(2+n)-m[4+8n+(4-\gamma_1)n^2]\Big)\Big]\Big\} \quad (29)$$

$$r_2 = -\frac{n-m}{\gamma m^2}\Big\{\gamma_1\gamma_2(n-m)^2-\gamma m\Big[5\gamma_2 m(1+n)+\gamma_1 n(3+5n)\Big]\Big\}, \quad (30)$$

$$r_3 = \frac{2(n-m)^2}{m^2}\Big\{\gamma_2 m+\gamma_1 n\Big\}. \quad (31)$$