

Cluster Size Distributions of Economic Agents of Many Types in Market

Masanao Aoki*

Department of Economics

University of California, Los Angeles

Fax Number 310-825-9528, e-mail aoki@econ.ucla.edu

February, 2000

Abstract

We use partition vector as state vector in models composed of a large number of exchangeable agents of possibly many types. To illustrate the use of this notion we derive size distributions of clusters (groups) formed by these interacting economic agents.

1 Introduction

Ricard Bellman was instrumental in introducing the notion of state to many disciplines including economics. He was also one of the leading pioneers in bringing new or non-traditional mathematical tools to analysis of economic models, many of which are now in the tool kits of some segments of economists. In view of his important contributions to economics, it is felt appropriate that this paper illustrates the use of partition vector, which is originally developed in ecology, genetics and statistics, in analysis of economic models composed of many interacting 'exchangeable' agents, because partition vectors characterize random exchangeable partitions.

The view of building macroeconomic models from interacting microeconomic agents has certainly been known in the economic literature, but has not been vigorously pursued in the recent past by the economic profession. Many economists still persist in using the discredited construct of representative agents in their models possibly for the supposed convenience in manipulating models. See Kirman (1992) for an example of critiques. Recently,

*Prepared for the Bellman issue of Journal of Mathematical Analysis and Applications. The author wishes to thank Profs. Costantini and Garibaldi for elucidating fine points of the Ewens distribution.

a new life has been brought into this area, however, by recent interdisciplinary efforts from statistical physics, stochastic processes, and population genetics. These efforts are reported under various names such as evolutionary economics, econo-physics, or more simple as agent-based modelings or heterogenous agents. This paper should be regarded in this light.

The word partition vector was apparently coined by Zabell (1992) even though the notion originates with Kingman (1977) in his work on exchangeable random partitions which generalizes the notion of random sequences introduced by de Finetti. Sachkov (1997) calls it specification of the second kind.

We consider a collection of a large number of interacting economic agents, such as firms, households, and sectors of economies, or even countries. To simplify our exposition suppose there are a fixed number n of them. They belong to a potentially large number of different types. Here, the word 'type' should be broadly interpreted. For example, firms belong to different categories in terms of capitalization, size of employees, plants and equipments, households may be classified into different categories or types by their income levels, places they purchase various goods, or more generally by their demand patterns in general, and so on. The word could even mean that agents are classified by the kinds of algorithms they use in their decision makings, searches for jobs, and so on. The types generally change over time. Agents may change their mind in their stock purchase patterns and switch their strategies, or may become bearish from bullish or vice versa.

We associate different types with each of the discrete choices agents make. Since agents in our models make discrete choices over time, we model the process of agent choice patterns as discrete-state continuous time Markov processes, also called jump Markov processes. Aoki (1996, 1998, 1999a, b) has several economic examples.

2 Jump Markov Process Models

Suppose there are potentially a large number K of types in a model of market. Let $\mathbf{n} = (n_1, n_2, \dots, n_K)$ describe how n agents are distributed over K types, $n = n_1 + n_2 + \dots + n_K$. Initially for ease of explanation we use this vector as state vector. Excluding pathological phenomenon of an infinitely many jumps in a small time interval, it is known that jump Markov processes are specified by transition rates. See Breiman (1969) for example.

In our model we need to specify entry rates, exit rates and rates of type changes.¹ They are:

$$w(\mathbf{n}, \mathbf{n} + \mathbf{e}_k) = c_k(n_k + h_k),$$

for $n_k \geq 0$, where \mathbf{e}_k is the vector with unit element in the k th position which

¹See Cox and Miller (1965, Chapt. 4) for related models.

is the only non-zero element of the vector,

$$w(\mathbf{n}, \mathbf{n} - \mathbf{e}_j) = d_j n_j,$$

$n_j \geq 1$, and

$$w(\mathbf{n}, \mathbf{n} - \mathbf{e}_j + \mathbf{e}_k) = \lambda_{jk} d_j n_j c_k (n_k + h_k),$$

with $\lambda_{jk} = \lambda_{kj}$, and where $j, k = 1, 2, \dots, K$. We assume that $d_j \geq c_j > 0$, and $h_j > 0$, and $\lambda_{jk} = \lambda_{kj}$ for all j, k pairs.

The first transition rate specifies entry rate of type k agents, and the second that of the exit or departure rate by type j agents and the last specifies the transition intensity of changing types by agents from type j to type k . In the entry transition rate specification $c_k n_k$ stands for attractiveness of larger group, such as network externality which makes it easier for others to join the cluster or group, and $c_k h_k$ stands for the innovation effects which is independent of the group size. Sutton (1992) calls this niche effects. The niche effects are something that affects a limited segments of economy such as a specific group or clusters. Transition rates for type changes are also found in Kelly (1979). As Sutton remarks, the classical models of firm growth such as Gibrat's law are not empirically supported. We need interactions or correlations among agents. The same is true for size distributions of groupings of agents. Parameter θ introduced in the next section plays this role. The jump Markov process thus specified has the steady state or stationary distribution

$$\pi(\mathbf{n}) = \prod_{j=1}^K \pi_j(n_j),$$

where

$$\pi_j(n_j) = (1 - g_j)^{-h_j} \binom{-h_j}{n_j} (-g_j)^{n_j}$$

where $g_j = c_j/d_j$.

These expressions are derived straightforwardly by applying the detailed balance conditions to the transition rates. See Kelly (1979, Chapt.1) for example.

To provide simpler explanation, suppose that $g_j = g$ for all j . Then, noting that $\prod_j (1 - g)^{-h_j} = (1 - g)^{-\sum_j h_j}$, the joint probability distribution is expressible as

$$\pi(\mathbf{n}) = \binom{-\sum h_k}{n}^{-1} \prod_{j=1}^K \binom{-h_j}{n_j}. \quad (1)$$

3 Partition Vectors and Ewens Distribution

Now, we introduce the partition vector $\mathbf{a} = (a_1, a_2, \dots, a_n)$, where a_k is the number of types or clusters with exactly k agents. Consequently we have an inequality

$$\sum_i a_i = K_n \leq K,$$

where K_n is the number of groups or clusters formed by n agents, and

$$\sum_i i a_i = n,$$

which is an accounting identity.

To further simplify our presentation, let us suppose that $h_j = h$ for all j .² Then

$$\pi(\mathbf{n}) = \binom{-Kh}{n}^{-1} \prod_{j=1}^K \binom{-h}{j}^{a_j}.$$

This is so because there are a_j of the n s which equal j .

Now suppose that K becomes very large and h very small, while the product Kh approaches a positive constant θ . We note that the negative binomial expression

$$\binom{-h}{j}^{a_j}$$

approaches $(h/j)^{a_j} (-1)^{j a_j}$ as h becomes smaller. Suppose $K_n = k \leq K$. Then, there are

$$\frac{K!}{a_1! a_2! \cdots a_n! (K-k)!}$$

many ways of realizing a vector. Hence

$$\pi(\mathbf{a}) = \binom{-\theta}{n} (-1)^n \frac{K!}{a_1! a_2! \cdots a_n! (K-k)!} \prod_j \left(\frac{h}{j}\right)^{a_j}. \quad (2)$$

Noting that $K!/(K-k)! \times h^k$ approaches θ^k in the limit of K becoming infinite and h approaching 0 while keeping Kh at θ , we arrive, in the limit, at the probability distribution, known as the Ewens distribution, or Ewens sampling formula very well known in the genetics literature, Ewens (1972), Kingman (1987), or Johnson, Kotz, and Balakrishnan (1997).

$$\pi_n(\mathbf{a}) = \frac{n!}{\theta^{[n]}} \prod_{j=1}^n \left(\frac{\theta}{j}\right)^{a_j} \frac{1}{a_j!},$$

where $\theta^{[n]} := \theta(\theta+1)\cdots(\theta+n-1)$. This distribution has been investigated in several ways. See Arratia and Tavaré (1992), or Hoppe (1987). Kingman (1980) states that this distribution arise in many applications. There are other ways of deriving this distribution. See Appendix. We next examine some of its properties.

²This assumption is not necessary. All we need is that the product Kh_j has the same positive limit as K goes to infinity and h_j to zero.

3.1 The number of clusters and value of θ

Ewens sampling formula has a single parameter θ . Its value influences the number of clusters formed by the agents. Smaller values of θ tends to produce a few large clusters, while larger values produce a large number of smaller clusters.

To obtain some quick feels for the influences of the value of θ , take $n = 2$ and $a_2 = 1$. All other a s are zero. Then

$$\pi_2(a_1 = 0, a_2 = 1) = \frac{1}{1 + \theta}.$$

This shows that two randomly chosen agents are of the same type with high probability when θ is small, and with small probability when θ is large. In fact, θ controls correlation between agents' types or classification.

Two extreme situations may convey the relation between the value of θ and the number of clusters. We note that the probability of n agents forming a single cluster is given by

$$\pi_n(a_j = 0, 1 \leq j \leq (n - 1), a_n = 1) = \frac{(n - 1)!}{(\theta + 1)(\theta + 2) \cdots (\theta + n - 1)}$$

while the probability that n agents form n singleton is given by

$$\pi_n(a_1 = n, a_j = 0, j \neq 1) = \frac{\theta^{n-1}}{(\theta + 1)(\theta + 2) \cdots (\theta + n - 1)}.$$

With θ much smaller than one, the former probability is approximately equal to 1, while the latter is approximately equal to zero. When θ is much larger than n the opposite is approximately true.

We can show that

$$P_n(K_n = k) = \frac{1}{\theta^{[n]}} c(n, k) \theta^k,$$

where $c(n, k)$ is known as the signless Stirling numbers of the first kind, and is defined by

$$\theta^{[n]} = \sum_1^n c(n, k) \theta^k.$$

See Hoppe (1987) for the derivation. Stirling numbers are discussed in van Lint and Wilson (1992, 104) for example.

3.2 Fractions

The expected value of a_j is given by

$$E(a_j) = \sum_{w(n)} a_j \pi_n(\mathbf{a}) = \frac{\theta}{j} \frac{n!}{(n - j)!} \frac{\theta^{[n-j]}}{\theta^{[n]}}, \quad (3)$$

where $w(n) := \{\mathbf{a} : \sum_j ja_j = n\}$. A quick way to see this is to recognize that the multiplication of π_n by a_j is equivalent to changing the partition vector from \mathbf{a} to \mathbf{a}' where the j th component is changed from a_j to $a_j - 1$ and the sum $\sum ia_i = n$ is changed to $\sum ia'_i = n - j$, because $a'_i = a_i$, $i \neq j$, but $a'_j = a_j - 1$.

Watterson (1976) shows how to calculate the variance and covariances by using the relation

$$E\{a_j(a_j - 1)\} = \left(\frac{\theta}{j}\right)^2 \frac{n!}{(n-2j)!} \frac{\theta^{[n-2j]}}{\theta^{[n]}},$$

and for $i \neq j$

$$E\{a_i a_j\} = \frac{\theta^2}{ij} \frac{n!}{(n-i-j)!} \frac{\theta^{[n-i-j]}}{\theta^{[n]}}.$$

What is characteristic is the fact that the standard deviations of a s are of the same order of magnitude as the means.

With θ small, we have approximate relations

$$E(a_j) \approx \frac{\theta}{j} \frac{n}{n-j} \frac{1}{1 + \theta h_{n-1, n-j}},$$

where $h_{n-1, n-j} := 1/(n-1) + 1/(n-2) + \dots + 1/(n-j)$. Similarly

$$E(a_j^2) - E(a_j) \approx \left(\frac{\theta}{j}\right)^2 \frac{n}{n-2j} \frac{1}{1 + \theta h_{n-2j, n-j}},$$

where $h_{n-2j, n-j} := 1/(n-j) + 1/(n-j-1) + \dots + 1/(n-2j)$.

Let $x_i = ia_i/n$. It is the fraction of agents of type i . It is a point in the $(K-1)$ -dimensional simplex, $\sum_{i=1}^K x_i = 1$ with $x_i \geq 0$, $i = 1, 2, \dots, K$. Its expected value is

$$E(x_j) = E(ja_j/n).$$

The order statistics of the fractions, $x_{(1)} \geq x_{(2)} \geq \dots$ are important in markets with highly correlated agents, because with θ smaller than 1, the sum of two or three largest fractions can be shown to be nearly one. See tables in Watterson and Guess (1977). This means that two or three large groups dominate the market behavior. For example, excess demands for some financial assets by these large groups may largely determine the market excess demand, and hence the price dynamics. See Aoki (1999) for an example of such a model.

3.3 Limiting Behavior

Arratia et al. (1992, 1994) have shown that the components of the partition vector converges in distribution to (Z_1, Z_2, \dots) as n goes to infinity where Z_i s are independent Poisson random variables with mean θ/i , and

$$\Pr(\mathbf{a}) = \Pr(z_i = a_i, i = 1, 2, \dots, n | \sum_{j=1}^n jZ_j = n).$$

They also have a bound

$$E(a_j) - \theta/j \leq \frac{\theta(1 - \theta)}{\theta + n - j}.$$

When we use partition vectors to characterize models with exchangeable random partitions, the multivariate Ewens distribution (MED in Johnson et al. (1997, 232) govern the distributions of the components. Here we give a couple of ways this distributions may be derived.

3.4 Expected Value of the Largest Share

Suppose that fractions x_1, x_2, \dots describe the population composition by types of exchangeable agents. Recall that the subscript 1, 2, and so on have no intrinsic meaning, but a mere convenience in referring to different clusters or groupings of agents by types. Denote the largest fraction by $x_{(1)}$. Here we follow Watterson (1976), and Watterson and Guess (1977) to show how to calculate its expected value. An entirely analogous procedure can calculate the joint probability density for r order statistics of the fractions. This is discussed in the section following this.

We can assume that x_K is $x_{(1)}$, i.e., assume without loss of generality that

$$0 \leq x_i \leq x_K = 1 - x_1 - \dots - x_{K-1},$$

for $i = 1, \dots, K - 1$. Change variables from xs to

$$y_i = x_i / (1 - x_1 - x_2 - \dots - x_{K-1}) \leq 1.$$

Then, we have

$$E(x_{(1)}) = K \int \dots \int p_K \phi(x_1, \dots, x_K) dx_1 dx_2 \dots dx_{K-1},$$

where ϕ is the symmetric Dirichlet distribution with parameter ϵ . Accordingly we assume that

$$\phi(x_1, x_2, \dots, x_K) = \frac{\Gamma(K\epsilon)}{(\Gamma(\epsilon))^K} \prod_{i=1}^K x_i^{\epsilon-1}$$

is the joint density function of xs . See Appendix on Dirichlet distributions. The use of Dirichlet distribution is based on a deep connection of this distribution with the representation of random exchangeable partition introduced by Kingman (1978), and later expounded by Zabell (1992). We do not stop here to explain these but go directly to calculate the expected size of the largest fraction governed by the Dirichlet distribution.

Noting that $1 - x_1 - \dots - x_{K-1} = (1 + \sum_{j=1}^{K-1} y_j)^{-1}$, and that the Jacobian is

$$\frac{\partial \mathbf{x}}{\partial \mathbf{y}} = (1 + \sum_j y_j)^{-K}$$

by straightforward calculation, we rewrite

$$E(x_{(1)}) = \frac{K\Gamma(K\epsilon)}{(\Gamma(\epsilon))^K} \int \cdots \int \prod_{j=1}^{K-1} y_j^{\epsilon-1} (1 + \sum_j y_j)^{-1-K\epsilon} dy_1 \cdots dy_{K-1},$$

where the range of integration is over the cube, $0 \leq y_j \leq 1$, $j = 1, \dots, K-1$. By multiplying the numerator and denominator by ϵ^{K-1} , we regard this integral as the expected value of $(1 + Z_{K-1})^{-1-K\epsilon}$ with respect to Y_j , $j = 1, \dots, K-1$ which are independent and identically distributed with the density $\epsilon^{K-1} \prod_{j=1}^{K-1} y_j^{\epsilon-1}$.

Although we cannot express the expected value in closed form we can give the Laplace transform of Z_{K-1} and obtain its limit as K goes to infinity. From independence of Y s

$$E(e^{-sZ_{K-1}}) = \{E(e^{-sY_1})\}^{K-1} = \left\{ \int_0^1 e^{-sy} \epsilon y^{\epsilon-1} dy \right\}^{K-1}.$$

Write the integrand as

$$\epsilon e^{-sy} y^{\epsilon-1} = \epsilon \{y^{\epsilon-1} + (e^{-sy} - 1)y^{\epsilon-1}\}$$

we let K goes to infinity and ϵ to zero while the product goes to a nonnegative value θ . Thus

$$E(e^{-sZ_{K-1}}) \rightarrow \exp\left\{\theta \int_0^1 (e^{-sy} - 1)y^{-1} dy\right\}.$$

Denote by Z the limiting random variable with the Laplace transform specified above. Then

$$E(x_{(1)}) = e^{\gamma\theta} \Gamma(\theta + 1) E\{(1 + Z)^{-1-\theta}\}.$$

In deriving this expression we note that $\Gamma(K\epsilon) \rightarrow \Gamma(\theta)$, $\Gamma(\epsilon) = \Gamma(1 + \epsilon)/\epsilon$, and that $\Gamma(1 + \epsilon) \approx \Gamma(1) + \Gamma'(1)\epsilon = 1 - \gamma\epsilon$, where $\gamma = -\Gamma'(1)$ is Euler's constant. See Abramovitz and Stegun [(1968, (5.1.1) and (5.1.11)]. By definition

$$\Gamma(1 + \theta) = \int_0^\infty x^\theta e^{-x} dx,$$

and we note that

$$\Gamma(\theta + 1)(1 + Z)^{-1-\theta} = \int_0^\infty s^\theta e^{-(1+Z)s} ds.$$

Taking expectation with respect to Z , we obtain

$$E(x_{(1)}) = e^{\gamma\theta} \int_0^\infty s^\theta e^{-s} E(e^{-sZ}) ds = \int_0^\infty e^{-s} e^{-\theta E_1(s)} ds,$$

where we use the relation

$$E(e^{-sZ}) = \exp\{-\gamma\theta - \theta \ln t - \theta E_1(s)\},$$

and where

$$E_1(s) = \int_s^\infty e^{-x} x^{-1} dx = -\gamma - \ln s - \sum_{n=1}^{\infty} \frac{(-s)^n}{nn!} = -\gamma - \ln s - \int_0^\infty (e^{-sy} - 1) dy.$$

The marginal probability density of the largest fraction is

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

for $1/2 < x \leq 1$. When x is not greater than $1/2$, the expression is more complex

$$f(x) = \Gamma(\theta + 1) e^{\gamma\theta} x^{\theta-2} g(x^{-1} - 1),$$

where $g(\cdot)$ is the density of the random variable Z introduced above, and characterized in terms of its Laplace transform.

3.5 Joint probability density for r largest fractions

We next derive the joint probability density for the largest r fractions on the K -dimensional simplex, $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(r)}$, where $x_i, i = 1, 2, \dots, K$ are the fractions. Denote the Dirichlet probability density on the K dimensional simplex by $\phi(x_1, x_2, \dots, x_k) = \mathcal{D}(\epsilon, K)$. Then, the probability density for the first r order statistics is given by

$$f(x_1, x_2, \dots, x_r) = K(K-1)(K-2) \dots (K-r+1) \int \phi(x_1, x_2, \dots, x_k) dx_{r+1} \dots dx_{K-1},$$

where $1 \geq x_1 \geq x_2 \dots \geq x_r > 0$, and where we substitute $x_K = 1 - x_1 - \dots - x_{K-1}$. Carrying out the integral

$$f(x_1, \dots, x_r) = \frac{K!}{(K-r)!} \frac{\Gamma(K\epsilon)}{\Gamma(\epsilon)^K} (x_1 \dots x_r)^{\epsilon-1} x_r^{(K-r)\epsilon-1} I,$$

where

$$I = \int \dots \int \prod_{r+1}^{K-1} y_j^{\epsilon-1} [1 - a - x_r(y_{r+1} + \dots + y_{K-1})]^{\epsilon-1} dy_{r+1} \dots dy_{K-1},$$

where the integration is carried out in area A

$$A = \{0 \leq y_j \leq 1, r+1 \leq j \leq K-1; (1-a-x_r)/x_r \leq y_{r+1} + \dots + y_{K-1} \leq (1-a)/x_r\}$$

where $a := x_1 + x_2 + \dots + x_r$.

As in the case of the largest fraction, introduce random variable Z with the density function g_{K-r-1} which is the $(K-r-1)$ -fold convolution of the density $\epsilon y_j^{\epsilon-1}$, $j = r+1, \dots, K-1$.

The integral is approximately given by

$$I = \frac{\epsilon^{-(K-r)}}{x_r} g_{K-r-1}\left(\frac{1-a}{x_r}\right).$$

Letting $K\epsilon$ approach θ , while K goes to infinity and ϵ to zero, we note that $K(K-1)\cdots(K-r+1)$ approaches K^r , and

$$\Gamma(\epsilon)^K = \left[\frac{\Gamma(1+\epsilon)}{\epsilon} \right]^K$$

which approaches $\epsilon^{-K}e^{-\gamma\theta}$, where we use the fact

$$\Gamma(1+\epsilon) \approx 1 - \gamma\epsilon,$$

where γ is the Euler's constant, $\gamma = .5772\cdots$.

Putting them all together we arrive at

$$f(x_1, x_2, \dots, x_r) = \theta^r e^{\gamma\theta} \Gamma(\theta) x_r^{\theta-1} g\left(\frac{1-a}{x_r}\right) (x_1 x_2 \cdots x_r)^{-1},$$

in the range $1 \leq x_1 \leq \cdots \leq x_r > 0$, and $\sum_1^r x_i \leq 1$.

We know from our result for the largest fraction that $f(x_1) = \theta x_1^{-1} (1 - x_1)^{\theta-1}$ for x_1 between $1/2$ and 1 , that is

$$\Gamma(1+\theta) e^{\gamma\theta} g\left(\frac{1-x_1}{x_1}\right) = \theta \left(\frac{1-x_1}{x_1}\right)^{\theta-1},$$

for x_1 between $1/2$ and 1 .

To obtain the expression for the density in the range $0 \leq x \leq 1/2$, we follow Watterson (1976) and differentiate the Laplace transform for the random variable Z

$$E(e^{-sZ}) = \exp\left[\theta \int_0^1 (e^{-sz} - 1) z^{-1} dz\right]$$

, with respect to s . Recall that this transform is derived in connection with the largest fraction. Then, divide the result by $-\theta$ to see

$$(1/\theta) \int_0^\infty e^{-sz} z g_\theta(z) dz = \int_0^\infty I_{(0,1]}(y) e^{-sy} dy \int_0^\infty e^{-sz} g_\theta(z) dz.$$

The right hand side is the product of two Laplace transforms. Hence the integrand of the left-hand side is the convolution of the uniform function on the unit interval and g_θ

$$(1/\theta) z g_\theta(z) = \int_{z-1}^z g_\theta(y) dy.$$

Setting z to 1 , we obtain

$$g_\theta(1) = e^{-\gamma\theta} / \Gamma(\theta).$$

Differentiating the integral equation with respect to z , we derive the differential equation which determines the function recursively

$$z g'_\theta(z) + (1-\theta) g_\theta(z) = -\theta g_\theta(z-1),$$

where $z \geq 0$. In the range $z \in [0, 1)$, this integro-differential equation yields the result we obtained above. In the next range $z \in [1, 2)$ we have

$$g_\theta(z) = z^{\theta-1} [g_\theta(1) - \theta \int_1^z g_\theta(u-1) u^{-\theta} du].$$

Changing variable of integration to $v = 1/u$, we note that the integration above becomes

$$\int_{1/z}^1 v^{-1} (1-v)^{\theta-1} dv.$$

The joint density for the first two largest fractions is given by

$$f_\theta^{(2)}(x, y) = \frac{e^{\gamma\theta} \theta^2 \Gamma(\theta) y^{\theta-1}}{x y} g_\theta\left(\frac{1-x-y}{y}\right) = \frac{\theta^2}{xy} (1-x-y)^{\theta-1}.$$

This expression is valid for the range $0 < y < x < 1$, $0 < x + y < 1$, and $x + 2y > 1$, that is $y > (1-x)/2$.

We know that

$$g(z) = \frac{z^{\theta-1}}{\Gamma(\theta) e^{\gamma\theta}},$$

for z between 0 and 1. For other values of z , we have a recursion

$$\frac{zg(z)}{\theta} = \int_{z-1}^z g(y) dy,$$

see Watterson and Guess (1977). Alternatively put, we have

$$g(z) = z^{\theta-1} [g(n) n^{1-\theta} - \theta \int_n^z g(y-1) y^{-\theta} dy],$$

in the range $n \leq z < n+1$. This can be verified by direct substitution into the differential equation for g_θ .

With the joint probability densities for the largest 2 or 3 fractions in the case of a large number of correlated agents in a market, we are in a position to approximately obtain the macroeconomic behavior of the market prices and quantities.

3.6 Examples of Alternative Distributions

So far, we have introduced Ewens distribution, well known in population genetics literature but unknown in economics, as a potentially useful distribution to describe market behavior. Here we derive some other size distributions with the same applications in mind.

3.6.1 Two-sided exponential distribution

In Amaral et al. (1997) empirical distributions for growth rates by firms have been found to have a two-sided exponential distributions, and they have

derived it as a steady state distribution of certain Fokker-Planck equation based on the hypothesis that firm sizes grow if its size is less than some desired size by a random factor, that is $S_{t+\Delta t} = S_t k e^{\epsilon_t}$ where ϵ_t is a mean zero random variable, and if S_t is greater than the desired size, then it shrinks by the same factor, $S_{t+\Delta t} = S_t (1/k) e^{\epsilon_t}$. From this the authors derive the equilibrium growth rate distribution to be two sided exponential function

$$p(r) = \text{const} \exp\{-c|r - \hat{r}|\},$$

where \hat{r} is mean growth rate and c is related to variance. Note that even though they started out with growth in size, they convert it to growth rates. They have done extensive empirical work in this and related work to support this form.

Here, we look at the size distribution itself assuming that sizes measured in some basic units are proportional to n , and there is a desired value for n , denoted by n^* . We interpret size broadly as Amaral et.al do. It may refer to the sales, amount of capitalization, numbers of employees etc. We apply our jump Markov process machinery and show that the steady state distribution of n is a two sided exponential about n^* .

Suppose that the transition rates are specified by

$$w(n, n + 1) = \alpha,$$

if $n < n^*$, and by

$$w(n, n + 1) = \beta,$$

if $n > n^*$. Assume that $\alpha > \beta$. Also assume that

$$w(n, n - 1) = \beta,$$

when $n < n^*$, and when $n > n^*$, then

$$w(n, n - 1) = \alpha$$

These transition rate specifications capture the same idea as in Amaral et. al. By the detailed balance condition we obtain

$$\pi_n = c_1 \left(\frac{\alpha}{\beta}\right)^n,$$

for $n < n^*$, and for $n > n^*$

$$\pi_n = c_2 \left(\frac{\beta}{\alpha}\right)^{n-n^*}.$$

This basic specification can be modified to let the distance $|n - n^*|$ influence the transition rates. One such example would be

$$w(n, n + 1) = \frac{e^{-\gamma(n-n^*)}}{e^{-\gamma(n-n^*)} + e^{-\gamma(n^*-n)}},$$

for $n < n^*$, and $w(n, n - 1)$ is specified analogously by replacing the numerator by $e^{-\beta(n^*-n)}$, for example. Here γ may represent uncertainty which prevails in the sector of the economy in which firms are located.

3.6.2 Clusters in partition vector

We adapt some examples from Kelly (1976, 1979).³

There are N basic units partitioned into distinct clusters or collections, with a_i being the number of groups consisting of i units. Recall that we mean by units some basic building blocks from which objects that cluster are made up.

The transition rate $w(\mathbf{a}, \mathbf{a} + e_i) = \alpha$ represents the process in which a basic unit or singleton (called isolate in Kelly (1979, chapt.8)) joins a group of size i at rate α , $i = 1, 2, \dots$. The transition rate $w(\mathbf{a}, \mathbf{a} - e_i) = \beta$ refers to the rate at which a singleton (one basic unit) leaves that group to become an isolate. The vector e_i has 1 in the i th position and zero everywhere else. Call a cluster of size i i -cluster.

We assume \mathbf{a} is a Markov process in which the transition rate $w(\mathbf{a}, \mathbf{a} - e_1 - e_i + e_{i+1}) = \alpha a_1 a_i$, $i \geq 2$. This refers to the rate at which an isolate joins an i -cluster, hence forming one more $i + 1$ -cluster. When two isolates form a new group of size 2, the transition rate is $w(\mathbf{a}, \mathbf{a} - 2e_1 + e_2) = \alpha a_1 (a_1 - 1)$. The rate at which an i -cluster breaks up into an isolate and a cluster of size $i - 1$ is represented by $w(\mathbf{a}, \mathbf{a} + e_1 + e_{i-1} - e_i) = i\beta a_i$, $i \geq 2$. The transition rate $w(\mathbf{a}, \mathbf{a} + 2e_1 - e_2) = 2\beta a_2$ refers to one cluster of size 2 divides into 2 isolates. In a more general setting, suppose that the transition rate of one r -cluster and one s -cluster form one u -cluster. It is written as

$$w(\mathbf{a}, \mathbf{a} - e_r - e_s + e_u) = \lambda_{rsu} a_r a_s,$$

when $r \neq s$. With $r = s$, we specify the transition rate by

$$w(\mathbf{a}, \mathbf{a} - 2e_r + e_u) = \lambda_{rru} a_r (a_r - 1),$$

and

$$w(\mathbf{a}, \mathbf{a} - e_u + e_r + e_s) = \mu_{rsu} a_u$$

is the transition rate of one u -cluster breaking up into one r -cluster and one s -cluster. In the simple example described above, we have

$$\lambda_{1,i,i+1} = \alpha,$$

and

$$\mu_{1,i-1,i} = i\beta,$$

for $i \geq 2$.

Assume that $\lambda_{rsu} = \lambda_{sru}$ and $\mu_{rsu} = \mu_{sru}$. We check the detailed balance conditions and verify that the equilibrium distribution is of the form

$$\pi(\mathbf{a}) = B \prod_r \frac{c_r^{a_r}}{a_r!},$$

³This subsection is based in part on Kelly (1979, Chap. 8). We use a_i rather than m_i used in Kelly.

provided there are positive numbers c_1, c_2, \dots , such that

$$c_r c_s \lambda_{rsu} = c_u \mu_{rsu}.$$

We can easily verify that the detailed balance conditions are satisfied. In the simple example, we note that

$$c_r = \frac{\beta}{\alpha r!}$$

satisfies the detailed balance conditions.

For a closed model with N fixed, the equilibrium distribution

$$\pi(\mathbf{a}) = B \prod_{i=1}^N \frac{1}{a_i!} \left(\frac{\beta}{\alpha i}\right)^{a_i}$$

is an example of assemblies analyzed by Arratia and Tavaré. Here $m_i = (\beta/\alpha)$ serves as the "number" of labelled structures on a set of size i , that is the number of the labelled structure in this example is independent of the size of the block.

In the more general development that follow the example, if we set

$$c_r = \frac{m_r}{r!}$$

then m_r is the number of the labelled structures on a set of size r .

4 Concluding Remarks

This paper presented Ewens and several other distributions as candidates for distributions of clusters sizes formed by a large number of economic agents who interact in a market. Although no application is described in this paper, Aoki (1999) has one simple application in which stocks of a holding company is traded by a large number of agents. With $\theta = .3$, two largest groups are shown to capture nearly 80 per cent of the market shares and hence dominate the market excess demands, which in turn determine the stationary distributions of price.

In this way it is shown that the tail distribution of the market clearing prices obey a power law. Power laws have not played much role in the economics literature. It has been shown in recent research, however, that power laws may prove to be an important device for imposing empirical constraints on macroeconomic models such as rational bubble models. See Lux and Sornette (1999).

References

Abramovitz, M., and I. A. Stegun., *Handbook of Mathematical Functions*, Dover Publications, Inc., New York, 1968

- Amaral, L. A., S. V. Buldyrev, S. Havlin, P. Maass, M. A. Salinger, H. E. Stanley, M.H.R. Stanley, "Scaling behavior in economics: The problem of quantifying company growth", *Physica A*, **244** 1-24, 1997.
- Arratia, R., and A.D. Barbour and S. Tavaré, "Poisson process approximations for the Ewens sampling formula", *Ann. Appl. Probab.* **2** 519-35, 1992.
- Arratia, R., and S. Tavaré, " ", 1994.
- Aoki, M., (1996) *New Approaches to Macroeconomic Modeling: Evolutionary Stochastic Dynamics, Multiple Equilibria, and Externalities as Field Effects*, Cambridge University Press, New York.
- , "Open Models of Share Markets with Several Types of Participants", presented at 1999 Weia Conference, Univ. Genova, Genova, Italy, June, 1999.
- Chen, W-C., "On the weak form of Zipf's Law", *J. Appl. Probab.* **17**, 611-22, 1980.
- Costantini, D., and U. Garibaldi, "A finitary characterization of the Ewens sampling formula", mimeo, Department of Statistics, Univ. Bologna, 1999
- Ewens, W.J., (1990) "Population genetic theory—The past and the future" in *Mathematical and Statistical Developments of Evolutionary Theory*, edited by S. Lessard, Kluwer Academic Publishers, London
- Johnson, Kotz, and Balakrishnan (1997)
- Hoppe, F., (1987) "The sampling theory of neutral alleles and an urn model in population genetics," *J. Math. Biol.* **25**, 123-59
- Kelly, F., (1979) *Reversibility and Stochastic Networks*, J.Wiley, New York.
- Kingman, J.F.C., *Poisson Processes*, Oxford Clarendon Press, 1992
- (1980), *Mathematics of Genetic Diversity*, SIAM, Philadelphia
- , (1977) "The population structure associated with the Ewens sampling formula," *Theor. Popul Biol.* **11**, 274-83.
- Kirman, A., "Whom or what does the representative individual represent?" *J. Econ. Perspective*, **6**, 117-36, 1992 van Lint, J. H., and R. M. Wilson, (1992) *A Course in Combinatorics*, Cambridge Univ. Press, New York
- Lux, T., and D. Sornette, "On Rational Speculative Bubbles and Fat Tails", mimeo, 1999.
- Sachkov, V. N., *Probabilistic Methods in Combinatorial Analysis*, Cambridge Univ. Press, Cambridge, 1997.
- Sutton, J., "Gibrat's Legacy", *J. Econ. Literature*, **35**, 40-59, 1997.
- Watterson, G.A., "The stationary distribution of the infinitely-many neutral alleles diffusion model" *J. Appl. Probab.* **13**, 639-51, 1976.
- , and Guess, H. A., "Is the most frequent allele the oldest?" *Theor. Popul. Biol.* **11** 141-60, 1977.
- Zabell, S., "Predicting the Unpredictable", *Synthese* **90** 205-32, 1992.

Appendix: Dirichlet Distributions

A quick heuristic derivation of the Ewens sampling formula is given.

Suppose that there are n exchangeable agents and each chooses one of K possible behavioral rules (decisions, classes etc). A possible state vector to describe a configuration of agents clustering or distributed among K classes is to give the number of agents of each type, $n = (n_1, n_2, \dots, n_K)$ with $\sum_i n_i = n$.

If we know the fraction of agents who are of type i is p_i , $i = 1, 2, \dots, K$, and if each agent chooses independently, then this configuration is realized with probability

$$P(n_1, n_2, \dots, n_K | N) = \frac{n!}{n_1! n_2! \dots n_K!} \prod_{i=1}^K p_i^{n_i}.$$

This is a multinomial distribution.

Suppose next that the joint probability of the fractions is given by a Dirichlet distribution

$$\Pr(p_1, p_2, \dots, p_K) = \mathcal{D}(\epsilon, K) = \frac{\Gamma(K\epsilon)}{\Gamma(\epsilon)^K} (p_1 p_2 \dots p_K)^{\epsilon-1},$$

where we substitute p_K out by $p_K = 1 - p_1 - p_2 - \dots - p_{K-1}$. These fractions are located on the $K - 1$ dimensional simplex. The exponent ϵ is the same for all fractions by assumption, i.e., all types are exchangeable which implies that the joint probability must be symmetric in the fractions.

When fractions are integrated out from the multinomial distribution, we obtain what Chen calls Dirichlet-Multinomial distribution, Chen (1980),

$$DM(n_1, n_2, \dots, n_K | n, \epsilon) = \frac{n!}{\prod_i n_i!} \frac{\Gamma(K\epsilon)}{\Gamma(\epsilon)^K} \times I,$$

where

$$I = \int p_1^{n_1+\epsilon} \dots p_K^{n_K+\epsilon} dp_1 \dots dp_{K-1} = \frac{\prod_i \Gamma(n_i + \epsilon)}{\Gamma(N + K\epsilon)}.$$

Now rearrange n s in the non-increasing order and consider k largest ones, denoted by $n_{(1)} \geq n_{(2)} \geq \dots \geq n_{(k)}$. Their probability is given by

$$\Pr(n_{(1)}, \dots, n_{(k)} | N, k) = \frac{K!}{k!(K-k)!} DM(n_1, \dots, n_K | N, \epsilon).$$

Let $\sum_{i=1}^k n_{(i)} = n$.

In the alternative state vector choices we let a_i denote the number of n_j s which equal i ,

$$a_i = |\{j; n_j = i\}|,$$

where we note that

$$\sum_{i=1}^n a_i = k,$$

and

$$\sum_{i=1}^n ia_i = n.$$

Then

$$P(a_1, a_2, \dots, a_n | \sum a_i = k, \sum_i ia_i = n) = \frac{k!}{\prod_j a_j!} \times \Pr(n_{(1)}, \dots, n_{(k)} | N, k, \epsilon).$$

We now consider a limiting case of letting ϵ approach zero and K infinity while $K\epsilon \rightarrow \theta$ for some positive real number.

Write $\Gamma(\epsilon) = \Gamma(1+\epsilon)/\epsilon$, and note that $\prod_j \Gamma(n_j + \epsilon)/n_j!$ approaches $\prod_j j^{-a_j}$ because the product $n_1 \cdots n_k = \prod_j j^{a_j}$.

In the limit then, the probability is given by

$$P(a_1, a_2, \dots, a_n | \sum a_i, \sum_i ia_i = n) = \frac{n!}{\theta^{[n]}} \prod_i \frac{\theta^{a_i}}{i^{a_i} a_i!},$$

where $\theta^{[n]} := \Gamma(N + \theta)/\Gamma(\theta)$. This is the Ewens sampling formula.