No. 120 (81-21)

THE VALIDITY OF THE SIMON'S FIRM-SIZE MODEL AND ITS REVISION

by

Hajime Eto
Kyoko Makino*

August 1981

* The University of Tsukuba, Graduate School of Management Science and Public Policy Studies

THE VALIDITY OF THE SIMON'S FIRM-SIZE MODEL AND ITS REVISION

ABSTRACT

Herein is examined the theoretical validity of the wellknown and widely-accepted model by Simon and others which attempts to provide a theoretical foundation for the skew distribution of firm sizes. The scope and limit of its theoretical validity for a model of firm-size problem are specified. An alternative theoretical model is presented which is free from the aforementioned theoretical limit.

1. Introduction

The fact that the size of business firms is skewly distributed was theoretically explained by Simon and Bonini [7] for the first time on the base of Simon's theory [5]. Since then this theory has been developed [3, 6] and accepted and seems now established [3, 4]. Simon [5] first gave two resonably realistic assumptions of the proportionate growth and the constant birth rate. Second he claimed that the Yule distribution is derived from these assumptions. Finally he showed that the Yule distribution fits the real data of the American firms. This paper will show (i) his derivation of the Yule process assumes other additional conditions which do not hold in a general case of economics, (ii) therefore his theoretical model fails despite a good fit to the empirical data, and (iii) this paper will specify in which case his additional conditions hold. Finally (iv) this paper will provide another theoretical model which also yields the Yule distribution under the conditions which hold in a more general case of economics.

2. Assessment of the Conditions

The derivation of the Yule distribution by Simon [3, 5] is carried out with aid of linguistic interpretation of the stochastic process. Specifically the process is modelled in terms of word frequencies. Starting with the beginning of a book, k words are read so far. Then f(i, k) denotes the frequency of distinct words which have occurred exactly i times in the series of k words. In proving that the frequency f(i, k) is subject to the Yule distribution, not only the two assumptions of the proportionate growth (A1) and of the constant birth rate (A2) but also the two properties (P1 and P2) of the word frequencies are exploited.

(A1) $f(i, k+1)-f(i, k)= u(k)((i-1)f(i-1,k)- if(i, k)),$

$$i = 2, \ldots, k+1$$

(A2) $f(1, k+1)-f(1, k) = \alpha - u(k)f(1, k)$

where u(k) denotes the proportional growth rate and α denotes the constant birth rate ("birth" means the new words which did not occur in the series of k words). (A2) is the boundary condition of (A1).

In evaluating u(k), (P1) is exploited.

(P1) $\sum_{i=1}^{k} i\, f(i, k) = k \qquad \text{for all } k$

Evidently (P1) holds for the word frequency problem because the left side means the total number of words which have occurred in the series of k words.

In obtaining the steady-state distribution, in other words, in making f(i, k) independent of k, the other property (P2) is assumed.

(P2) $f(i, k+1)/f(i, k) = (k+1)/k$ for all k

(P2) means the proportionate growth of each frequency in k but is independent of (A1). Rather it is easy to see that (P2), as the balanced growth, depends on (P1).

Assuming (A1), (A2) and (P2) and utilizing (P1), he derivates the steady-state frequency f(i);

$$f(i) = (1 + \rho)B(i, \rho+1)f(1) \qquad \text{for all } i \qquad (1)$$

where B denotes the Beta function and ρ is defined

$$\rho = 1/(1 - \alpha) \qquad (2)$$

With aid of the empirical fact;

$$f(1) \sim 1/(2 - \alpha) \qquad (3)$$

the equation (4) is rewritten;

$$f(i) \sim \rho B(i, \rho+1) \qquad \text{for all } i \qquad (4)$$

which is the density of the Yule distribution.

To be more detailed, his derivation is based on (P1) and (P2) as follows. In evaluating u(k) from (A1) and (A2), the following relation is obtained;

$$u(k) = (1 - \alpha)/\sum_{i=1}^{K} i\, f(i, k) \qquad \text{for all } k \qquad (5)$$

Substituting (P1) in (5), u(k) is obtained;

$$u(k) = (1 - \alpha)/k \qquad \text{for all } k \qquad (6)$$

Also in obtaining the steady-state frequency f(i) independent of k, (P2) implies the following relation;

$$f(i, k)/f(i-1, k) = f(i, k+1)/f(i-1, k+1) \equiv v(i) \quad (7)$$

$$i = 2, \ldots, k+1$$

Owing to (A1), (P2) and (6), this yields;

$$v(i) = ((1 - \alpha)(i - 1))/(1 + i(1 - \alpha)) \quad (8)$$

Hence his derivation will loose its validity if (P1) and (P2) are invalid. Since they are consistent with each other and also with (A1) and (A2), they have no mathematical problem. As was discussed above, (P1) is valid in the word frequency problem. (P2) may be natural in this problem. His derivation follows only from (A1), (A2), (P1) and (P2) plus the widely accepted axiom of the probability theory that $1 - \alpha$ denotes the probability of the event compliment of the birth. As (A1) and (A2) are mathematically and linguistically acceptable, his derivation of the Yule distribution is mathematically valid and its application to the word frequency problem is valid.

In applying the Yule distribution to the firm-size problem and in claiming (A1) and (A2) as its theoretical model by which the application of the Yule distribution is theoretically justified [3, 4, 7], the reasonableness of the theoretical assumptions was not carefully examined. Let the validity of (A1), (A2), (P1) and (P2) in the firm-size problem be examined here.

(A1) and (A2) are acceptable as a discretized model when k is considered to denote the time. But (A1) holds for $i = 1, \ldots, \infty$. Now (P1) is meaningless because i is an economic term and k is a physical term. Instead of (P1), now we have (Q1);

$$\text{(Q1)} \quad \sum_{i=1}^{\infty} i\, f(i, k) = G(k) \qquad \text{for all } k$$

where G(k) denotes the sum of sizes (e.g., asset or sales of firms at k. Accordingly (P2) must change its form to retain the interpretation of the balanced growth of each

$$\text{(Q2)} \quad f(i, k+1)/f(i, k) = G(k+1)/G(k) \qquad \text{for all } k$$

After the same argument as Simon and Bonini [7], (9), (10) and (11) are obtained in place of (6), (2) and (8) respectively.

$$u(k) = (1 - \alpha)/G(k) \qquad \text{for all } k \qquad (9)$$

$$\rho(k) = (G(k+1) - G(k))/(1 - \alpha) \qquad \text{for all } k \qquad (10)$$

$$v(i, k) = ((1-\alpha)(i-1))/((G(k+1) - G(k)) + i(1-\alpha)) \qquad \text{for all } i \qquad (11)$$

Then (1), (5) and (7) can retain the same form meanwhile (3) must be replaced by (12) in order that (4) holds.

$$f(1, k) = (G(k+1) - G(k))/(1-\alpha + (G(k+1) - G(k)) \qquad \text{for all } k \qquad (12)$$

In order that $\rho(k)$, $v(i, k)$ and $f(1, k)$ are independent of k, the following function equation must hold;

$$G(k+1) - G(k) = g \qquad \text{for all } k \qquad (13)$$

where g denotes an unspecified constant. The unique solution to this function equation is the linearity of G(k) in k, that is;

$$G(k) = gk + \text{constant} \qquad \text{for all } k \qquad (14)$$

Proposition 1. The Yule distribution is theoretically valid for the firm-size problem under (A1) and (A2) if and only if the total economy grows (decays) linearly or stops the growth.

The same comment applies to his alternative formulation of the process [3, 5] which takes the same conditions as before. Indeed the equation (8) which assumes the linear growth plays again the essential role. But the real economy may grow exponentially.

A special kind of firm-size model could be found for which their derivation of the Yule distribution is valid. Let firms be grouped according to the location of their head offices. In other words let the firms whose head offices are located in the same city belong to the same group. Take k firms in the decreasing order of sizes and let f(i, k) denote the frequency of distinct groups in which exactly i firms have been listed up among the top k firms. These frequencies form a non-randomly selected sample distribution. As a concentration of firms in particular cities is a general trend, (A1) and (A2) may be reasonable this time. (P1) holds necessarily by definitions of i and k and (P2) is acceptable now. This may rather be called the city-size problem in terms of firm size. As the city size in terms of population is empirically known being subject to the Yule distribution [3, 5, 6], this new problem is not only theoretically but probably also empirically represented by the above model.

3. An Alternative Theoretical Model and the Associated Derivation of the Yule Distribution

The linear or zero growth assumption is not generally acceptable in the firm-size problem. Hence another theoretical model will be proposed here from which the Yule distribution is derived. As the proportionate growth now (A3) is assumed in place of (A1) and (A2) for $\gamma > 0$.

(A3) $\quad \mathrm{Prob}(i \rightarrow i+1,\ \Delta t) \sim \gamma i \Delta t$

where the left side denotes the probability that a firm of size i increases its size in an infinitesimal time. It is natural in the continuous time and discrete size model to assume the growth of size by the unit in an infinitesimal time. It is also natural in the age of inflation and of big technology to neglect the decrease of asset because it needs to increase its scale for technological survival and accordingly its monetary value at least nominally. Even as to the sales amount, its nominal decrease may be negligible.

Solving (A3) yields the "probability" $p(i|t)$ that the firm size grows from i to i+1 in time t under the initial condition of the size being i. [2, 8]

$$p(i|t) = \exp(-\gamma t)(1 - \exp(-\gamma t))^{i-1} \qquad (15)$$

which is not exactly the probability in that its integration over the time is less than the unity.

Theorem 1.

$$\int_0^\infty p(i|t)dt = (\gamma i)^{-1}$$

Proof. Apply the integration by change of variable to the right side of (15). Q.E.D.

Let $p^*(i|t)$ be defined as follows.

$$p^*(i|t) = \gamma i p(i|t)$$

Corollary 1. $p^*(i|t)$ possesses the property of the probability.

Actually $p(i|t)$ or $p^*(i|t)$ yields a distribution which approaches asymptotically to the log-normal distribution. This is consistent to the fact that a slightly different model of the proportionate growth yields the log-normal distribution [1].

Second it is assumed that a firm grows with a certain time of growth whose duration time d is subject to the exponential distribution for $\delta > 0$.

(A4) $d(t) = \delta \exp(-\delta t)$

The exponential distribution assumption is chosen because the duration or life time is often assumed so in bio-science, reliability engineering and communication engineering.

The assumption that a firm gows in a certain time duration and then suddenly stops growth is an extremitized simplification of the fact that a firm grows with a cyclic pulse. Since our concern is not with the state at each time but with the steady state long after the start time, the important thing is the total length of growth time. Clearly it is irrelevant whether the growth time is periodical or once at all.

Now the probability that a firm is at size level i for sufficiently large t is given by p(i) to follow;

$$p(i) = \int_0^\infty \exp(-\gamma t)(1 - \exp(-\gamma t))^{i-1}\,\delta\exp(-\delta t)dt$$
$$= \delta B(i, 1+\delta/\gamma)/\gamma = \rho B(i, \rho+1) \quad \text{for all } i \qquad (16)$$

with $\rho = \delta/\gamma$. Here the Yule distribution is derived from (A3) and (A4) with no additional condition.

Proposition 2. The Yule distribution is theoretically valid for the firm-size problem under (A3) and (A4) without the linear growth condition.

As was shown by Simon [3, 5], the upper tail of the Yule distribution is longer when ρ is smaller. In our new definition of ρ, the upper tail is longer when $1/\delta$ is relatively larger or γ is relatively larger. Recalling the wellknown fact of the exponential distribution that $1/\delta$ is the mean of d(t) in (A4), the following statement is now valid.

Proposition 3. The firm-size distribution is skewer when the duration of the growth time is longer relative to the growth rate.

4. Condlusion

The seemingly established model for explaing theoretically the firm size being subject to the Yule distribution was herein reexamined. This theoretical model, though mathematically correct, was found to be based on the implicit assumption which requires the linear growth of the total economy. An alternative model was presented which also explains the Yule distribution of the firm size. The latter model was shown free from the linear growth assumption and its new assumptions was justified.

5. References

1. Aitchison, J. and J. A. C. Brown: The Log-Normal Distribution; Cambridge Univ. Press, Cambridge, UK, 1957
2. Harris, T. E: The Theory of Branching Process; Springer, Berlin, 1972
3. Ijiri, Y. and H. A. Simon: Skew Distributions and the Sizes of Business Firms; North-Holland Pub. Co., Amsterdam, 1977
4. Nakayama, I. (ed): Encyclopedia of Modern Statistics (in Japanese); Toyo Keizai, Tokyo, 1962
5. Simon, H. A.: On a Class of Skew Distribution Functions; Biometrika 52, 425-440, 1955
6. Simon, H. A.: On Judging the Plausibility of Theories (in B. v. Rootselaar and J. F. Staal(ed): Logic, Methodology and Philosophy of Sciences; vol. 3); North Holland Pub. Co., Amsterdam, 1968
7. Simon, H. A. and C. P. Bonini: The Size Distribution of Business Firms; American Economic Review, 48, 607-617, 1958
8. Yabronsky, A. L.: On Fundamental Regularities of the Distribution Scientific Productivity; Scientometrics, 2, 3-34, 1980