Modeling Growth Stocks via Size Distribution[*]

S. C. Kou and S. G. Kou [†]

This version November 2001

[†]S. C. Kou, Department of Statistics, Science Center 603, Harvard University, Cambridge, MA 02138, `kou@stat.harvard.edu` S. G. Kou, Department of IEOR, 312 Mudd Building, Columbia University, New York, NY 10027 `kou@ieor.columbia.edu`

Modeling Growth Stocks via Size Distribution

This version November 2001

Abstract

The inability to predict the earnings of growth stocks, such as biotechnology and internet stocks, leads to the high volatility of share prices and difficulty in applying the traditional valuation methods. This paper attempts to demonstrate that the high volatility of share prices can nevertheless be used in building a model that leads to a particular size distribution, which can then be applied to price a growth stock relative to its peers. The model focuses on both *transient* and *steady* state behavior of the market capitalization of the stock, which in turn is modeled as a birth-death process. In addition, the model gives an explanation to an empirical observation that the market capitalization of internet stocks tends to be a power function of their relative ranks.

Issuing stocks is arguably the most important way for growth companies to finance their projects, and in turn helps transfer new ideas into products and services for society. Although the components of growth stocks may change over time (perhaps consisting of railroad and utility stocks in the early 1900's, and biotechnology and internet stocks in 2001), studying the general properties of growth stocks is essential to understand financial markets and economic growth.

However, uncertainty is manifest for growth stocks. For example, (a) growth stocks tend to have low or even negative earnings; (b) the volatility of growth stocks is high (both their daily appreciation and depreciation rates are high); (c) it is difficult to predict the upward and downward trends. Consequently, it poses a great challenge to derive a meaningful mathematical model within the classical valuation framework, such as the net present value method.

Since it appears that as far as growth stocks are concerned, we are only sure about their uncertainty, one may wonder whether there is much more to say about them. The current paper attempts to illustrate that a mathematical model for growth stocks can, nevertheless, be built, mainly by utilizing the high volatility of their share prices.

One motivation of the current study comes from a report on internet stocks in the Wall Street Journal[1] (Dec 27, 1999): People[2] at Credit Suisse First Boston observed that "there is literally a mathematical relationship between the ranking of the (internet) stock and its capitalization". More precisely, it is suggested the emergence of an almost linear downward

---

[1] We are grateful to Professor Morris Cohen at the Wharton School, University of Pennsylvania, to point out the article in the Wall Street Journal during a talk.

[2] This observation is summarized later in a research report by Mauboussin and Schay (2000).

pattern when the market capitalizations of internet stocks are plotted against their associated ranks on a log-log scale, with rank one being the largest market capitalization. Even more interestingly, the same article also reported that this phenomenon does not seem to hold for non-growth stocks[3]. The report challenges people to investigate whether such a phenomenon happens simply by chance or if there is a certain mechanism behind it.

The model proposed in the current paper provides an explanation of this phenomenon. Roughly speaking, the result suggests that if the market capitalization of the stocks is modeled as a birth-death process, then for the stocks with high volatility (such as biotechnology and internet stocks) an almost linear curve appears, on the log-log scale, for the size distribution[4] of the market capitalization. Meanwhile for non-growth stocks the model implies that such a phenomenon should not be expected, primarily because of the slow convergence of the birth-death process to its steady state distribution due to a low volatility. Furthermore, the model also suggests a way to price growth stocks (not just internet stocks) relatively to their peers.

Studying size distributions in various social problems has a long history, dating back at least to Pareto (1896), Yule (1924, 1944), Gibrat (1931), and Zipf (1949). In the 1950s, economists began to use various processes, including birth-death processes, to model size distributions in economics, including the sizes of business firms (see, for example, Simon, 1955, Ijiri and Simon, 1977, Lucas, 1978, Steindl, 1965, 1968, Simon and Bonini, 1958, Axtell, 2001), income distribution (see, for example, Rutherford, 1955, Mandelbrot, 1960, Shorrocks, 1975, Feenberg and Poterba, 1993), and city size distribution (see, for example, Glaeser, Scheinkman, and Shleifer, 1995, Krugman, 1996a, 1996b, Gabaix, 1999). However, most of the theory developed so far focuses on the steady-state size distribution and pays no special attention to the *transient* behavior of size distribution.

The contribution of the current paper is two fold.

(1) From a theoretical point of view, we give a detailed analysis (see Section III) of the transient behavior of size distribution, which is not well addressed in the size distribution literature. The *analysis of the transient behavior is crucial* to our analysis (see Section IV) as it explains why the size distribution theory can be applied to growth stocks but not to non-growth stocks.

---

[3]Based on the data of a single day, Mauboussin and Schay (2000) later also stated that saving and loan stocks may show a "strong power law characteristics" as well. However, our analysis in Section VI, based on the data of three years (from 1998 to 2000), does not seem to confirm any consistent patterns of the power law for saving and loan stocks.

[4]The size distribution is the distribution of some values of interest in terms of their relative ranks within a group.

(2) From an applied point of view, we point out that the theory of size distribution may have an interesting application in pricing growth stocks (see Section V), which is difficult for traditional methods, such as the net present value approach.

The current study also differs from the literature of using birth-death processes to model city size distribution in two aspects. First, in the city size distribution the exponent of the power law (i.e. the slope of regressing log city size on log city rank) is very close to -1 (see, for example, Krugman, 1996b, Gabaix, 1999). But here, the exponent (i.e. the slope of regressing log-market-capitalization on log-rank[5]) seems to be much smaller than -1, as will be seen in Section VI. Secondly, as pointed out in Gabaix (1999) and Krugman (1996a, pp. 96-97), it could take a birth-death process too much time to converge to the steady state distribution (which is the power law), if the volatility of city growth rates is not large; this, consequently, posts a serious problem for using birth-death processes to model city size distribution. However, in our case the volatility of growth stocks tends to be much higher than that of non-growth stocks, and that of city growth rates. Therefore, the growth stocks tend to converge to the steady state much faster, resulting in a clear linear pattern of the size distribution (as shown in Section VI). This also underlines the importance of studying the transient behavior of size distribution.

The paper is organized as follows. Section I proposes the basic model, while Sections II and III analyze both the *transient* and *steady* state properties of the model. The model is then applied in Section IV to derive the size distribution of growth stocks, and to explain why the method can be used for growth stocks but not for non-growth stocks. Implication for relative pricing of growth stocks is provided in Section V. Numerical illustrations are presented in Section VI. The advantage and disadvantage of the model are discussed in the last section. All the proofs are deferred to the appendices.

I The Model

Consider at time t a growth stock with total market capitalization $X(t)$, taking values in non-negative integers $X(t) = i$, $i = 0, 1, 2, \ldots$ (the market capitalization is defined as the product of the total outstanding shares and the market price of the stock). The unit of $X(t)$ could be, for example, millions or billions of dollars.

---

[5]In the city size distribution literature, people typically regress log rank on log city size, namely treating log rank as the regressor. Here we treat log rank as the regressee, i.e. regress log market capitalization on log rank, because the market capitalizations are of more interest than the ranks.

Model. The market capitalization $X(t)$ of the growth stock follows a birth-death process: given $X(t)$ being in state i, the instantaneous changes are as follows:

$$i \to i+1, \quad \text{with rate } i\lambda + g, \quad i \geq 0,$$

$$i \to i-1, \quad \text{with rate } i\mu + h, \quad i \geq 1,$$

where the parameters

$$\lambda, \mu > 0, \quad g > 0, \quad h \geq 0, \quad \lambda < \mu.$$

In other words, $X(t)$ follows a birth-death process with the infinitesimal generator given by the infinite matrix

$$\begin{pmatrix} -g & g & 0 & 0 & \cdots \\ \mu+h & -\lambda-\mu-g & \lambda+g & 0 & \cdots \\ 0 & 2\mu+h & -2\lambda-2\mu-g & 2\lambda+g & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

In the standard notation, $X(t)$ is a birth-death process[6] with the birth rate λ_i and the death rate μ_i satisfying

$$\lambda_i = i\lambda + g, \ \mu_i = i\mu + h, \ i \geq 1, \ \lambda_0 = g, \ \mu_0 = 0. \tag{1}$$

The two parameters λ and μ represent the instantaneous appreciation and depreciation rates of $X(t)$ due to market fluctuation; the model assumes that they influence the market capitalization proportionally to the current value. The requirement $\lambda < \mu$ is postulated here to ensure that the birth-death process has a steady state distribution. The existence of a steady state distribution is necessary for the discussion of the size distribution, and is a standard assumption in the literature; see, for example, the book by Ijiri and Simon (1977).

In general, because of the difficulty of predicting the instantaneous upward and downward price movements, for both growth stocks and non-growth stocks λ and μ must be quite close, $\lambda/\mu \approx 1$. In addition, for growth stocks, both λ and μ must be large, due to the high volatility. These observations will become assumptions (A3) and (A4) in Section IV.B.

The parameter $g > 0$ models the rate of increase in $X(t)$ due to non-market factors,[7] such as, the effect of additional shares being issued through public offerings, or the effect of warranties on the stock being exercised (resulting in new shares being issued). For growth stocks, g is at least as large as that for non-growth stocks. The parameter h attempts to capture the rate of

---

[6]The state 0 only means that the size is below a certain minimal level. It does not imply, for example, the company goes bankrupt.

[7]The model is not very interesting if $g = 0$, as the steady state distribution degenerates to a single point mass at zero.

decrease in $X(t)$ due to non-market factors, such as the effect of dividend payments. For most growth stocks $h \approx 0$, as no dividends are paid.
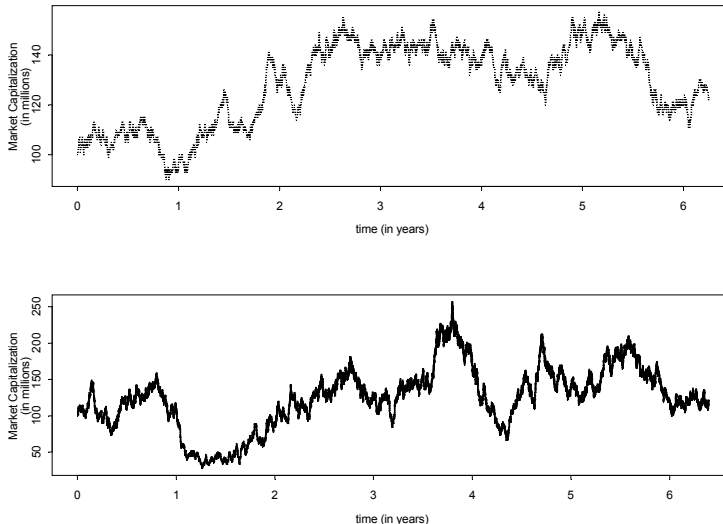

Figure 1: Two sample paths of the birth-death process. In the upper panel $X(0) = 100$, $\lambda = 2.49$, $\mu = 2.51$, $g = 1$, $h = 0$; in the lower panel, $X(0) = 100$, $\lambda = 24.9$, $\mu = 25.1$, $g = 10$, $h = 0$.

Figure 1 provides an illustration of the model by showing the sample paths of two realizations of the birth-death process (1) for about 6.5 years. In the upper panel, the instantaneous jump rates, λ and μ, are small, while in the lower panel λ and μ are large. The sample paths suggest two points: (1) for reasonably large λ and μ, the jumps of the birth-death processes is almost unnoticeable, and the overall sample paths fit in well with our intuition of market fluctuation; (2) although $\lambda < \mu$, the sample paths may still have some strong upside movements if λ is close to μ; for example, in the lower panel, the market capitalization increases from about 20 to about 250 (more than 12 times) within a short period (about 2.5 years).

The model proposed here is a variation and a generalization of the models proposed in Simon (1955) and Shorrocks (1975) to study business and income sizes, etc. The key difference here is that we provide a detailed analysis of both *transient* and *steady* states, not just the steady state analysis. The transient analysis not only presents some mathematical challenges (see Section III), but also is essential to understand why the theory of size distributions is useful

5

for growth stocks but not for non-growth stocks (see Section IV).

II   General Properties of the Model

II.A   Properties of the Steady State Distribution

The steady state measure is given by

$$\pi_0 = 1, \ \pi_n := \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n}, \ n = 1, 2, ...$$

Normalizing $\{\pi_n\}$ provides the steady state distribution of the birth-death process:

$$\lim_{t \to \infty} P(X(t) = n) = \pi_n/S, \ \ S := \sum_{n=0}^{\infty} \pi_n;$$

(see Proposition 1 below for the finiteness of S under the setting of (1)). In our case,

$$\pi_n = \left(\frac{\lambda}{\mu}\right)^n \frac{(g/\lambda)(1 + g/\lambda)(2 + g/\lambda) \cdots ((n-1) + g/\lambda)}{(1 + h/\mu)(2 + h/\mu) \cdots (n + h/\mu)}, \quad n \geq 1.$$

Using the gamma function, it can be succinctly expressed as

$$\pi_n = \frac{\Gamma(1 + h/\mu)}{\Gamma(g/\lambda)} \left(\frac{\lambda}{\mu}\right)^n \frac{\Gamma(n + g/\lambda)}{\Gamma(n + 1 + h/\mu)}, \quad n \geq 0. \tag{2}$$

Proposition 1. (Steady-State Properties).

(1) *The birth-death process (1) is positive recurrent; i.e. it will visit every state $\{0, 1, 2, \ldots\}$ with probability one, and the expected visiting time of any state is finite.*

(2) *As $n \to \infty$,*

$$\pi_n \cong \frac{\Gamma(1 + h/\mu)}{\Gamma(g/\lambda)} \left(\frac{\lambda}{\mu}\right)^n n^{g/\lambda - h/\mu - 1}, \tag{3}$$

where throughout this paper, $a \cong b$ means $\lim a/b = 1$. This asymptotic order, in particular, implies that $S = \sum_{n=0}^{\infty} \pi_n$ is finite.

(3) *The moment generating function of the steady state distribution is given by*

$$M(\theta) := \sum_{n=0}^{\infty} e^{\theta n} \pi_n/S = F(\frac{g}{\lambda}, 1; 1 + \frac{h}{\mu}; \frac{\lambda}{\mu} e^{\theta})/F(\frac{g}{\lambda}, 1; 1 + \frac{h}{\mu}; \frac{\lambda}{\mu}), \tag{4}$$

where $F(a, b; c; z)$ is the hypergeometric function (see page 556 of Abramowitz and Stegun, 1972):

$$F(a, b; c; z) := \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^{\infty} \frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)} \frac{z^n}{n!}.$$

6

In particular, the mean and the second moment of the steady state distribution are

$$m_1 \quad := \quad M'(0) = \frac{1}{S}\frac{g}{\mu+h}F(1+\frac{g}{\lambda},2;2+\frac{h}{\mu};\frac{\lambda}{\mu})$$

$$m_2 \quad := \quad M''(0) = \frac{1}{S}\frac{g}{\mu+h}\left\{F(1+\frac{g}{\lambda},2;2+\frac{h}{\mu};\frac{\lambda}{\mu})+2\frac{\lambda+g}{2\mu+h}F(2+\frac{g}{\lambda},3;3+\frac{h}{\mu};\frac{\lambda}{\mu})\right\}.$$

(4) Let the tail probability of the steady state distribution be

$$F(n) := \lim_{t\to\infty} P(X(t) \geq n) = \sum_{k=n}^{\infty} \pi_k/S.$$

Then, as $n \to \infty$,

$$F(n) \cong \frac{1}{S}\frac{\Gamma(1+h/\mu)}{\Gamma(g/\lambda)}\left(1-\frac{\lambda}{\mu}\right)^{-1}\left(\frac{\lambda}{\mu}\right)^{n}n^{g/\lambda-h/\mu-1}. \tag{5}$$

Proof. See Appendix A. □

Note that, instead of the original parameters, only the three ratios, λ/μ, h/μ, and g/λ, determine the steady state distribution. Thus the steady state properties only reflect the relative magnitude of the parameters λ, μ, g and h, rather than the absolute magnitude. (This contrasts with the realizations of the birth-death process, such as Figure 1, in which the dynamic behavior of sample path does depend on the absolute magnitude of λ, μ, g and h.)

II.B   Transient Mean and Variance

Proposition 1 only provides the steady state properties of the market capitalization in our model. However, these properties are only relevant *if the convergence from the transient states to the steady states is fast enough*, i.e. if the convergence can be observed in a timely fashion[8]. This is a point stressed in Gabaix (1999) and Krugman (1996a, pp. 96-97).

There are several ways to judge the convergence speed. In this subsection we shall focus on the mean and variance of the transient distribution, which can lead to a measure of the convergence rate; see Section IV.A. A more accurate measure (which is of course more difficult to study) is the convergence rate for the transition probabilities, which attempts to capture the convergence rate for the whole distribution rather than just the first two moments; this will be analyzed in the next section.

---

[8]Another possibility is that a birth-death process has been run for a long time; i.e. the stock has been traded in market for a long period. However, the parameters, λ, μ, g and h may have changed during the period, thus altering the steady state distribution.

Denote the transition probability at time t to be $p_{i,j}(t) := P(X(t) = j | X(0) = i)$, the transient expectation at time t to be $m_1(t) := EX(t) = \sum_{j=0}^{\infty} j p_{i,j}(t)$, and the second moment to be $m_2(t) := EX^2(t) = \sum_{j=0}^{\infty} j^2 p_{i,j}(t)$.

Proposition 2. (Transient Mean and Variance) *Suppose the birth-death process starts from $X(0) = i$. The first moment $m_1(t)$ at time t satisfies the following differential equation:*

$$m_1'(t) = (\lambda - \mu) m_1(t) + g + h(1 - p_{i,0}(t)), \tag{6}$$

whose solution is given by

$$m_1(t) = i e^{(\lambda - \mu)t} + \frac{g}{\mu - \lambda}[1 - e^{(\lambda - \mu)t}] + h \int_0^t e^{(\mu - \lambda)(s-t)}(1 - p_{i,0}(s))ds.$$

The second moment $m_2(t)$ satisfies

$$m_2'(t) = 2(\lambda - \mu) m_2(t) + (\lambda + \mu + 2g - 2h) m_1(t) + g + h(1 - p_{i,0}(t)), \tag{7}$$

with the solution given by

$$m_2(t) = i^2 e^{2(\lambda - \mu)t} + \frac{g}{2(\mu - \lambda)}[1 - e^{2(\lambda - \mu)t}] + h \int_0^t e^{2(\mu - \lambda)(s-t)}(1 - p_{i,0}(s))ds +$$
$$+ (\lambda + \mu + 2g - 2h) \int_0^t e^{2(\mu - \lambda)(s-t)} m_1(s)ds.$$

Proof. See Appendix[9] A. $\square$

---

[9]The differential equations (6) and (7) also provide another way to calculate the mean and variance of the steady state distribution. Letting $t \to \infty$ in (6) and (7), and using the fact that $\lim_{t \to \infty} p_{i,0}(t) = \pi_0/S = 1/S$, we know that the mean and second moment of the steady state distribution satisfy

$$0 = (\lambda - \mu) m_1 + g + h(1 - 1/S),$$
$$0 = 2(\lambda - \mu) m_2 + (\lambda + \mu + 2g - 2h) m_1 + g + h(1 - 1/S),$$

which gives the mean, the second moment, and variance of the steady state distribution as

$$m_1 = \frac{g}{\mu - \lambda} + \frac{h}{\mu - \lambda}(1 - \frac{1}{S}), \quad m_2 = \frac{\mu + g - h}{(\mu - \lambda)^2}[g + h(1 - \frac{1}{S})],$$
$$\text{var} = m_2 - m_1^2 = \frac{1}{(\mu - \lambda)^2}[g + h(1 - \frac{1}{S})][\mu - h(2 - \frac{1}{S})].$$

8

III   The Transient Behavior of the Model

We shall see in the next section that the model implies that in the steady state the size distribution of the birth-death process follows an almost linear curve (on a log-log scale), as empirically observed. A natural question is then: why this phenomenon is observed for growth stocks, but not for non-growth stocks. Basically, the answer hinges on the fact that due to the high volatility of the growth stocks (in our model meaning that λ and μ are big) the birth-death process converges very fast to the steady state distribution, whereas for the non-growth stocks the convergence is so slow (because that λ and μ are not big) that essentially the steady state size distribution cannot be observed in practice.

This section provides necessary results of the transient behavior of the model, especially the convergence speed to the steady state distribution, to be used in the next section. As we mentioned, most of the literature on the size distribution focuses on the steady state properties, and, except for some numerical results (see for example Shorrocks[10], 1975), the theoretical properties of the transient behavior are hardly addressed in the literature. In this sense, this section constitutes the main technical contribution of the current paper to the size distribution literature.

The speed of convergence of a birth death process can be measured by the decay parameter (see Kijima, 1997), which is defined by

$$\gamma := \sup\{\alpha \geq 0 : p_{i,j}(t) - (\pi_j/S) = O(e^{-\alpha t}) \text{ for all } i, j \geq 1\},$$

where recall $p_{i,j}(t)$ is the transition probability at time t and π_j/S is the steady state probability.

Notice that the decay parameter γ affects the convergence *in an exponential way*. In other words, a small difference in γ can have a remarkable effect on the speed of convergence, which in turn suggests that the steady state analysis of the size distribution in our model based on the birth-death process is only relevant when the decay parameter is large.

Theorem 1. (The Decay Parameter) *For the birth-death process in the model, if* $h = 0$ *then the decay parameter*

$$\gamma = \mu - \lambda;$$

---

[10]Shorrocks (1975) showed, through numerical calculation, that if the convergence rate is not large enough, it may take 15 to 181 years for some birth-death processes to reach the steady state.

otherwise, if $h > 0$ then

$$\mu - \lambda \le \gamma < \mu - \lambda + h[1 - \min(\frac{\lambda}{\mu}, \frac{\lambda + g}{\mu + h})].$$

The derivation of this theorem is the main technical contribution of the current paper to the study of size distribution, and is deferred to Appendix B.

IV The Size Distribution for the Growth Stocks

In this section we shall apply the results obtained in the last two sections, of both steady state and the transient behavior of the model, to study the size distribution of growth stocks. Since for most growth stocks, there is no dividend payment, we shall assume from this section on that

$$h = 0. \tag{8}$$

IV.A Basic Transient and Steady State Properties for $h = 0$

Under the assumption (8), Proposition 1 implies that the steady state measure becomes

$$\pi_n = \frac{1}{\Gamma(g/\lambda)} \left(\frac{\lambda}{\mu}\right)^n \frac{\Gamma(n + g/\lambda)}{n!}, \quad n \ge 0,$$

with

$$F(n) = \sum_{k=n}^{\infty} \pi_k / S = \pi_n F(n + g/\lambda, 1; n + 1; \lambda/\mu)/S, \quad n \ge 0.$$

In addition, $S = \sum_{k=0}^{\infty} \pi_k = F(g/\lambda, 1; 1; \lambda/\mu) = \left(1 - \frac{\lambda}{\mu}\right)^{-g/\lambda}$, thanks to the following property of hypergeometric function: $F(a, b; b; z) = (1 - z)^{-a}$. This, together with (5), yields

$$F(n) = \lim_{t \to \infty} P(X(t) \ge n) \cong \frac{1}{\Gamma(g/\lambda)} \left(1 - \frac{\lambda}{\mu}\right)^{g/\lambda - 1} \left(\frac{\lambda}{\mu}\right)^n n^{g/\lambda - 1}. \tag{9}$$

By (4), the moment generating function of the steady state distribution, under $h = 0$, is

$$M(\theta) = \left(\frac{1 - \frac{\lambda}{\mu} e^{\theta}}{1 - \frac{\lambda}{\mu}}\right)^{-g/\lambda} = \left(\frac{\mu - \lambda e^{\theta}}{\mu - \lambda}\right)^{-g/\lambda}.$$

Consequently, the mean and the second moment of the steady state distribution are

$$m_1 = M'(0) = \frac{g}{\mu - \lambda}, \quad m_2 = M''(0) = \frac{g(\mu + g)}{(\mu - \lambda)^2},$$

10

and the variance is given by

$$\text{var} = m_2 - m_1^2 = \frac{\mu g}{(\mu - \lambda)^2}.$$

For the properties of the transient behavior, first note that, by Theorem 1, the decay parameter, which measures the speed of convergence to steady state in an exponential way, is given by

$$\gamma = \mu - \lambda.$$

Secondly, by Proposition 2, the differential equations[11] of the mean and second moment for the transient states become

$$
\begin{aligned}
m_1'(t) &= (\lambda - \mu)m_1(t) + g, \quad m_1(0) = i, \\
m_2'(t) &= 2(\lambda - \mu)m_2(t) + (\lambda + \mu + 2g)m_1(t) + g, \quad m_2(0) = i^2,
\end{aligned}
$$

whose solutions are

$$
\begin{aligned}
m_1(t) &= ie^{(\lambda - \mu)t} + \frac{g}{\mu - \lambda}[1 - e^{(\lambda - \mu)t}], \\
m_2(t) &= i^2 e^{2(\lambda - \mu)t} + i\frac{\lambda + \mu + 2g}{\lambda - \mu}(e^{2(\lambda - \mu)t} - e^{(\lambda - \mu)t}) + \\
&\quad + \frac{g}{2(\mu - \lambda)}[1 - e^{2(\lambda - \mu)t}] + \frac{g(\lambda + \mu + 2g)}{2(\mu - \lambda)^2}(1 - e^{(\lambda - \mu)t})^2.
\end{aligned}
$$

The exponents in $m_1(t)$ and $m_2(t)$ are all related to $(\lambda - \mu)$, which also points out, from a different viewpoint, that $(\mu - \lambda)$ should affect the speed of convergence in an exponential way. In addition, it is easily seen that

$$\lim_{t \to \infty} m_1(t) = \frac{g}{\mu - \lambda} = m_1, \quad \lim_{t \to \infty} m_2(t) = \frac{g(\mu + g)}{(\mu - \lambda)^2} = m_2.$$

IV.B The Size Distribution

Consider M (here M is an unknown quantity) growth stocks governed by the same birth-death process as indicated in the model, among which the K largest stocks (in terms of their market capitalization) are included in a group to be studied. Suppose we rank the market capitalization from 1 to K and denote the resulting ranked values as $X_{(1)}$, $X_{(2)}$, ..., $X_{(K)}$, with $X_{(1)}$ being the largest, and $X_{(2)}$ the second largest etc. Then the empirical tail distribution $\tilde{F}(x)$ (the empirical version of F) evaluated at $X_{(i)}$ is simply $\tilde{F}(X_{(i)}) = i/M$, $i = 1, ..., K$. Now we make two assumptions.

---

[11]Karlin and McGregor (1958) derived a differential equation for $p_{i,j}(t)$, in the case of $h = 0$, and solved it by using orthogonal polynomials.

(A1): The birth-death process has reached the steady state.

(A2): For each stock included in the group, the market capitalization is large; in other words, even $X_{(K)}$ is large.

According to (9), in the steady state, for large capitalization n,

$$\log F(n) \cong \log \left\{ \frac{1}{\Gamma(g/\lambda)} \left(1 - \frac{\lambda}{\mu}\right)^{g/\lambda - 1} \right\} + n \log \left(\frac{\lambda}{\mu}\right) - (1 - g/\lambda) \log(n).$$

Therefore, empirically with $X_{(i)} = n$, we shall expect that

$$\log \tilde{F}(X_{(i)}) = \log(i/M) \approx \log \left\{ \frac{1}{\Gamma(g/\lambda)} \left(1 - \frac{\lambda}{\mu}\right)^{g/\lambda - 1} \right\} + X_{(i)} \log \left(\frac{\lambda}{\mu}\right) - (1 - g/\lambda) \log(X_{(i)}).$$

Rearranging the terms above yields

$$\log X_{(i)} \approx C - \frac{1}{1 - g/\lambda} \log i + \frac{1}{1 - g/\lambda} X_{(i)} \log \left(\frac{\lambda}{\mu}\right), \quad 1 \leq i \leq K, \tag{10}$$

where the constant term

$$C = \frac{1}{1 - g/\lambda} \log \left\{ \frac{1}{\Gamma(g/\lambda)} \left(1 - \frac{\lambda}{\mu}\right)^{g/\lambda - 1} \right\} + \frac{\log(M)}{1 - g/\lambda}. \tag{11}$$

Since M is unknown, C is essentially a free parameter. Equation (10) has several immediate implications, if we make two more assumptions.

(A3): For both growth and non-growth stocks, $\lambda/\mu \approx 1$.

(A4): For growth stocks, both λ and μ must be large.

Assumption (A3) is postulated because generally it is hard to predict instantaneous upward and downward price movements for both growth stocks and non-growth stocks; thus, λ and μ must be quite close. Assumption (A4) reflects the high volatility[12] of growth stocks. Note that (A4) implies that the decay parameter $\gamma = \mu - \lambda$ (which affects the convergence in an exponential way) may also be large, thus leading to a fast convergence to the steady state distribution and justifying (A1).

By (A3), $\lambda/\mu \approx 1$; so the last term in (10) must generally be small. Consequently, if one plots the logarithm of the market caps against the logarithm of the ranks, of "large-cap" [13] growth stocks that satisfy Assumption (A2) above, (10) suggests that it should be almost linear. This explains the first half of the empirical observation reported in the Wall Street Journal.

---

[12] Kerins, Smith, and Smith (2001) shows empirically that the volatility of internet stocks may be at least five times that of traditional stocks.

[13] Here the word "large-cap" is used in a loose sense, and should not be confused with similar words used in the exchanges. Here it means that the market capitalization is large enough so that the asymptotic result (9) holds.

In addition, the results imply that the same phenomenon of the size distribution should hold not only for large-cap internet stocks but also for other large-cap growth stocks, such as large-cap biotechnology stocks, with large λ and μ, and $\lambda/\mu \approx 1$.

IV.C   Why the Model Does Not Apply to Non-Growth Stocks

There are at least two reasons why the almost linear relationship between the logarithm of the market capitalization and the logarithm of the ranks does not appear for non-growth stocks. First, the birth-death process model may not be valid for non-growth stocks. Secondly, even if the model is valid for non-growth stocks, in order to empirically observe such a linear phenomenon as implied by (10), a few conditions must be satisfied, as (10) is based on the steady-state distribution:

(C1): In terms of time, the convergence from the transient states to the steady state must be fast enough. This in turn depends on the magnitude of the decay parameter γ; in other words, γ must be large.

(C2): In terms of market capitalization, X must be large enough, as required by the asymptotic results in (9) and (10).

(C3): To observe the almost linear relationship between the logarithm of the market capitalization and the logarithm of the ranks, λ/μ must be close to one in order to make the last term in (10) disappear.

For the large-cap (thus satisfying the condition (C2) above) growth stocks, by assumptions (A3) and (A4), both λ and μ are large, and $\lambda/\mu \approx 1$. If $\mu - \lambda$ is large, then the decay parameter γ is also large, thus resulting in a fast convergence to the steady state.

For non-growth stocks, the volatility parameters, which in our model are λ and μ, are generally not large. As a consequence, the decay parameter $\gamma = \mu - \lambda$ (which affects the convergence in an exponential way) cannot be large in general. In other words, although in the steady state plotting the logarithm of the market capitalization against the logarithm of the relative ranks may display a linear relationship, the linear relationship may not emerge at all within a reasonable amount of time, due to the slow convergence from the transient state to the steady state. Furthermore, if the convergence rate is slow, many factors can lead the process to depart from the original steady state, e.g. changing of λ and μ, etc.

13

V   Relative Pricing of Growth Stocks

The model can be used to price large-cap growth stocks relatively within a peer group with similar parameters λ, μ, and g (for example, it may not be sensible to group biotechnology stocks with internet stocks as their parameters may be quite different). In particular, (10) provides a link between the market capitalization of the stocks and their relative ranks within the group. However since it involves a nuisance parameter C, a better equation can be obtained by eliminating C first, as is typical in many standard statistical procedures. To do this, observe that when $i = 1$ we have

$$\log X_{(1)} \approx C - \frac{1}{1 - g/\lambda} \log 1 + \frac{1}{1 - g/\lambda} X_{(1)} \log\left(\frac{\lambda}{\mu}\right). \tag{12}$$

Taking the difference between (10) and (12) cancels out the nuisance constant C and gives

$$\log \frac{X_{(i)}}{X_{(1)}} \approx -\frac{1}{1 - g/\lambda} \log i + \frac{1}{1 - g/\lambda} \left(X_{(i)} - X_{(1)}\right) \log\left(\frac{\lambda}{\mu}\right), \quad 1 \le i \le K. \tag{13}$$

As a key formula in this section, equation (13) provides a way to price growth stocks relative to their peers. More specifically, one can precede with the following two steps: (a) obtain the parameters λ/μ and g/λ by running a regression according to (13). This can be done, for example, by choosing λ/μ and g/λ to minimize the squared errors for $\log(X_{(i)}/X_{(1)})$, subject to the constraints that $\lambda/\mu < 1$ and $g/\lambda > 0$. (b) Once these parameters are obtained, the theoretical market capitalization of the stock can be calculated according to equation (13), with the input being its rank. This, in turn, leads to a theoretical value of the shares price, after dividing the market capitalization by the total number of outstanding shares.

Since the last term in (13) is typically small (due to $\lambda/\mu \approx 1$), one can also use, as a quick approximation, a simplified version of (13):

$$\log \frac{X_{(i)}}{X_{(1)}} \approx -\frac{1}{1 - g/\lambda} \log i, \quad 1 \le i \le K, \tag{14}$$

with the constraint $g/\lambda > 0$.

Note that the model suggests that the exponent of the power distribution (i.e. the slope of the regression line), $-\frac{1}{1-g/\lambda}$, is less than -1, which will be confirmed by the data in Section VI. This is quite different from the city size distribution, in which the exponent is very close to -1; see Krugman (1996b), Gabaix (1999).

We want to point out that the total number of stocks, K, included in the peer group in (13) or (14) should be as large as possible, as long as it satisfies the requirements that, for example,

14

the stocks within the peer group must have similar characteristics (so that they are governed by the same parameters λ, μ, and g), and their market capitalizations should all be large enough (so that assumption (A2) is satisfied). Big K helps in two ways: (a) it makes the estimation of the parameters more accurate; (b) If K is small, then the result may be altered by omitting some stocks with large capitalization; for example, if there are 10 growth stocks and the third largest one is ignored in the regression, then the estimated parameters will be biased.

The regression using (13) or (14) is, however, robust against possible truncation errors, thanks to the fact that the relative ranks are used . For example, if there are totally 200 growth stocks and only the top 100 stocks with the large market capitalization are included in the estimation, then (10) will not alter.

Another good property worthy of mentioning is that (14) is scale-invariant. Indeed, if the unit of $X(t)$ changes by a factor of A (i.e. the new unit becomes A times the original unit), then g/λ in (14) would not change at all. Equation (13) is almost scale-invariant: if the unit of $X(t)$ scales up by a factor of A, then g/λ in (13) remains the same, while λ/μ becomes $(\lambda/\mu)^A$. However, since $\lambda/\mu \approx 1$, the difference between λ/μ and $(\lambda/\mu)^A$ is generally insignificant unless A is very large.

In certain sense, the pricing method via (13) or (14) reminds us of the relative pricing idea in the valuation of contingent claims, such as in the Black-Scholes model (Black-Scholes, 1973), in which, given the price of a stock, the price of an option of the stock is calculated relative to the value of the stock. Essentially, the model here provides a way for relative pricing, by evaluating the price of a growth stock relative to its peers within the group (the contribution of the peer group is to provide an estimate of λ/μ, g/λ, and the relative ranks). This echoes a principle underlying the Black-Scholes model that when absolute pricing is difficult relative pricing may be easier.

VI   Numerical Illustrations

To illustrate the results in the previous sections for biotechnology stocks, we plot in Figure 2 the logarithm of their market capitalization relative to the largest biotechnology stock versus the logarithm of their ranks. In other words, $\log(X_{(i)}/X_{(1)})$ are plotted against $\log i$. This can be viewed as choosing $X_{(1)}$ as the unit of measurement. The six panels shown, which involve 139 biotechnology stocks[14], reflect January 2, 1998 and every 150 trading days thereafter. In

---

[14]The 139 stocks include most of the stocks listed in the Nasdaq biotechnology index and the BTK biotechnology index. See Appendix C for a list of these 139 biotechnology stocks.

each panel, the total market capitalization of these 139 stocks are first computed by taking the product of the number of outstanding shares and the share price; then the stocks with a market capitalization not smaller than 0.5% of that of the largest stock are plotted. The relationship (10) requires large market capitalization, and here "large-cap" are ad hoc adopted as stocks having market capitalization at least as large as 0.5% of that of the largest stock. One advantage of categorizing "largeness" relatively is that it automatically takes into account that different groups of stocks could have different sizes (for example, even within growth stocks, internet stocks tend to be larger than biotechnology stocks).
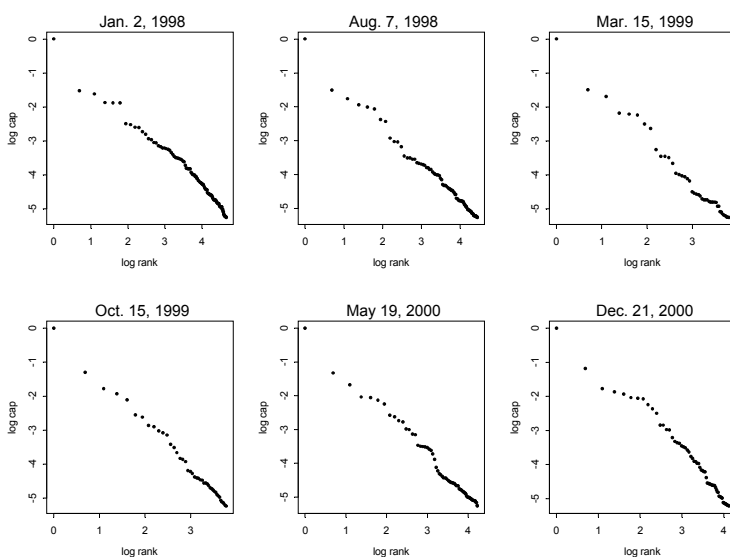

Figure 2: Size distribution of biotechnology stocks

It is worth noting that the six days shown in Figure 2 include the days when the biotechnology stocks were performing well, as well as the days when the biotechnology stocks were grounded heavily. Nevertheless, in all six plots there is clearly a linear trend, a pattern predicted by the model.

Contrastingly, in Figures 3 and 4, for the same six trading days, the logarithm of the market capitalization of the 20 Dow transportation[15] stocks and 88 saving and loan stocks[16] relative to

---

[15]The 20 Dow transportation stocks are listed in Appendix C, among which the smallest one has a market capitalization about 2% of that of the largest.

[16]The 88 saving and loan stocks are listed in Appendix C.

the largest one is plotted against the logarithm of their rank. The plot of Dow transportation stocks is far from linear. For the saving and loan stocks, although in some days there may be a linear pattern (e.g. Dec. 21, 2000), the pattern disappears in other days and is not consistent at all. The non-linear pattern is again expected from the model, since the convergence of non-growth stocks to the steady state distribution (governed by the decay parameter) is generally too slow to be observed in practice.
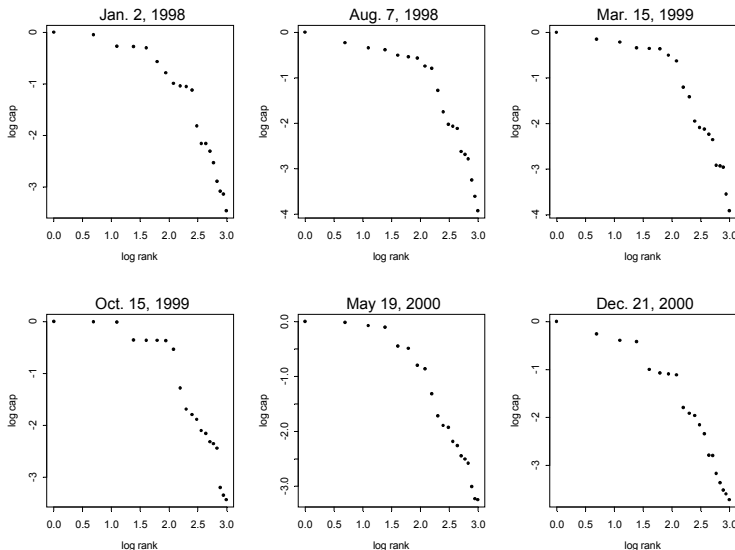

Figure 3: Plot for the 20 Dow transportation stocks

For the biotechnology stocks in Figure 2, the parameters g/λ and λ/μ are estimated by fitting the model (13) to the data. The estimates $\widehat{g/\lambda}$ and $\widehat{\lambda/\mu}$ can be simply obtained by minimizing the squared errors for $\log(X_{(i)}/X_{(1)})$:

$$(\widehat{g/\lambda}, \widehat{\lambda/\mu}) = \arg\min_{(\frac{g}{\lambda}, \frac{\lambda}{\mu})} \sum_{i=1}^{K} \left[\log \frac{X_{(i)}}{X_{(1)}} - \left\{ -\frac{1}{1-g/\lambda} \log i + \frac{1}{1-g/\lambda} \left(X_{(i)} - X_{(1)} \right) \log \left(\frac{\lambda}{\mu} \right) \right\} \right]^2,$$

(15)

For example, for the first panel in Figure 2 (January 2, 1998), g/λ and λ/μ are estimated to be 0.08 and $1 - 1.38 \times 10^{-9}$ respectively. For growth stocks, since g (which models, for example, new shares being issued and warranties being exercised) is at least as large as that of non-growth stocks, the above numbers, hence, suggest that both λ and μ must be large and that $\lambda/\mu \approx 1$, thus confirming our earlier assumptions (A3) and (A4). Table 1 reports the estimated $\widehat{g/\lambda}$ and
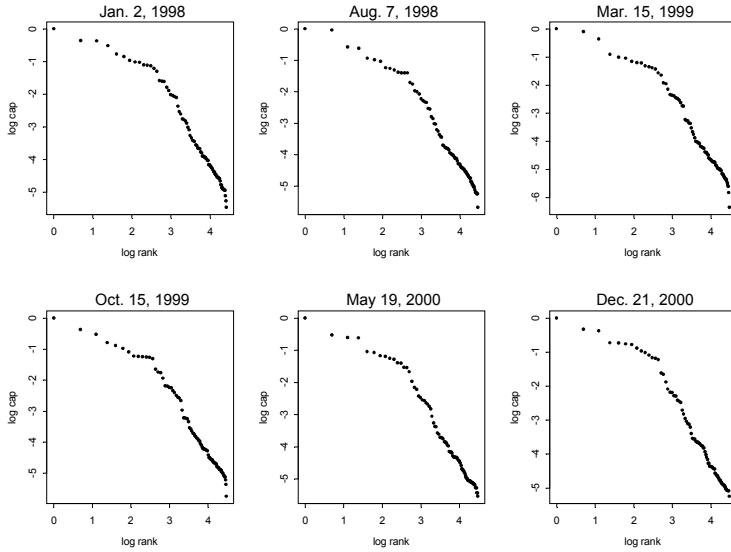
Figure 4: Plot for saving and loan Stocks

$\widehat{\lambda/\mu}$ as well as the R^2, which measures the goodness-of-fit[17], for all the six panels in Figure 2. Note that $\widehat{g/\lambda}$ are all small and that $\widehat{\lambda/\mu}$ are very close to 1.

Using the estimated values of g/λ and λ/μ, the dashed lines in Figures 5 show the relationship between the log-market capitalization and the log-rank, as suggested by the model. They agree well with the empirical observation. The R^2 being at least 97% directly supports the visual impression.

	$\widehat{g/\lambda}$	$1 - \widehat{\lambda/\mu}$	R^2
Jan 2, 98	0.080	1.38×10^{-9}	97.8%
Aug 7, 98	0.165	1.25×10^{-9}	98.2%
Mar 15, 99	0.295	1.06×10^{-9}	98.3%
Oct 15, 99	0.272	1.09×10^{-9}	99.2%
May 19, 00	0.197	1.20×10^{-9}	98.6%
Dec 21, 00	0.265	5.65×10^{-9}	97.5%

Table 1: The R^2 and estimated g/λ and λ/μ for Biotechnology Stocks

As a further illustration, Figure 6 shows the empirical and estimated size distribution for

---

[17]Like in the linear regression, here R^2 is simply defined as $1 -$ (variance of the residuals)/(variance of the observed responses).
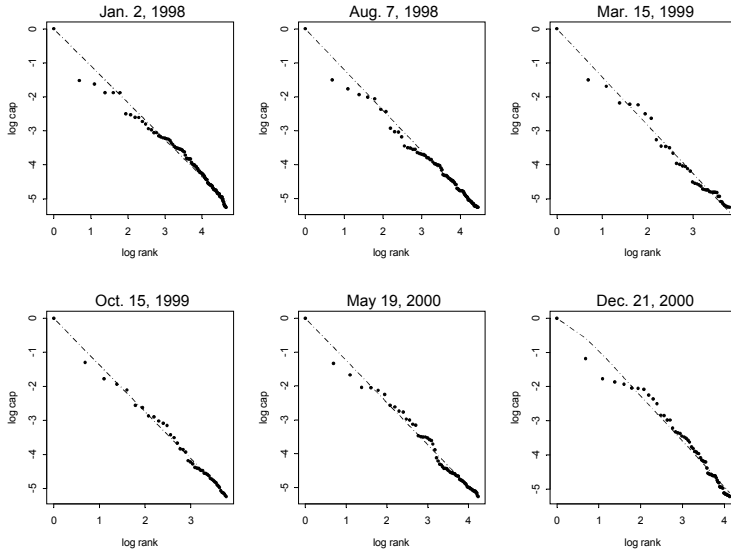
Figure 5: Empirical and estimated size distribution for biotechnology stocks

internet stocks. The six panels represent January 4, 1999 and every 100 trading days onward[18]. Again the expected linear pattern emerges. Table 2, for the internet stocks, reports the estimated parameters and the R^2, which is at least 94%.

	$\widehat{g/\lambda}$	$1 - \widehat{\lambda/\mu}$	R^2
Jan 4, 99	0.365	1.47×10^{-6}	97.3%
May 27, 99	0.298	1.60×10^{-6}	96.8%
Oct 19, 99	0.211	1.18×10^{-9}	99.0%
Mar 13, 00	0.135	1.12×10^{-7}	94.0%
Aug 3, 00	0.234	1.15×10^{-9}	99.5%
Dec 26, 00	0.315	4.51×10^{-7}	99.4%

Table 2: The R^2 and estimated g/λ and λ/μ for Internet Stocks

Equation (15) is considered here mainly because (a) it is easy to implement and (b) the focus here is more illustrative than precise estimation. We shall point out that there are other ways, such as likelihood based method, to estimate the parameters, which might be more efficient.

We conclude this section by presenting the picture of the recent market. Figure 7 shows

---

[18]Totally 70 internet stocks are involved. See Appendix C for the list. The plot starts from January 4, 1999 because there were not many internet stocks before 1999.
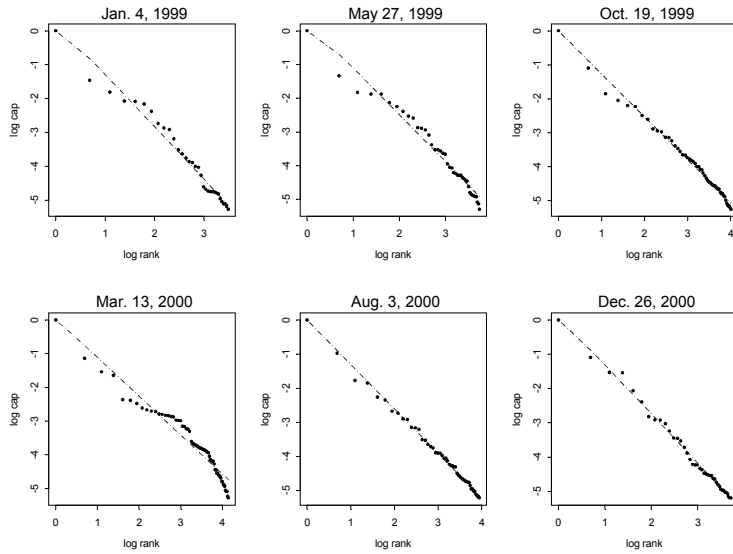
Figure 6: Empirical and estimated size distribution for internet stocks

the size distribution of biotechnology, internet stocks as of August 22, 2001. The clear linear pattern for biotechnology and internet stocks expected by the model again emerges. Table 3 reports the estimated parameters and the R^2. Note that the "internet bubble" has burst then; for example, the American stock exchange internet index (IIX) was 688.52 on 3/27/2000 and was only 141.21 on 8/22/2001. The fitting is well even under this severe market downturn.
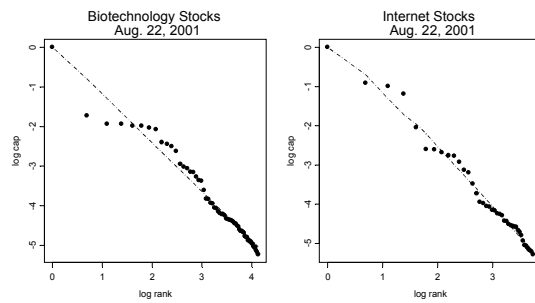

Figure 7: The size distribution for the recent market (August 22, 2001)

	$\widehat{g/\lambda}$	$1 - \widehat{\lambda/\mu}$	R^2
Biotech Stocks	0.192	9.23×10^{-7}	96.4%
Internet Stocks	0.362	3.43×10^{-6}	98.5%

Table 3: The R^2 and estimated parameters for the recent market (August 22, 2001)

VII   Discussion

By utilizing the high volatility of growth stocks, the paper proposes, based on both the transient and steady state behavior of birth-death processes, a model for growth stocks, which are otherwise quite difficult to price using traditional valuation methods.

There are two useful properties of the model. First, the model leads to a relative pricing formula, equation (13), which can be used to value growth stocks, including both biotechnology and internet stocks, relative to their peers. The method only uses regression and relative ranks, which are easy to implement. Secondly, the model remains valid irrespective to the market ups and downs, mainly because the model compares the value of a stock against the other stocks within its peer group.

There are several limitations of the model. (a) An problem that we are currently investigating is the possible effect of merger and acquisition. For example, currently (as of 2001) internet stocks have many more activities of merger and acquisition than biotechnology stocks. Thus, from this point of view, the current model is perhaps more suitable for biotechnology stocks than for internet stocks. (b) The model only applies to growth stocks with a large enough market capitalization, i.e. large-cap growth stocks. It does not attempt to provide a solution to small-cap growth stocks. (c) The model focuses on market capitalization, and does not take other possible factors, e.g. outstanding debt of companies, into account. One intuitive explanation of why the fit is good without including the debt is that most growth companies may not use debt as a major way of financing; thus the debt level may be low for growth stocks. (d) The model is not meant to be a tool for trading purposes. There are at least two reasons for this. First, we did not provide a dynamics of the relative ranks for growth stocks; therefore, the model is only meant as an understanding of growth stocks as a whole rather than as a model for any individual growth stock. Secondly, if after fitting the model to the data, one finds an "outlier", i.e. an individual stock whose market capitalization lies far away from the regression line, this does not necessarily mean that a trading opportunity arises; rather, one should pay special attention to the outlier and try to investigate other possible factors, such as debt level and merger and acquisition activities, related to the outlier.

21

A major disadvantage of the model, as pointed out by Herbert Simon and others for models based on the size distribution, is that they may put too much emphasis on the role of chance and too little on specific economic factors that might affect the distribution; see, for example, Lydall (p. 21, 1968). In this regard, it is encouraging to point out that Lucas (1978) provides an equilibrium justification of some size distributions for business firm sizes. We are investigating whether an equilibrium justification can be made for the current model as well.

Because of these limitations, as a cautionary remark, the model is only intended to provide a quick and first-order approximation[19] to a difficult yet important problem: how to value volatile growth stocks without any earnings.

Appendix

A   Proofs of the Propositions 1 and 2

Proof of Proposition 1. (1) To show that the birth-death process is positive recurrent, it is enough to check that $\sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n} = \infty$ and $\sum_{n=0}^{\infty} \pi_n < \infty$; see Kijima (1997, p. 245). The result follows as $\sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n}$ has the same order of $\sum_{n=0}^{\infty} \frac{1}{n\lambda+g} \left(\frac{\mu}{\lambda}\right)^n n^{1-g/\lambda+h/\mu} = \infty$, and $\sum_{n=0}^{\infty} \pi_n$ has the same order of $\sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n n^{g/\lambda-h/\mu-1} < \infty$, thanks to the assumption that $\mu > \lambda$.

(2) The equation (3) follows from the fact that $\lim_{z \to \infty} z^{b-a} \frac{\Gamma(z+a)}{\Gamma(z+b)} = 1$.

(3) First we consider $\sum_{k=n}^{\infty} \pi_k$, which, according to (2), is

$$\sum_{k=n}^{\infty} \pi_k = \alpha \sum_{k=n}^{\infty} \left(\frac{\lambda}{\mu}\right)^k \frac{\Gamma(k+g/\lambda)}{\Gamma(k+1+h/\mu)},$$

where $\alpha := \frac{\Gamma(1+h/\mu)}{\Gamma(g/\lambda)}$. The definition of the hypergeometric function yields

$$\begin{aligned} \sum_{k=n}^{\infty} \pi_k &= \alpha \left(\frac{\lambda}{\mu}\right)^n \frac{\Gamma(n+g/\lambda)}{\Gamma(n+1+h/\mu)} F(n+\frac{g}{\lambda}, 1; n+1+\frac{h}{\mu}; \frac{\lambda}{\mu}) \\ &= \pi_n F(n+\frac{g}{\lambda}, 1; n+1+\frac{h}{\mu}; \frac{\lambda}{\mu}), \quad \text{for } n \geq 0. \end{aligned} \tag{16}$$

In particular, we obtain that

$$S = \sum_{n=0}^{\infty} \pi_n = \pi_0 F(\frac{g}{\lambda}, 1; 1+\frac{h}{\mu}; \frac{\lambda}{\mu}) = F(\frac{g}{\lambda}, 1; 1+\frac{h}{\mu}; \frac{\lambda}{\mu}).$$

---

[19]Using a real option approach, Schwartz and Moon (2000) propose another interesting model to value internet stocks.

The moment generating function is given by

$$
\begin{aligned}
M(\theta) &= \sum_{n=0}^{\infty} e^{\theta n} \pi_n / S = \frac{1}{S} \sum_{n=0}^{\infty} \alpha \left(\frac{\lambda}{\mu} e^{\theta} \right)^n \frac{\Gamma(n + g/\lambda)}{\Gamma(n + 1 + h/\mu)} \\
&= F(\frac{g}{\lambda}, 1; 1 + \frac{h}{\mu}; \frac{\lambda}{\mu} e^{\theta})/S \\
&= F(\frac{g}{\lambda}, 1; 1 + \frac{h}{\mu}; \frac{\lambda}{\mu} e^{\theta})/F(\frac{g}{\lambda}, 1; 1 + \frac{h}{\mu}; \frac{\lambda}{\mu}).
\end{aligned}
$$

The results about the mean and the second moment follow easily via the following property of the hypergeometric function (see also formula 15.2.1 of Abramowitz and Stegun, 1972):

$$
\frac{d}{dz} F(a, b; c; z) = \frac{ab}{c} F(a + 1, b + 1; c + 1; z).
$$

(4) By (16),

$$
F(n) = \pi_n F(n + \frac{g}{\lambda}, 1; n + 1 + \frac{h}{\mu}; \frac{\lambda}{\mu})/S.
$$

So we only have to study the limiting behavior of $F(n + \frac{g}{\lambda}, 1; n + 1 + \frac{h}{\mu}; \frac{\lambda}{\mu})$. Using formula 15.3.5 of Abramowitz and Stegun (1972), $F(a, b; c; z) = (1 - z)^{-b} F(b, c - a; c; \frac{z}{z-1})$, we have

$$
F(n + \frac{g}{\lambda}, 1; n + 1 + \frac{h}{\mu}; \frac{\lambda}{\mu}) = (1 - \frac{\lambda}{\mu})^{-1} F(1, 1 + \frac{h}{\mu} - \frac{g}{\lambda}; n + 1 + \frac{h}{\mu}; \frac{\lambda}{\lambda - \mu}).
$$

But from the definition of the hypergeometric function, it is easily seen that (see Section 2.3.2 of Erdélyi et. al, 1953, Vol. 1) $F(1, 1 + \frac{h}{\mu} - \frac{g}{\lambda}; n + 1 + \frac{h}{\mu}; \frac{\lambda}{\lambda - \mu}) \to 1$. Therefore we obtain

$$
F(n) \cong \frac{\alpha}{S} (1 - \frac{\lambda}{\mu})^{-1} \left(\frac{\lambda}{\mu} \right)^n n^{g/\lambda - h/\mu - 1},
$$

from which the proof of Proposition 1 is terminated. $\square$

Proof of Proposition 2. We start from the forward Kolmogorov equations of a birth-death process (see Karlin and Taylor, 1975, page 136):

$$
\begin{aligned}
p'_{i,0}(t) &= -\lambda_0 p_{i,0}(t) + \mu_1 p_{i,1}(t), \\
p'_{i,j}(t) &= \lambda_{j-1} p_{i,j-1}(t) - (\lambda_j + \mu_j) p_{i,j}(t) + \mu_{j+1} p_{i,j+1}(t), \quad j \geq 1,
\end{aligned}
$$

which in our case is

$$
\begin{aligned}
p'_{i,0}(t) &= -g p_{i,0}(t) + (\mu + h) p_{i,1}(t), \\
p'_{i,j}(t) &= (\lambda(j - 1) + g) p_{i,j-1}(t) - ((\lambda + \mu)j + g + h) p_{i,j}(t) + (\mu(j + 1) + h) p_{i,j+1}(t), \quad j \geq 1.
\end{aligned}
$$

Multiplying the jth equation by j and taking a sum yields

$$m_1'(t) = (\lambda - \mu)m_1(t) + g + h(1 - p_{i,0}(t)),$$

with the initial condition $m_1(0) = i$. The solution is

$$m_1(t) = ie^{(\lambda-\mu)t} + \frac{g}{\mu - \lambda}[1 - e^{(\lambda-\mu)t}] + h\int_0^t e^{(\mu-\lambda)(s-t)}(1 - p_{i,0}(s))ds.$$

Similarly, multiplying the jth equation by j^2 and summing leads to

$$m_2'(t) = 2(\lambda - \mu)m_2(t) + (\lambda + \mu + 2g - 2h)m_1(t) + g + h(1 - p_{i,0}(t)),$$

with the initial condition $m_2(0) = i^2$. The solution for this differential equation is

$$
\begin{aligned}
m_2(t) \quad = \quad & i^2 e^{2(\lambda-\mu)t} + \frac{g}{2(\mu - \lambda)}[1 - e^{2(\lambda-\mu)t}] + h\int_0^t e^{2(\mu-\lambda)(s-t)}(1 - p_{i,0}(s))ds + \\
& + (\lambda + \mu + 2g - 2h)\int_0^t e^{2(\mu-\lambda)(s-t)}m_1(s)ds,
\end{aligned}
$$

from which the result follows. $\square$

B Calculating the Decay Parameters

To study the decay parameter of the birth death process, we start from the following proposition rephrased from Lemma 5.14 of Kijima (1997).

Proposition B.1. (Kijima, 1997) *There exists a sequence* $\{k_i\}$ *such that* $k_0 = \infty$, $k_i > 0$ *for all* $i \geq 1$, *and*

$$y = \lambda_i + \mu_{i+1} - \frac{\lambda_i \mu_i}{k_i} - k_{i+1}, \quad i = 0, 1, 2, \ldots,$$

if and only if $y \leq \gamma$.

Using the above proposition, we have

Proposition B.2. The decay parameter for the birth and death process (1) γ must satisfy $\gamma \geq \mu - \lambda$.

Proof. Consider the sequence $\{k_i\}$ defined by

$$\begin{cases} k_0 = \infty \\ \mu - \lambda = \lambda_i + \mu_{i+1} - \frac{\lambda_i \mu_i}{k_i} - k_{i+1}, & i \geq 0 \end{cases},$$

i.e.

$$\begin{cases} k_1 = \lambda + g + h \\ k_{i+1} = \lambda_{i+1} + \mu_i - \frac{\lambda_i \mu_i}{k_i}, & i \geq 1 \end{cases}.$$

Let $l_i = k_i - \lambda_i$, $i \geq 1$. Then

$$\begin{cases} l_1 = h \\ l_{i+1} = \frac{l_i}{\lambda_i + l_i} \mu_i, & i \geq 1 \end{cases}.$$

It is easy to see that $l_i \geq 0$ for all $i \geq 1$, which says that $k_i > 0$ for every $i \geq 1$. By Proposition B.1, we must have $\gamma \geq \mu - \lambda$. $\square$

The following result is useful in deriving an upper bound for the decay parameter γ.

Proposition B.3. *For any constant $c > 0$, consider the sequence k_i defined by*

$$\begin{cases} k_1 = \lambda + g + h - c, & c > 0, \ h \geq 0 \\ k_{i+1} = \lambda_{i+1} + \mu_i - \frac{\lambda_i \mu_i}{k_i} - c, & i \geq 1 \end{cases}.$$

Let $l_i := k_i - \lambda_i$, $i \geq 1$, which has the following recurrence relation

$$\begin{cases} l_1 = h - c, & c > 0, \ h \geq 0, \\ l_{i+1} = \frac{l_i}{\lambda_i + l_i} \mu_i - c, & i \geq 1 \end{cases}. \tag{17}$$

Then the sequence $k_i > 0$ for all $i \geq 1$ if and only if $l_i > 0$ for all $i \geq 1$.

Proof. Suppose $l_i > 0$ for all $i \geq 1$. Then immediately $k_i > \lambda_i > 0$ for every i. We shall prove the other direction by contradiction.

Suppose $k_i > 0$ for all $i \geq 1$ and $l_m \leq 0$ for some m. The recurrence relation gives (i) $l_{m+1} \leq -c$, (ii) $l_{i+1} \leq l_i \frac{\mu_i}{k_i}$, $\forall i \geq 1$. Therefore,

$$l_{m+2} \leq l_{m+1} \frac{\mu_{m+1}}{k_{m+1}} \leq -c \frac{\mu_{m+1}}{k_{m+1}}.$$

In general, for any $M > m + 1$,

$$l_M \leq -c \prod_{j=m+1}^{M-1} \frac{\mu_j}{k_j} < 0.$$

25

But note that $0 < k_j = l_j + \lambda_j < \lambda_j$, for $j \geq m+1$. We have

$$l_M \leq -c \prod_{j=m+1}^{M-1} \frac{\mu_j}{\lambda_j} < 0.$$

Since $\frac{\mu_j}{\lambda_j} \to \frac{\mu}{\lambda} > 1$, the above line tells us that $l_M \to -\infty$ exponentially fast. Therefore we must have $k_j = l_j + \lambda_j < 0$ for some $j \geq m+1$, contradicting the assumption. $\square$

Proof of Theorem 1. There are only two possibilities for γ, either $\gamma = \mu - \lambda$ or $\gamma > \mu - \lambda$. If $\gamma = \mu - \lambda$, then the statement in the theorem automatically holds.

Now suppose $\gamma = \mu - \lambda + c$, with $c > 0$. The recurrence relationship of l_i in (17) implies the simple fact that

$$l_{i+1} > d \geq 0 \text{ if and only if } l_i > (c+d)\frac{\lambda_i}{\mu_i - c - d}.$$

In particular, for any number $d \geq 0$,

$$\text{if } l_{i+1} > d \geq 0, \text{ then } l_i > (c+d)\frac{\lambda_i}{\mu_i} = c\frac{\lambda_i}{\mu_i} + d\frac{\lambda_i}{\mu_i}.$$

Using it once again, we know that if $l_{i+1} > d \geq 0$ then

$$l_{i-1} > c\frac{\lambda_{i-1}}{\mu_{i-1}} + \left(c\frac{\lambda_{i-1}}{\mu_{i-1}} + d\frac{\lambda_{i-1}}{\mu_{i-1}}\right)\frac{\lambda_i}{\mu_i} = c\frac{\lambda_{i-1}}{\mu_{i-1}} + c\frac{\lambda_{i-1}}{\mu_{i-1}}\frac{\lambda_i}{\mu_i} + d\frac{\lambda_{i-1}}{\mu_{i-1}}\frac{\lambda_i}{\mu_i}.$$

In general, simple induction gives that if $l_{i+1} > d \geq 0$ then for any $j \leq i$,

$$l_j > c\left\{\frac{\lambda_j}{\mu_j} + \frac{\lambda_j}{\mu_j}\frac{\lambda_{j+1}}{\mu_{j+1}} + \cdots + \frac{\lambda_j}{\mu_j}\frac{\lambda_{j+1}}{\mu_{j+1}}\cdots\frac{\lambda_i}{\mu_i}\right\} + d\frac{\lambda_j}{\mu_j}\frac{\lambda_{j+1}}{\mu_{j+1}}\cdots\frac{\lambda_i}{\mu_i}.$$

Letting $d = 0$, $j = 1$, $i \to \infty$ and using the fact that

$$\frac{\lambda_i}{\mu_i} = \frac{i\lambda + g}{i\mu + h} \geq \xi := \min(\frac{\lambda}{\mu}, \frac{\lambda + g}{\mu + h}) \text{ for any } i \geq 1$$

yield

$$l_1 > c\{\xi + \xi^2 + \cdots\} = c\xi/(1 - \xi).$$

But, by the definition in (17), $l_1 = h - c$, which gives

$$c < h(1 - \xi), \quad \forall h \geq 0. \tag{18}$$

Now if $h = 0$, then (18) leads to a contraction as c is assumed to be positive. Thus, when $h = 0$, γ must be equal to $\mu - \lambda$. If $h > 0$, then (18) yields

$$\gamma = \mu - \lambda + c < \mu - \lambda + h(1 - \xi) = \mu - \lambda + h[1 - \min(\frac{\lambda}{\mu}, \frac{\lambda + g}{\mu + h})],$$

from which the conclusion follows. $\square$

C Lists of the Stocks Used in the Numerical Illustration

Except the stocks (e.g. non-U.S. stocks) that are not included in the Center for Research in Security Prices (CRSP) historical database and the stocks no longer exist because of merger or bankruptcy, we use all the biotechnology stocks included in the Nasdaq biotech index (IXBT) and the Amex biotech index (BTK); all the internet stocks included in the Amex internet index (IIX), the Dow Jones composite internet index (DJINET), the Street.com internet index (DOT), the Amex Internet Infrastructure HOLDRS (IIH), the Amex B2B Internet HOLDRS (BHH), and the Amex Internet HOLDRS (HHH); and all the saving and loan stocks included in the Philadelphia exchange bank index (BKX), the S&P bank index (BIX), the regional bank HOLDRS (RKH), the Nasdaq Financial-100 index (IXF).

To save space, instead of the full names, we only list the stocks by their ticket symbols.

Dow Transportation Stocks

abf	alex	amr	bni	cnf	csx	dal	fdx	gmt	jbht
nsc	nwac	road	r	luv	ual	unp	u	usfc	yell

Internet Stocks

adbe	agil	akam	amtd	amzn	artg	athm	beas	brcm	bvsn
ckfr	cmgi	cmrc	coms	cs	csco	dclk	ebay	et	epny
exds	fmkt	goto	hlth	homs	icge	imgx	inap	inkt	insp
intu	isld	issx	itwo	iwov	jnpr	kana	lvlt	macr	navi
neta	novl	nxcd	omkt	pegs	pcln	ppro	prsf	psix	q
qcom	qrsi	retk	rnwk	rsas	scnt	sgi	spln	sqst	sunw
tibx	tmcs	tmpw	usix	vert	vign	vitr	vntr	vrsn	yhoo

Biotechnology Stocks

adrx	aimm	akrn	alks	allp	alxn	amgn	amln	anik	apht	aria	arql
astm	atis	atlc	atrx	avgn	avii	avir	avxt	axph	bcii	bcrx	bgen
blsi	blud	bste	btgc	btrn	carn	cbst	cege	ceph	cers	cgpi	chir
clgy	cnct	corr	crxa	ctic	cvas	cvtx	cypb	cyph	cyto	cytr	dcrn
dige	drmd	dsco	dusa	emis	enmd	enzn	epix	ergo	gene	genz	gern
ggen	gild	glfd	glia	gnlb	gnta	gztc	heph	hgsi	hysq	iart	iccc
icos	idph	ilxo	imcl	imgn	immu	imnr	imnx	incy	inhl	inkp	ipic
isip	kosp	lgnd	ljpc	lynx	matx	mcde	medi	medx	mlnm	mogn	mrvt
mygn	nabi	nbix	neot	nerx	nfld	novn	npro	npsp	nrgn	onxx	org
orph	osip	oxgn	pars	pcyc	pdli	pgnx	regn	rgen	rzym	sang	scio
scln	scri	sepr	sero	snap	snus	supg	teva	tgen	tktx	trms	vicl
vion	virs	vphm	vrtx	vvus	zmtx	zona					

				Saving and Loan Stocks					
amfi	asbc	asfc	aso	bk	bokf	bpop	caty	cbcf	cbsh
cbss	cfbx	chfc	cma	cmb	cors	cvbf	ewbc	fban	fcnca
fctr	ffbc	fitb	fmbi	fmer	ftbk	ftu	fult	gbbk	harb
hban	hbhc	ibnk	iboc	icbc	irwn	jpm	key	mafb	mel
mi	mrbk	ncc	npbc	ntrs	nwsb	nycb	ok	oldb	one
pbct	pbks	pfgi	pnc	rbnc	rgbk	rigs	rsln	sabb	sbib
sivb	skyf	snv	sotr	srce	stba	sti	stt	sub	susq
swbt	trbs	trmk	trst	tsfg	ubsi	ucbh	umbf	usb	wabc
wb	wbpr	wbst	wfsl	wm	wsbc	wtny	zion		

References

[1] Abramowitz, M. and Stegun, I.A. (1972). *Handbook of Mathematical Function.* U. S. National Bureau of Standards, 10th Printing.

[2] Axtell, R. L. (2001). Zipf distribution of U.S. firm sizes. *Science*, Vol 293, Sept. 7th, pp. 1818-1820.

[3] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *J. Polit. Econ.,* 81 , 637-659.

[4] Erdélyi, A. et. al (1953). Bateman Manuscript Project. *High Transcendental Functions.* Vol. 1, McGraw-Hill.

[5] Feenberg, D. and Poterba, J. (1993). Income inequality and the incomes of very high-income taxpayers: evidence from tax returns, in *Tax Policy and the Economy*, Vol. 7, J. Poterba ed. MIT press, pp. 145-177.

[6] Gabaix, X. (1999). Zipf's law for cities: an explanation. *Quarterly Journal of Economics.* Vol. 154, pp. 739-767.

[7] Gibrat, R. (1931). *Les inégalités économiques.* Paris, France.

[8] Glaeser, E., Scheinkman, J., and Shleifer, A. (1995), Economic Growth in a Cross-Section of Cities. *Journal of Monetary Economics*, Vol 36, pp. 117-143.

[9] Ijiri, Y. and Simon, H. A. (1977). *Skew Distributions and the Sizes of Business Firms.* North-Holland Publishing Company.

[10] Lydall, H. F. (1968). *The Structure of Earnings.* Oxford University Press.

[11] Karlin, S. and McGregor, J. (1958). Linear growth, birth and death processes. *Journal of Mathematics and Mechanics.* Vol. 7, pp. 643-662.

[12] Karlin, S. and Taylor, H. (1975). *A First Course in Stochastic Processes.* 2nd Ed. Academic Press.

[13] Kerins, F., Smith, J. K. and Smith, R. (2001). New venture opportunity cost of capital and financial contracting. Working paper, Washington State University.

[14] Kijima, M. (1997). *Markov Processes for Stochastic Modeling.* Chapman & Hall.

[15] Krugman, P. (1996a). *The Self-Organizing Economy.* Blackwell, Cambridge, MA.

[16] Krugman, P. (1996b). Confronting the urban mystery. *Journal of the Japanese and International Economies.* Vol 10, pp. 399-418.

[17] Lucas, R. (1978). On the size distribution of business firms. *Bell Journal of Economics.* pp. 508-523.

[18] Mandelbrot, B. (1960). The Pareto-Lévy Law and the distribution of income. *International Economic Review*, pp. 79-106.

[19] Mauboussin, M. J. and Schay, A. (2000). Still powerful: the internet's hidden order. Equity research report. Credit Suisse First Boston Corporation, July 7, 2000.

[20] Pareto, V. (1896). *Cours d' Economie Politique.* Geneva, Switzerland.

[21] Rutherford, R. (1955). Income distributions: a new model. *Econometrica*, Vol. 23, pp. 425-440.

[22] Schwartz, E. and Moon, M. (2000). Rational pricing of internet companies. *Financial Analysts Journal.* May/June, pp. 62-75.

[23] Shorrocks, A. F. (1975). On stochastic models of size distributions. *Review of Economic Studies.* pp. 631-641.

[24] Simon, H. A. (1955). On a class of skew distribution functions. *Biometrika.* Vol. 52, pp. 425-440.

[25] Simon, H. A. and Bonini, C. P. (1958). The size distribution of business firms. *American Economics Review.* Vol 48, pp. 607-617.

[26] Steindl, J. (1965). *Random Processes and the Growth of Firms.* Hafner, New York City.

[27] Steindl, J. (1968). Size distributions in economics, in *International Encyclopedia of the Social Sciences*, Vol. 14, Silks, ed., Macmillan Press, New York City.

[28] Yule, G. U. (1924). A mathematical theory of evolution, based on the conclusions of Dr. J. R. Willis, F.R.S. *Phil. Trans.* B, Vol. 213, pp. 21-83.

[29] Yule, G. U. (1944). *The Statistical Study of Literary Vocabulary.* Cambridge University Press.

[30] Zipf, G. (1949). *Human Behavior and the Principle of Least Effort.* Addison-Wesley Press.