Firm Growth Dynamics: The importance of large jumps

ARATA Yoshiyuki
University of Tokyo
Firm Growth Dynamics: The importance of large jumps

ARATA Yoshiyuki
Department of Economics, The University of Tokyo

Abstract

How a firm grows is one of the important themes in industrial organization literature. Recent empirical studies have demonstrated that the distribution of firms’ growth rates is not Gaussian as predicted by the celebrated Gibrat's law (Gibrat, 1931), but rather is quite well fitted by the Laplace distribution. These findings challenge the existing theoretical models and also our understanding of the mechanism of firm growth. To explain the empirical distributions, we consider the firm growth dynamics in the framework of the Lévy process and infinitely divisible distributions. Our analysis shows that the growth of a firm does not result from the accumulation of small shocks as the existing models assume. Instead, it is characterized by a handful of large shocks to the firm, i.e., jumps. The result has important implications for our understanding of the nature of innovations.

Keywords: Laplace distribution, Variance-Gamma process, Firm growth, Gibrat's law
JEL classification: L1, C1, D2
1. Introduction

How a firm grows is one of the important themes in the Industrial Organization literature. The early works include Ashton (1926), Gibrat (1931), Simon and Bonini (1958), Little (1962) and Steindl (1965), to name a few. In his seminal work, Gibrat (1931) attempts to explain the skewed pattern of firm size distribution. Gibrat’s law (Gibrat (1931)), that the growth rate of a firm is independent of its size, continues to receive much attention in the theoretical and empirical literature (c.f. Sutton (1997)). Gibrat’s approach models the size of an existing firm, $S_t$, as an accumulation of infinitesimally small shocks:

$$S_{t+1} = (1 + \epsilon^1_t)(1 + \epsilon^2_t)(1 + \epsilon^3_t)...(1 + \epsilon^n_t)S_t$$

Taking the logarithm of both sides, under weak conditions, the Central Limit Theorem (CLT) leads to a log-normal firm size distribution. Note that, in this model, the distribution of growth rates defined by $\log(S_{t+1}) - \log(S_t)$ becomes Gaussian. Following Gibrat (1931), Ijiri and Simon (1977) and later Sutton (1998) offer an explanation for skewed firm size distribution, widely known as the island model. They assume that firms compete for business opportunities in order to grow, and that the size of a firm reflects the number of opportunities it has gained. They show that the firm size distribution converges to the Yule distribution.

A different stream of literature that focus on the R&D investment decision and heterogeneity of firms has proposed theoretical models for firm-level empirical studies. In Ericson and Pakes (1995), firms are exposed to uncertainty arising from investment, leading to variability in their fortunes. Firms maximize the expected present value of their profits based on their beliefs about other firms’ behavior. They show that the beliefs of firms are fully consistent with the optimal decisions of all firms at an equilibrium. Klette and Kortum (2004), following Penrose (1959), propose a more parsimonious model. A firm is composed of the goods it produces and can expand into a new market through innovation. When the firm takes over a particular market through innovation, its competitor loses the market (creative destruction). They argue that their model can account for firm-level evidence, for example, the persistence of differences in R&D intensity among firms and why productivity growth is not strongly related to firms’ R&D.

Recent discoveries have led to interest in firm growth theory. Since Stanley et al. (1996), subsequent papers have shown that the distribution of firm growth rates is not Gaussian; rather, it closely follows the Laplace distribution. The finding is not consistent with the Gibrat model. More interestingly, the same characteristics of firm growth rate distributions appear at the finer sectoral level. This is a robust feature that holds at various levels of aggregation. On the other hand, the characteristics of firm size distribution differ significantly across industries (see Bottazzi et al. (2007)). This suggests that firm size distribution is just a consequence of aggregation, and therefore, the explanation

---

1In their analysis, new firms play an important role in the derivation of the Yule distribution. “What distinguishes the Yule distribution from the log-normal distribution is not the first assumption - the law of proportionate effect - but the second - the assumption of a constant “birth rate” for new firms. ... If we assume a random walk, but with a steady introduction of new firms from below, we obtain the Yule distribution”(Ijiri and Simon (1977), p.143).
of the distribution of firm growth rates is arguably more important for understanding firm growth dynamics. For this reason, we focus on the distribution of firm growth rates, not firm size in this paper.

The non-normality of growth rates challenges the existing models based on Gibrat’s model because it indicates that the distribution approaches Gaussian. Obviously, it is important to investigate why Gibrat’s model fails and which assumption of the model is violated. Several studies have attempted to explain these problems. For example, Bottazzi and Secchi (2006) assume a success brings success type of dynamic. A successful firm has a higher probability to achieve another success. They assert that this positive feedback generates a big leap and that the resultant distribution has a fatter tail. A different model, proposed by a group of physicists (Buldyrev et al. (2007b)) assumes that the distribution of firm growth rates is the convolution of two random variables: one is the number of products a firm produces and the other the size of each product. As we will see later, both models fail to explain the observed Laplace distribution.

This paper provides an alternative explanation for why Gibrat’s model fails and shows how the model can be modified to account for the Laplace shape. We relax one of the assumptions of Gibrat’s model, and show that the family of infinitely divisible distributions is appropriate to describe firm growth dynamics. We further show that the Variance-Gamma process corresponding to the Laplace distribution represents the underlying firm dynamics. We focus on the sample path properties and show that firm growth is not the consequence of an accumulation of many small shocks. Rather, the firm occasionally jumps. Put differently, large shocks dictate the fortunes of firms. The process of growth can be well described by a handful of large shocks, and therefore we can identify which events really lead to the prosperity (or decay) of the firm. This is in sharp contrast to the assumptions of the existing models such as Klette and Kortum (2004).

The rest of the paper is organized as follows. Section 2 summarizes the empirical evidence obtained in recent studies. Some evidence on the growth of Japanese firms is also provided to reconfirm the stylized facts. Section 3 develops a general framework to extend Gibrat’s law and investigates the sample path properties of the firm growth process. Section 4 discusses the implications of our analysis and concludes the paper.

2. Empirical Evidence

Recent studies show that the empirical distribution of growth rates of firms is not Gaussian. Stanley et al. (1996) analyze the publicly traded U.S manufacturing firms and find that the growth rates of firms display a tent-shaped density. They propose the Laplace distribution to describe the density:

$$p(x; \gamma, a) = \frac{1}{2a} \exp\left(-\frac{|x - \gamma|}{a}\right)$$

This density has a fatter tail than Gaussian; that is, there is a relatively higher probability that a firm experiences more extreme events than predicted by Gaussian. On the other hand, contrary to power-law distributions, this density has finite moments for all orders, because the tail decays exponentially.
Subsequent papers have confirmed this finding. Amaral et al. (1997) analyze all publicly traded U.S. manufacturing firms and find an exponential tent-shaped form rather than the bell-shaped Gaussian. Bottazzi and Secchi (2003) and Bottazzi and Secchi (2006) analyze the Italian manufacturing industry and Italian firms, respectively, and confirm the Laplace shape. Alfarano and Milaković (2008) analyze the Forbes Global 2000 list of the world’s largest companies and report a similar pattern.

Following these works, we study publicly traded Japanese companies for the years 2000–2010. The dataset we use is developed by Nikkei Digital Media Inc and contains data on thousands Japanese firms. We exclude banks, insurance, and security companies in our dataset. This is a commonly used procedure in the literature. The dataset also includes information on whether mergers and acquisitions occur in a year. This information can be used to identify whether the growth of a given firm is internal or external, such as M&A. Large outliers in the dataset are mainly due to external events. We simply exclude them because we focus on the process of internal growth in the present analysis.

In what follows we take the number of employees as the definition of firm size. Let \( S_t \) be the number of employees of a firm at time \( t \). It is customary to study firm growth on logarithmic scales, so the growth rate is defined as follows:

\[
g_t = \log S_{t+1} - \log S_t
\]

Figure B.1 shows a simple histogram of the distribution of growth rates of Japanese firms and the fitted Laplace distribution with the maximum likelihood estimate \( \alpha = 0.0541 \). It clearly displays a tent-shaped form as previous studies have observed. In Figure B.2, the histogram is replotted with a logarithmic vertical axis. When plotted in this way, it closely follows closely a straight line, which represents a Laplace distribution. Both depict marked departures from Gaussian and are rather Laplace.

3. Firm Growth Model: Variance-Gamma Process

As discussed above, non-Gaussian distribution of growth rates means the rejection of Gibrat’s model. Recent studies have tried to explain such empirical distributions through different mechanisms. For instance, Bottazzi and

\[f(x) = \frac{1}{2ab^{1/2}(1/b+1)} \exp\left(-\frac{1}{b} \frac{x-m}{a}\right)
\]

This family includes both Laplace \((b = 1)\) and Gaussian \((b = 2)\) distributions as special cases. They find that the shape parameter, \( b \), is significantly lower than 1; that is, the distribution of growth rates is more fat-tailed than the Laplace, contrary to previous works on the U.S. and Italian data.

In the Appendix, we briefly review the tail equivalence and show that our main conclusion holds even in these cases.

We take consolidated companies as the definition of firms.

The Laplace distribution emerges across different measures of size, such as sales or value added.

In what follows, we subtract the average growth rates from the growth rates; that is, we rescale \( \bar{g} = 0 \)
Secchi (2006) assume that “the probability that a given firm obtains new opportunities depends on the number of opportunities already caught” (p.251). Success leads to further success. This positive feedback generates big leaps and a fat-tailed distribution. Assuming that the Polya Urn statistics describe the feedback, Bottazzi and Secchi (2006) prove that the resulting distribution converges to the Laplace distribution. However, their analysis relies crucially on the assumption that “the process of opportunity assignment is repeated anew each year, i.e., that no memory of the previous year’s assignment is retained when the new year’s opportunities are assigned” (p.251). Otherwise, as time t passes, a single lucky firm would end up gaining all the opportunities and eventually diverging in size. Why is the feedback mechanism cut off abruptly at the end of the year? Why are firms not carrying over their technology or R&D investment to the next year? Although ingenious, the crucial assumption must be said to be arbitrary.

A different kind of model is proposed by a group of physicists (Buldyrev et al. (2007b)). They model a business firm as a portfolio of products. They assume that the number of products is a random variable whose distribution converges to an exponential distribution. Each product is assumed to follow an independent random process. Hence, the distribution of firm growth rates is a combination of these two random variables. In this model, the conditional distribution of growth rates of firms with the same number of products converges to a Gaussian distribution. The Laplace shape is due only to the aggregation of firms with a different number of products. This implies that if we disaggregate the data by, for example, firm size or sector, the Laplace shape would disappear. However, Stanley et al. (1996) show that the conditional distribution \( p(r|x_0) \), where \( x_0 \) is firm size (an initial sales value), is well described by Laplace. Moreover, Sakai and Watanabe (2010) empirically test this model and conclude that the Laplace shape is not generated as described by Buldyrev et al. (2007b).

Before we proceed to our own model, let us first discuss why Gibrat’s model fails and which assumption of the model should be changed. The basic assumptions of Gibrat’s model are as follows:

1. The growth rate of a firm is independent of its size.
2. The successive growth rates of a firm are independent of each other.
3. The growth rate of a firm consists of small shocks that satisfy the conditions of the standard CLT.

In the literature, a large number of studies have been devoted to test the validity of assumptions 1 and 2 (see Mansfield (1962), Singh and Whittington, (1975), Evans (1987) and Fujiwara et al. (2004); see also Santarelli et al. (2006) for a review). Although the results are inconclusive, they seem to suggest that assumptions 1 and 2 are not implausible and can be taken as a good starting point. For example, Fujiwara et al. (2004) analyze firms in European countries and show that the conditional distribution \( P(R|x_1) \) for large firms is independent of \( x_1 \), where \( x_1 \) is the size of a firm at a certain time, \( x_2 \) the size at a later time, and \( R = x_2/x_1 \). These findings strongly suggest that the observed Laplace distribution arises from the violation of assumption 3.

6^"One cannot conclude that the Law is generally valid nor that it is systematically rejected"(Santarelli et al. (2006),p.43). This is why many economists continue to test Gibrat’s law, more than 80 years after Gibrat’s publication.
We assume that the growth rate consists of independent random shocks and does not require the conditions of the standard CLT. A firm grows for all sorts of reasons, for example, the introduction of new products, quality improvement of existing products, cost reduction induced by technological progress, creation of new markets and effective advertisement. On the other hand, firm size can also decrease due, for example, to competition with other firms, serious conflicts with a labor union, change of consumer preferences and emergence of superior substitutes for the firm’s product. We can assume that these shocks are different in terms of their impacts on the firm’s performance. Some shocks are expected to have an unproportional impact on the firm’s performance. Therefore, we need a broad generalization of the standard CLT. A fundamental limit theorem on sums of independent random variables has been proven by Khintchin (1937). The theorem states that if the distribution of sums converges, it is infinitely divisible \(^7\). Infinite divisibility means that the random variable can be expressed by the sum of an arbitrary number of independent and identically distributed random variables. Gaussian and stable distributions belong to this family. Moreover, if an infinitely divisible distribution is given, there exists a corresponding Lévy process. The Lévy process is a natural generalization of Gibrat’s model in that assumptions 1 and 2 are satisfied. For convenience, we summarize the limit theorem and the definition of the Lévy process in the Appendix.

As is well known in the field of probability, the Laplace distribution is infinitely divisible. The characteristic function of the Laplace distribution with \( \gamma = 0 \) is written as

\[
\hat{\mu}(z) = \frac{1}{1 + az^2} = \frac{1}{(1 + i\alpha z)(1 - i\alpha z)}
\]

(2)

Factor \( \frac{1}{(1 - i\alpha z)} \) is the characteristic function of an exponential distribution, \( p(x) = 1_{(0,\infty)} \frac{1}{\alpha} \exp\left(-\frac{x}{\alpha}\right) \). This means that the Laplace distribution is obtained by the convolution of two exponential distributions. The characteristic function of an exponential distribution can be written in the form of (A.1) (Sato (1999), p.45),

\[
\hat{\mu}(z) = \frac{1}{1 - i\alpha z} = \exp\left[\int_{0}^{\infty} (e^{ix} - 1) \frac{e^{-\frac{x}{\alpha}}}{x} dx\right]
\]

(3)

where \( A = 0, \nu(dx) = 1_{(0,\infty)}(x)x^{-1}e^{-\frac{x}{\alpha}} \) and \( \gamma = 0 \) in (A.1). When \( \gamma \neq 0 \), a drift term, \( i(\gamma, z) \), in (A.1) is added. As discussed above, there exists a Lévy process corresponding to a given infinitely divisible distribution. This process, called the Variance-Gamma process, is what we need for our purpose. This was first introduced in option pricing theory by Madan and Seneta (1990), and the mathematical properties of the process and its generalization, pure jump processes, are fully explored by Ferguson and Klass (1972).

Let \( X_t \) be the logarithm of the size of a firm, \( \log S_t \). Equation (2) means that we can represent \( X_t \), that is, the Variance-Gamma process, as the difference between two independent and identical processes,

\[
X_t = Y^1_t - Y^2_t.
\]

---

\(^7\)A probability measure \( \mu \) on \( \mathbb{R}^d \) is infinitely divisible if for any positive integer \( n \), there is a probability measure \( \mu_n \) on \( \mathbb{R}^d \) such that \( \mu = \mu_n \), where \( \mu_n \) is the \( n \)-fold convolution of the probability measure \( \mu \) with itself.
Here, the characteristic function of $Y_t$ is written as

$$\hat{P}_Y(t) = \exp \left[ t \int_0^\infty (e^{ix} - 1) \frac{e^{-\frac{x^2}{2}}}{x} dx \right]. \tag{5}$$

Note that $Y_t$ is increasing as a function of $t$, a.s. This means that firm growth can be divided into two components, a positive growth process and a negative one, both having the same distribution.

By using the Lévy-Itô decomposition, we can express the sample paths of the Lévy process as a sum of the jump part and continuous part. From (5), $A = 0$, $Y_t$ has no continuous part, so the process is a pure jump process. This is in sharp contrast to the Brownian motion, which Gibrat’s model predicts. Because the Brownian motion has only a continuous part ($\nu(dx) = 0$), the positive growth is assumed to have small continuous movements in Gibrat’s model.

On the contrary, the positive growth process in our model, $Y_t$, has discrete movements, that is, jumps. Note that $\int x \nu(dx) = \infty$; that is, the number of jumps in a finite interval is infinite. This is due to small jumps around 0, and by dropping these small jumps, we can well approximate the process by a compound Poisson process:

$$\hat{P}_{Y_{n,t}}(t) = \exp \left[ t \int_{\epsilon_n}^\infty (e^{ix} - 1) \frac{e^{-\frac{x^2}{2}}}{x} dx \right], \quad \epsilon_n > 0 \text{ and } \epsilon_n \to 0 \text{ as } n \to \infty. \tag{6}$$

$P_n$ converges to the distribution of $Y_t$ (for more details, see Sato (1999) and Madan and Seneta (1990)). When the firm growth process follows a compound Poisson process, jumps occur randomly according to a Poisson process and the size of each jump is also an independent random variable. In the above case, the jumps whose size is within the interval $(x, x + dx)$ occur as a Poisson process with intensity $\nu(dx) = x^{-1} e^{-\frac{x^2}{2}} dx$. That is, $Y_t$ is well approximated by

$$Y_{n,t} = \sum_{i=1}^{N(t)} J_i, \tag{7}$$

where $N(t)$ follows a Poisson distribution with rate $\lambda = \int_{\epsilon_n}^\infty \frac{e^{-\frac{x^2}{2}}}{x} dx$ and $J_i$ is an independent random variable with distribution $\lambda^{-1} \frac{e^{-\frac{x^2}{2}}}{x} dx$ on $[\epsilon_n, \infty)$. Figure B.3 depicts a typical sample path of $Y_{n,t}$.

The properties of the sample path discussed above shed new light on firm growth dynamics. Suppose that a firm increases its size through innovation, for example, the introduction of new products. An idea for an innovation arrives at the Poisson rate $\lambda$. The impact of each innovation on the firm size is also a random variable, $J_i$; that is, major innovations contribute greatly to the growth of the firm, while others turn out to be almost useless. A successful innovation would bring about rapid expansion of the firm. In this way, the firm size sometimes jumps. In other words, firm growth is not a consequence of infinitesimally small shocks as existing models assume. If we collect a small number of such large jumps (i.e., successful innovations), that would largely explain the firm’s growth. This suggests that we would be able to identify which events lead to the ups and downs of a firm.

This is reminiscent of the concept of radical innovation in the literature on innovation management. Radical innovations involve the development or application of something fundamentally new that creates a wholly new industry.

---

$^8$Note that $\int x \nu(dx) < \infty$
or causes a complete transformation of the market structure. In our model, the radical innovations correspond to large jumps. Radical innovations are critical to the long-term growth of firms and differ significantly from incremental innovation (Ettlie et al. (1984), Chandy and Tellis (2000), Leifer et al. (2000)). Leifer et al. (2000) noted that once a radical innovation was introduced into the existing market,

“products based on one technology were undermined by radically new ones — and incremental improvement to the old technology has done litte more than delay the eventual rout.” (p.3)

Small firms that engage and succeed in radical innovation usually bring down giants and gain the leading position in the market. Our finding that the growth of firms is determined by large jumps is consistent with these phenomena. Note that while the studies mentioned above are based mainly on interviews and case studies of large firms, our findings are derived explicitly from the distribution of growth rates.

Finally, by using the estimate $a = 0.0541$ in the previous section, we can identify the structure of the jumps. For example, a positive growth shock larger than 3% has intensity $\psi(0.03, \infty) \approx 0.498$. This means that such shocks occur, on average, once every two years. In case of an 8% growth shock, $\psi(0.08, \infty) \approx 0.103$, and once every ten years. Note that these jumps are comparable to annual growth rates (see Figure B.1). The probability of larger jumps can be calculated in a similar manner.

4. Concluding Remarks

In the existing literature, a large firm is viewed as the consequence of many small successes. For example, Bottazzi and Secchi (2006) develop a model to generate a Laplace distribution based on the well-known island model. What makes this different from previous studies is the feedback effect; that is, the larger the number of opportunities already obtained, the higher the probability for a given firm to obtain new opportunities. In their model, when the number of shocks per firm (in their notation $M/N$ where $M$ is a finite number representing business opportunities and $N$ the number of firms) goes to infinity, the distribution of growth rates converges to a Laplace distribution. Bottazzi and Secchi (2006) state that

“…competitive success is seen not as the outcome of a single lucky event granting one firm a persistent, dominant, position, but rather as a firm’s ability to build its new success, through a permanent struggle within an extremely volatile environment, on the basis of its past, successful, behavior.” (p252)

The feature of their model is the absence of large shocks.

Klette and Kortum (2004) view a firm as the portfolio of goods that it produces. The size of a firm is represented by the number of goods it produces. Let $n$ denote the number of the firm’s products. A firm grows through innovation and by obtaining new markets and shrinks when some other firm innovates on a good in its portfolio. The firm that successfully innovates on a particular good takes over the market for that good at the expense of its competitors in the sense of Shumpeterian creative destruction. This creative destruction is described by a Poisson process with $\mu > 0$;
that is, the incumbent firm loses one of the markets at the Poisson hazard rate $\mu$. The firm’s *innovation production function* $I$, the arrival rate of innovation, is homogeneous of degree 1 in $n$ and can be written as $I = n\lambda$. Hence, the time evolution of probability density $p(n, t)$ is described as follows

$$
\frac{dp(n, t)}{dt} = (n - 1)p(n - 1, t) + (n + 1)p(n + 1, t) - n(\lambda + \mu)p_0(n, t), \quad n \geq 1
$$

$$
\frac{dp(0, t)}{dt} = \mu p(1, t)
$$

The meaning of these equations is clear. For example, the first term on the right-hand side of the equation represents that firms whose size is $n - 1$ innovate and increase their size by 1. The reason their model is tractable and simple enough to obtain an analytic solution is that they view a large firm as just a combination of many small firms. As Klette and Kortum (2004) state, “A firm of size $n_0 > 1$ at date 0 will evolve as though it consists of $n_0$ independent divisions of size 1.

$\ldots$ the evolution of the entire firm is obtained by summing the evolution of these independent divisions, each behaving as a firm starting with a single product would.”(p995)

For large firms ($n \gg 1$), each shock is very tiny compared to their size ($1/n \ll 1$). Therefore, this model is a typical example showing that the a larger firm is the result of successive small innovations, that is, an accumulation of small shocks. Like Bottazzi and Secchi (2006), there are no major shocks that lead to the rapid expansion of firms.

However, firm growth dynamics are different from what these previous studies envision. In sharp contrast to them, our analysis shows that firm growth is characterized by large jumps. Sizable shocks occasionary hit a firm and the firm grows discontinuously. If we collect a few large jumps of a particular firm, we can describe the firm’s growth path quite well. In other words, we can identify the events (or shocks) that lead to the ups and downs of a firm. Hence, there exist particular shocks leading to the rapid expansion of larger firms. This is impossible when shocks are infinitesimally small. This is where our analysis sheds new light and deviates substantially from the existing literature.

We also discuss the robustness of our conclusion in the Appendix. There are some studies showing that the distribution of growth rates has a fatter tail than a Laplace. Although the Variance-Gamma process cannot be applied directly to such cases, by using the subexponential family and tail equivalence, we show that the existence of a fatter tail is equivalent to having large jumps. This complements our analysis and strengthens our conclusion. It also guarantees the robustness of our conclusion that large jumps determine firm growth dynamics.

Radical innovation (or large jumps) is a key element in the long-term success of a firm. Firms that dominate the existing market but lag behind in competition for radical innovation often fail to maintain the leading position. The Laplace distribution of growth rates clearly represents these phenomena. Moreover, radical innovation is an engine of economic growth. It revolutionizes existing markets or opens up whole new industries. As Schumpeter (1942) famously states, “creative destruction is the essential fact about capitalism.” Hence, the investigating the statistics of growth rates also has an important meaning in macroeconomics and firm growth theory. Multidisciplinary study of firm growth dynamics, especially radical innovation or large jumps, is very promising.
Acknowledgements

This research was conducted as part of a Research Institute of Economy, Trade and Industry (RIETI) research project (Issues Faced by Japan’s Economy and Economic Policy Part II : Population Decrease, Sustained Growth, Economic Welfare). It was also supported by Grant-in-Aid for JSPS Fellows (Grant Numbers 25-7736). The paper has greatly benefited from comments by Hiroshi Yoshikawa. I am also indebted to the participants in seminars and conferences at the University of Tokyo, RIETI, for their helpful comments and suggestions.

Appendix A. A Generalization of the CLT and Lévy processes

The limit theorem proven by Khintchin (1937) generalizes the standard CLT (for details, see Sato (1999)).

**Theorem Appendix A.1.** Let \( \{Z_{nk} \} \) be a null array\(^9\) on \( \mathbb{R}^d \) with row sums \( S_n = \sum_{k=1}^{r_n} Z_{nk} \). If, for some \( b_n \in \mathbb{R}^d, n = 1, 2, ..., \) the distribution of \( S_n - b_n \) converges to a distribution \( \mu \), then \( \mu \) is infinitely divisible (Khintchin (1937); see also Sato (1999), p.47).

It is well known that the characteristic function, \( \hat{\mu}(z) \), of infinitely divisible distributions can be represented by the Lévy-Khintchin formula:

\[
\hat{\mu}(z) = \exp \left[ -\frac{1}{2} \langle z, A z \rangle + i \langle y, z \rangle + \int_{\mathbb{R}^d} (e^{i\langle z, x \rangle} - 1 - i\langle z, x \rangle 1_D(x)) \nu(dx) \right], \quad z \in \mathbb{R}^d
\]  

(A.1)

where \( \langle ., . \rangle \) is an inner product, \( A \) is a symmetric nonnegative-definite \( d \times d \) matrix, \( \nu \) is a measure on \( \mathbb{R}^d \) satisfying

\[ \nu(0) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} (|x|^2 \wedge 1) \nu(dx) < \infty, \]

\( D \) is the closed unit ball, and \( y \in \mathbb{R}^d \).

Now, in the case of an infinitely divisible distribution, a corresponding Lévy process exists; that is, *if \( \mu \) is an infinitely divisible distribution on \( \mathbb{R}^d \), then there is a Lévy process in law such that \( P_{X_1} = \mu \) (Sato (1999), p.35). Lévy processes are defined as follows.

**Definition Appendix A.1.** A stochastic process \( \{X_t : t \geq 0\} \) on \( \mathbb{R}^d \) is a Lévy process if the following conditions are satisfied.

1. \( X_0 = 0 \ a.s. \)\(^{10}\).

2. For any choice of \( n \geq 1 \) and \( 0 \leq t_0 < t_1 < ... < t_n \), the random variables \( X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, ..., X_{t_n} - X_{t_{n-1}} \) are independent (independent increments property).

3. The distribution of \( X_{s+t} - X_t \) does not depend on \( s \) (temporal homogeneity or stationary increments property).

---

\(^9\)A double sequence of random variables \( \{Z_{nk} : k = 1, 2, ..., r_n; n = 1, 2, ...\} \) on \( \mathbb{R}^d \) is called a null array if for each fixed \( n \), \( Z_{1n}, Z_{2n}, ..., Z_{rn} \) are independent, and for any \( \epsilon > 0, \lim_{n \to \infty} \max_{1 \leq k \leq r_n} P(|Z_{nk} | > \epsilon) = 0 \)

\(^{10}\)The Condition 1, \( X_0 = 0 \), is not essential, because we only add a constant.
4. It is stochastically continuous\textsuperscript{11}.

5. There is $\Omega_0 \in \mathcal{F}$ with $P[\Omega_0] = 1$ such that for every $\omega \in \Omega_0$, $X_t(\omega)$ is right-continuous in $t \geq 0$ and has left limits in $t > 0$.

Appendix B. Tail equivalence

In this section, we briefly review a mathematical framework to analyze the sample path properties of firm growth under the assumption that the process is a Lévy process. We denote the right tail of a measure $\mu$ by $\bar{\mu}$, that is, $\bar{\mu}(x) = \mu(x, \infty)$, and the $\gamma$-exponential moment of $\mu$ by $\bar{\mu}(\gamma) = \int_{-\infty}^{\infty} e^{\gamma x} \mu(dx)$, where $\gamma \geq 0$. $f(r) \sim g(r)$ means that $\lim_{r \to \infty} f(r)/g(r) = 1$. The convolution of distributions $\mu$ and $\rho$ is denoted by $\mu * \rho$. We then introduce two classes of distributions, called exponential class (denoted by $L(\gamma)$) and convolution equivalent ($S(\gamma)$).

Definition Appendix B.1. Let $\mu$ be a distribution on $\mathbb{R}$. Suppose that $\bar{\mu}(r) > 0$ for all $r \in \mathbb{R}$.

1. $\mu \in L(\gamma)$ if $\bar{\mu}(r + a) \sim e^{\gamma a} \bar{\mu}(a)$ for all $a \in \mathbb{R}$.
2. $\mu \in S(\gamma)$ if $\mu \in L(\gamma)$, $\bar{\mu}(\gamma) < \infty$, and $\bar{\mu} * \bar{\mu}(r) \sim 2 \bar{\mu}(\gamma) \bar{\mu}(r)$ for all $a \in \mathbb{R}$.

When $\gamma = 0$, the distributions in $L(0)$ and $S(0)$ are called long-tailed and subexponential, respectively. The definition means a very slow decay of distributions in $L(0)$ and $\lim_{\epsilon \to 0} e^{\epsilon x} \bar{\mu}(x) = \infty$.

A close relationship exists between $S(\gamma)$ and the Lévy processes. We denote the Lévy measure of an infinitely divisible distribution by $\nu$ (see A.1). Let

$$
\nu_c(dx) = \frac{1}{\bar{v}(c)} 1_{(c, \infty)}(x)\nu(dx)
$$

(B.1)

denote the jump distribution for $\bar{v}(c) > 0$.

Watanabe (2008) proves the following theorem (see also Pakes (2004) and Pakes (2007)).

Theorem Appendix B.1. Let $\gamma \geq 0$. Let $\mu$ be an infinitely divisible distribution on $\mathbb{R}$ satisfying

$$
\int_{-\infty}^{\infty} \exp(\imath z x) \mu(dx) = \exp(\phi(z)), \quad z \in \mathbb{R}
$$

with

$$
\phi(z) = \int_{-\infty}^{\infty} (e^{\imath z x} - 1 - 1_{|x| \leq 1}(x)\imath z x)\nu(dx) + iaz - \frac{1}{2} b z^2
$$

Then the following are equivalent.

1. $\mu \in S(\gamma)$

---

\textsuperscript{11}A stochastic process $(X_t)$ on $\mathbb{R}^d$ is stochastically continuous if for every $t \geq 0$ and $\epsilon > 0$, $\lim_{\epsilon \to 0} P[|X_t - X_0| > \epsilon] = 0$. 

10
2. \( v_1 \in S(\gamma) \)
3. \( v_1 \in L(\gamma), \bar{\mu}(\gamma) < \infty, \text{ and } \bar{\mu}(x) \sim \bar{\mu}(y) \tilde{v}(x) \).

This theorem characterizes class \( S(\gamma) \) of infinitely divisible distributions and shows that the tail is completely determined by \( \tilde{v}(x) \), that is, the distribution of large jumps. We can ignore the effects of small jumps and the Brownian component on the tail of the distribution.

If the tail of the distribution of growth rates is fatter than a Laplace, the distribution is expected to belong to \( S(\gamma) \), or especially \( S(0) \). In fact, one can prove that the family of distributions with a regular varying tail is a proper subset of \( S(0) \) (for more details, see Embrechts et al. (1997) and the following discussion). Therefore, if the distribution of growth rates has a power law tail, as asserted by Fu et al. (2005), Buldyrev et al. (2007a) and Buldyrev et al. (2007b), it belongs to \( S(0) \).

We then turn to the Subbotin family with parameter \( 0 < b < 1 \) (Bottazzi et al. (2011)). Let \( \mu^b \) be a member of the Subbotin family with \( 0 < b < 1 \). Define \( \mu^b_\gamma = 1_{[0,\infty)}(x)\mu^b(dx) + \mu^b((-\infty,0)]\delta_0(dx) \). Corollary 2.1 in Pakes (2004) implies that for \( \gamma \geq 0, \mu^b_\gamma \in S(\gamma) \) if and only if \( \mu^b_\gamma \in S(\gamma) \). We then choose an absolutely continuous probability distribution \( \rho \) satisfying \( \lim_{x \to \infty} \tilde{\rho}/\tilde{\varphi} = c \) where \( 0 < c < \infty \), so that \( \tilde{\rho} = c_1 \exp(-c_2 x^b) \) for \( x > c_3 \), where \( c_1, c_2 \) and \( c_3 \) are constants. Because \( \rho \in S(0) \) implies \( \mu^b \in S(0) \) (see Embrechts et al. (1979)), all we have to do is to examine whether \( \rho \in S(0) \).

In fact, from Proposition A.3.16 in Embrechts et al. (1997), the necessary and sufficient condition of the membership of \( S(0) \), we can immediately obtain \( \rho \in S(0) \). Therefore, we obtain the desired result, \( \mu^b \in S(0) \).

In either case, we can apply Appendix B.1 with \( \gamma = 0 \). When \( \gamma = 0, \bar{\mu} = 1 \). Hence, Appendix B.1 states that \( \bar{\mu}(x) \sim \tilde{v}(x) \); that is, \( \mu \) and \( v \) decay at almost the same rate. In other words, we can estimate the rate at which large shocks occur by estimating the tail of the observed distribution of growth rates. Moreover, Theorem Appendix B.1 states that small jumps and the Brownian component are irrelevant to the shape of the tail. Only large jumps affect it. Therefore, the recent findings that the distribution of growth rates has a fatter tail than a Laplace distribution in some countries suggest the importance of large jumps. This complements our analysis developed in Section 3.

References


Figure B.1: Growth rates for Japanese firms subtracting the average growth rate (= −0.00490). The solid curve is (1) with a maximum likelihood estimate \( a = 0.0541(0.000622) \). (The value in the parentheses is the Std. Error.) The dotted curve is a Gaussian distribution with its standard deviation 0.0958.
Figure B.2: The same data in Figure 1 with logarithmic scale. The y-axis is $\log(P(T))$.

Figure B.3: A sample path of $Y^1_{1,n}$. $Y^1_{1,n}$ is constant until a shock occurs. When a shock occurs, $Y^1_{1,n}$ jumps by $J_i$. 