Firm Size Dynamics of Industries:

Stochastic Growth Processes,
Large Fluctuations,
and the Population of Firms
as a Complex System

By

Daniel Teitelbaum and Robert Axtell
NuTech Solutions, Inc.
Arlington, VA

for

SBA
Office of Advocacy

Release Date: January 2005

Firm Size Dynamics of Industries:
Stochastic Growth Processes, Large Fluctuations, and the Population of Firms as a Complex System

By Daniel Teitelbaum and Robert Axtell, NuTech Solutions, Inc., Arlington, VA
under contract no. SBAHQ-03-Q-0015 (2005, 22 pages)

Purpose

With the emergence of various longitudinal data sources for firms, studies on business size distributions, business turnover, and business growth/decline have had a resurgence among the academic community. This study was conducted to better understand the dynamic nature of the growth and decline of firms in various size classes and industries.

Overall Findings

Firm growth rates do not fit the often assumed normal distribution. As a result, there are more slow-growing firms, more fast-growing firms, and fewer medium-growing firms than are generally modeled by economists. This finding held for all industries analyzed and growth rates were unrelated to establishment size.

Highlights

• The distribution of the growth of firms mapped a pyramid shape of the Laplace distribution rather than the inverted u-shape of normal distribution. Hence, business growth tends more to the tails (or extreme gains and losses) and less to the median than is commonly believed.

• Looking at the North American Industry Classification Systems' (NAICS) 20 major industry sectors revealed differing mean growth rates but similar distributions and variances. Particularly surprising was the similar results from two industries that had negative mean growth rates. The finance industry

was unique in that its pyramid shape had convex curved sides. This may have been a result of the time period chosen as equity markets were exploding from 1998 to 1999.

• Both small and large size classes exhibited similar distributions of growth rates and variances. The similar variance findings differed from previous research.

• This research is an exciting first step into growth analysis that generated some interesting questions such as would more detailed industry classes produce differing results, would these results hold up for the growth distribution of firms by age instead of size and is there persistence among the growth of establishments over time?

• With the findings from this report, economists will have to reevaluate their econometric tools when modeling the growth of businesses.

Scope and Methodology

The researchers utilized special tabulations from the U.S. Census Bureau's Statistics of U.S. Business (SUSB). SUSB excludes farms but includes nearly all employer firms in the U.S. Establishments (or business locations) surviving from 1998 to 1999 were the basis for the study which employed a method to approximate firm size associated with the establishments. Establishments that opened and closed during the period were not included in the analysis. Note that the Office of Advocacy is a partial funder of the SUSB.

The log of growth rates for establishments were charted and checked with a Kolmogorov-Smirnov goodness-of-fit test to determine statistically if they more closely matched a normal or Laplace distribution.

This report was peer-reviewed consistent with Advocacy's data quality guidelines. More information on this process can be obtained by contacting the Director of Economic Research at advocacy@sba.gov or (202) 205-6533.

Ordering Information

The full text of this report and summaries of other studies performed under contract to the U.S. Small Business Administration's Office of Advocacy are available at *www.sba.gov/advo/research.* Copies are also available from:

National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161
(800) 553-6847 or (703) 605-6000
(703) 487-4639 (TDD)
www.ntis.gov
NTIS order number: PB2005-101701
Pricing information:

Paper copy	A03 ($26.50)
Microfiche	A01 ($14.00)
CD-ROM	A00 ($18.95)
Download	A00 ($ 8.95)

To receive email notices of new Advocacy research, press releases, regulatory communications, and publications, including the latest issue of *The Small Business Advocate* newsletter, visit *http://web.sba.gov/list* and subscribe to the appropriate Listserv.

Executive Summary

This report describes several important features of firm growth rate dynamics, particularly as they manifest themselves across industries. The primary analytical perspective is that of stochastic dynamics and complex systems, in which the fluctuations in sizes of firms over time—that is, their distribution of growth rates—determine much of the overall structure of firms. Specifically, we report three major findings, some of which were previously known for some subset of U.S. firms and are here shown to obtain for the entire universe of U.S. business firms, and others that are apparently novel.

First, growth rate dynamics overall and across industries are well characterized by the so-called Laplace distribution, i.e., a two-sided exponential distribution. Firm growth rates are *not* normal. This has important implications for overall firm survival since the Laplace distribution has much heavier tails than the normal, meaning that some U.S. firms experience much stronger fluctuations than would be encountered if the economy were more normal (Gaussian). We argue that it is precisely small businesses that bear the brunt of such fluctuations, since they are sufficiently numerous that a certain fraction of them experience such giant fluctuations. But it is also the case that the largest *positive* fluctuations are experienced by some of the smallest firms. In fact, such fluctuations are largely responsible for what makes the small business sector so productive, an engine of economic growth.

A second finding has to do with the relatively modest differences in growth rate distributions across (two digit) sectors. Specifically, we find that the heavy-tailed character of growth rate distributions is nearly ubiquitous. This is surprising since the number of firms within such sectoral classifications differ by orders of magnitude, age cross-sections vary substantially, technological shocks are presumably quite different, and so on. The only significant deviations from Laplace-distributed growth are toward even *heavier*-tailed growth, not toward normality.

Third, we have discovered that for data on establishments, the variance in growth rates does not depend on establishment size. This is a surprising finding from a number of perspectives. Essentially what it means is that a large establishment is as likely to lay-off 10% of its workers, say, as is a small establishment, or that it is as probable for a small firm of 10 employees to hire 2 new people as it is for a size 100 firm to hire 20 new workers. We then attempt to relate this result to the firm-level notion of declining growth rate variance with size, by considering the multi-establishment character of firms.

From these findings we draw a variety of conclusions, some of which may be relevant for policy. Since extreme events/large fluctuations are an inherent part of firm growth, especially for small firms, government policies whose goal is to shield firms from fluctuations may be ineffective. However, policies designed to help firms ride out difficult times may go a long way in promoting firm survival. Similarly, the heavy-tailed character of positive growth fluctuations may be largely responsible for the intrinsic dynamism of small businesses, and any policy aimed at dampening these fluctuations may have, as an unwanted side effect, a weakening of the ability of the small business sector to drive growth. An important methodological conclusion that the present investigation highlights is the need for more systematic longitudinal data on firms, primarily multi-establishment ones. For economists a further sharp result is that econometric work utilizing normal specifications of log growth rates is badly misspecified.

The growth of firms is a critical, even foundational, process to modern, industrial economies, involving product, labor and financial markets. A comprehensive, empirically grounded understanding of this process has yet to be fully developed. The present study, highlighting important dynamical aspects of firm growth, contributes to this research enterprise.

I   Background

The important role played by firm *size* in the economy has been frequently studied in the economics research literature (for instance, Ijiri and Simon [1977]). Additionally, the role of firm size *dynamics* has been the subject of extensive research, primarily framed as how firm growth rates depend on size, industry, age of plants and establishments, firm and industry lifecycles, and so on (e.g., Hart and Prais [1956], Hart [1962], Hymer and Pashigan [1962], Evans [1987a, 1987b], Hall [1987], Dunne and Hughes [1994], Hart and Oulton [1996]). The role of firm sizes *within* industries has also received significant attention in the literature, primarily from a comparative perspective across the full span of industries. For example, Quandt [1966] studied how the firm size distribution within an industry varied across industries. The role of industry *life cycles* has also been studied extensively and is not unrelated to product life cycles (e.g., Baum and McGahan, forthcoming).

Much is known from this previous work in industrial organization (IO). First and foremost, the overall distribution of firm sizes is highly skewed, with a small number of very large firms coexisting in the economy along with larger numbers of smaller firms (Simon and Bonini [1958]). Within industries, a similar situation obtains, with most dominated by a few large firms, although the exact form of such skew distributions seems to depend crucially on the industry in question (Schmalansee [1989]). Secondly, concerning firm size dynamics—primarily framed as firm growth in the IO literature—there are conflicting results. Some studies have confirmed the original Gibrat [1931] hypothesis, that growth rates are independent of firm size, while others seem to refute this, finding that growth rates fall with size. Thirdly, most studies of the dependence of growth rate on firm and establishment age demonstrate that average growth rate falls with age. There are other facts about firm size dynamics that are known as well, including the notion that there is more variance in job destruction than in job creation time series (Davis, Haltiwanger and Schuh [1996]), that wages increase with firm size (Brown and Medoff [1989]), and so on.

All of the previous studies have been essentially *econometric* in nature, utilizing extant Census and Dun & Bradstreet (Compustat) databases to estimate rather simple—usually linear or log linear—models of firm size dynamics, both within industries and for the economy overall. However, these dynamics are reasonably complex, and it is not clear that static econometric techniques do full justice to the rich data that exist. This criticism of the existing literature has been reinforced recently in a series of papers that use analytical methods beyond conventional econometric techniques to explore these data, techniques derived from the study of complex systems.

'Heavy-Tailed' Firm Growth

Stanley *et al.* [1996] demonstrated certain of these novel techniques in a paper published in *Nature*. There they showed that two previously unknown regularities exist in these data, and that these regularities are very robust.

Utilizing Compustat data they discovered first that firm size dynamics are very well described by a Laplace distribution of log growth rates.[1] Note the size of a firm at time t by S_t. This can be measured in a variety of ways—e.g., the number of employees, the firm's total receipts or revenue, its market capitalization—without significantly altering the basic results.[2] Call the firm's growth rate at time $t+1$, $R_{t+1} = S_{t+1}/S_t$, and the logarithm of this quantity, r_{t+1}. Then this first finding of Stanley and co-workers simply says that across all firms and over time the probability density function (PDF) of r, $f(r)$, follows

$$f(r) = \frac{1}{\sqrt{2}\sigma}\exp\left(\frac{-2|r - \bar{r}|}{\sqrt{2}\sigma}\right),$$

where $\bar{r}$ is the average log growth rate, a quantity very near zero in typical firm data. This observation has since been confirmed with non-U.S. datasets (e.g., Bottazzi and Secchi [2004]).

This non-normal distribution of firm growth rates is interesting both conceptually—one would think that from central limit theorem type arguments alone one could marshal a strong case for growth rates to be normally distributed—as well as numerically, for the Laplace distribution is 'heavy-tailed' with respect to a normal having the same mean and variance. That is, there are many more firms that experience extreme growth rates—both positive and negative—than one would expect if such rates were normally distributed. We can see this in Figure 1 below, where both the normal and Laplace distributions are plotted in semi-log coordinates (which turn the familiar 'bell-shaped' normal into a parabola, and the Laplace becomes 'tent' shaped).
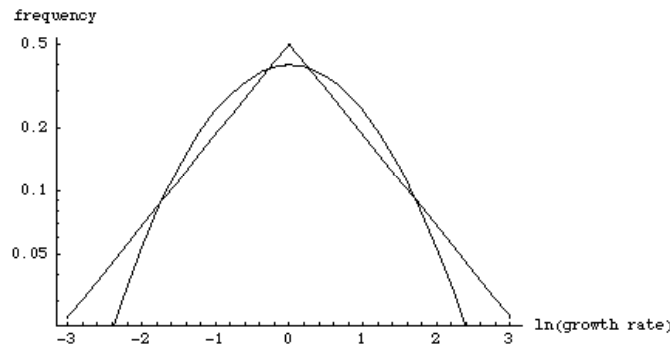
Figure 1: Comparison of the normal (curved line) and Laplace (straight line) distributions in semi-log coordinates; note that there is much more probability mass in the tails of the Laplace distribution (i.e., heavy tails).

To be somewhat more precise about the character of the heavy tails, consider the probability of extreme events. For example, 6σ and larger events occur with probability one in one billion for Gaussian random variables, while for Laplace

---

[1] For background on the Laplace distribution, see Johnson, Kotz and Balakrishnan [1994].
[2] The position of individual firms within a size distribution does depend on the size measure adopted. It is the overall shape of shape of such distributions that seems to be insensitive to how size is defined.

distributed variables such events occur somewhat more often than one in one thousand, a million-fold difference! Table 1 below shows how much more frequent tail events are in the Laplace distribution than in the normal.

n	$\Pr_{normal}[r \geq n\sigma]$	$\Pr_{Laplace}[r \geq n\sigma]$	$\Pr_{Laplace}/\Pr_{normal}$
2	2.3×10^{-2}	6.7×10^{-2}	2.97
3	1.3×10^{-3}	2.5×10^{-2}	18.4
4	3.2×10^{-5}	9.2×10^{-3}	289
5	2.9×10^{-7}	3.4×10^{-3}	1.18×10^{4}
6	9.9×10^{-10}	1.2×10^{-3}	1.25×10^{6}
7	1.3×10^{-12}	4.6×10^{-4}	3.56×10^{8}
8	6.2×10^{-16}	1.7×10^{-4}	2.70×10^{11}
9	1.1×10^{-19}	6.2×10^{-5}	5.47×10^{14}
10	7.6×10^{-24}	2.3×10^{-5}	2.98×10^{18}

Table 1: Characterization of the 'heavy tails' of the Laplace distribution

Given that there are approximately 6 million employer firms in the U.S., we shall see that such large deviations are empirically relevant.

Much of the literature in IO assumes that firm growth rates are normally distributed, so this basic result is not only surprising, but begs for an explanation and challenges conventional econometrics with its focus on normally distributed shocks.

Dependence of Growth Rate Variance on Size

The second important finding of Stanley *et al.* [1996] was that the variance in the Laplace distribution decreases with firm size according to a power law. Specifically, they found that $\sigma \sim s^{-1/6}$, and once again it did not matter very much how size was defined. Empirically, a similar relation has been found by Sutton [2002], using data on Japanese manufacturing, and by Bottazzi and Secchi [2003] using Compustat data. This relationship means that large firms have less variation in their growth than do smaller firms, in accord with intuition. In particular, as firm size increases by 10^6 the standard deviation in growth rates decreases by an order of magnitude. Given that the largest firm in the U.S.—Wal-Mart—has just about one million employees, the standard deviation in its growth rate should be fully an order of magnitude smaller than that of the smallest, single employee firms, on average.

As reasonable as this result seems, there is one outstanding and counter-intuitive aspect of it, again suggesting a departure from conventional statistics. If the processes within a firm that are responsible for firm growth were independent it would be reasonable to believe—assuming central limit-type behavior, again—that growth rate variance should fall like $\sqrt{s}$. Alternatively, if firm growth rate components were all perfectly correlated then we would expect that overall growth rates would be independent of size, and thus small and large firms would have the same growth rate variance, i.e., $\sigma \sim s^0 = $ constant. However, the empirical result is interior to each of these extremes; the exponent of $-1/6$ is

between –1/2 and 0, suggesting that there is significant correlation in firm growth rates, correlation that vitiates many of the assumptions of previous work in empirical IO. Today we are only beginning to understand this phenomenon (Wyart and Bouchaud [2003]), which may have important implications for growth rate fluctuations at the aggregate level (Gabaix [2003]). Perhaps of even greater interest to SBA is the fact that it is presumably the larger negative fluctuations (in percentage terms) among smaller firms that are mainly responsible for firm exit events. Developing a better understanding of the critical role such fluctuations play seems to be essential for developing public policy more focused on the needs of small to medium-sized businesses.

Extensions and Generalizations

Over the past few years, these results have been generalized theoretically and further tested empirically by Bottazzi and Secchi [2004]. They introduced a generalization of the Laplace distribution, the so-called Subbotin distribution [Subbotin 1923] as a family of densities which can be normal or Laplace, depending on the shape parameter, *b*

$$f(r) = \frac{1}{2ab^{1/b}\Gamma(1+1/b)} \, \exp\left(-\frac{1}{b}\left|\frac{r-\bar{r}}{a}\right|^b\right)$$

If the *b* parameter is unity the Subbotin distribution specializes to the Laplace distribution. In the case of *b* = 2 the normal distribution is recovered, while for large *b* the Subbotin approximates the uniform distribution. The distribution is shown for these and other values of the *b* parameter in Figure 2 below.
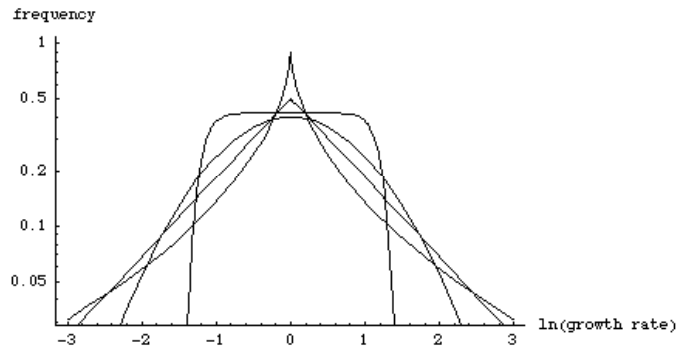
Figure 2: Comparison of the Subbotin distribution for different values of the shape parameter, *b* (for *a* = 1, mean value of 0): *b* = 1/2 is the most peaked distribution, *b* = 1 is the 'tent'-shaped , *b* = 2 is the parabola-shaped (normal) distribution, and *b* = 10 is the nearly uniform distribution

It is also possible empirically that this distribution can occur as an asymmetric one (e.g., Reichstein and Jensen [2003] and others), in which case one set of parameter applies to each side of the mean value.[3]

---

[3] For purposes of brevity we shall not present the functional form of the asymmetric Laplace or Subbotin distributions here, as the U.S. data turn out to be reasonably symmetric.

A Population of Firms as a Complex System

The intrinsically stochastic and 'heavy-tailed' character of firm growth rates, combined with the intra-firm coupling of growth rate processes, suggests that the overall behavior of the system of U.S. business firms is quite complex, and very different in character from the linear, Gaussian world of the IO textbooks. Large fluctuations and interdependence seem to characterize real firm dynamics. These are also features of complex systems (cf. Goldenfeld and Kadanoff [1999]).

A further element in the stream of work in IO influenced by complexity ideas is Axtell's [2001] discovery that, using data on all U.S. business firms that filed tax returns for 1997, the size distribution of firms follows a Pareto distribution with exponent very near unity, i.e., a so-called 'Zipf' distribution. Figure 3 below shows the main result.[4]
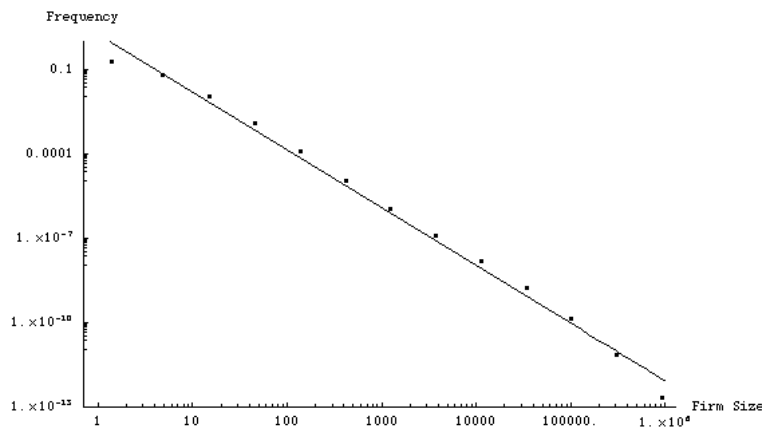
Figure 3: Size distribution of U.S. business firms in 1997 (Census data)

Since the time of Gibrat it had been thought that the lognormal distribution accurately described such data, but all previous analyses were based on limited samples. While the great number of small firms has been well-known since data on such entities began being gathered, the fact that a single power law fits essentially the entire distribution is further suggestive of the *complex systems* character of firm dynamics (e.g., Stanley [1995]).

The Zipf distribution is extremely skew. One interesting aspect of the empirical skewness in firm sizes is that the variance of this distribution is not meaningful. Indeed, it can be computed for a finite sample—there were about 5.5 million businesses with employees in the U.S. in 1997—but it diverges, meaning that it does not approach a limiting value as more data are added. In fact, the mean value of this distribution is barely meaningful—the actual average firm size of about 20 employees does not provide much information on the 'typical' firm size, for the median is closer to 2 while the mode is 1.

---

[4] In order to properly construct Figure 3 it was necessary to tabulate the data in a different way from that commonly used in previous analyses performed for SBA. Tabulations of the data in this paper better reflect the underlying skew character of firm size distributions than those usually employed by Census and SBA.

One important way to relate skew firm sizes, on the one hand, to stochastic growth models on the other, is through the so-called Kesten [1973] process. Today, little is know about the extent to which these complexity ideas—'power law' size distributions and 'heavy-tailed' growth rate distributions—transfer over to industry-specific data. Indeed, this is the main thrust of our analysis below.

II   Approach and Methodology

In looking at firm dynamics through the lens of complex systems, we first investigate the extent to which the results of Stanley *et al*. [1996] apply to the entire universe of U.S. business firms. For as Axtell [2001] showed, working with comprehensive datasets can yield quite different results than analyses of, for example, only publicly-traded firms (e.g., Stanley *et al.* [1995]). Then, the extent to which such results transfer to specific industries is studied. In the next section we describe the data we have used. Here we briefly describe the kinds of hypotheses to be tested and motivate their importance.

The questions we have tried to answer include the following:
  1. Does the Laplace distribution adequately describe firm growth rates across industries (Stanley *et al.* [1996] only analyzed the manufacturing sector)?
  2. Which industries are characterized by such 'heavy-tailed' distributions of growth rates? If only certain industries, are there features that explain these commonalities?
  3. Is the decline of variation in growth rates with size a universal phenomenon—in the sense of obtaining across all industries—or are the parameters and/or functional form of the dependence related to features of particular industries? For example, is it the case that so-called 'scaling' laws like that obtained by Stanley *et al.* [1996] have different exponents across industries? This is important for small businesses for having an exponent nearer –1/2 than 0 means that a firm can grow modestly and greatly dampen the kinds of fluctuations it experiences. To the extent that the log growth rate variance dependence on firm size is sector-specific, what are the features common to sectors that are responsible for this?
  4. Do newer, younger, more dynamic industries experience greater growth rate fluctuations, or are there reasons why their fluctuations are less dramatic—e.g., the ready availability of capital in technologically innovative areas?
  5. How do the dynamics of particular firms change across the industry lifecycle?

Our approach to resolving these questions will be a kind of hybrid pattern matching-econometric technique, where we will look for patterns in firm size and growth data and then fit simple functional forms to these data. We are essentially aiming to broaden the research agenda on the determinants of firm

growth, beyond a narrow econometrics focus, and believe that the kind of analysis we employ has certain advantages.

But why do we expect to uncover the kinds of patterns described above, e.g., heavy-tailed growth rates, and so on? How can these features be present in the data, especially given the usual intuition about the ubiquity of the normal distribution? It turns out there is a kind of central limit theory for how the Laplace distribution can arise, as well as its cousin the Subbotin. Due to space constraints we will not go into details here, but rather will refer the reader to recent work in this area. Suffice it to say here that when the number of summands in a random series is itself a random variable—it is a constant in the usual central limit theorem—and is approximately geometrically-distributed, then the attracting distribution is the Laplace, not the normal. In the context of firms we imagine the number of 'growth events' any particular firm experiences is a random variable, and that over time such events will shape a firm's size and structure, its form and function. Therefore, there is a theoretical underpinning to much of the purely empirical work to come, although we will not dwell on this here.

III   Data and Their Limitations

We utilized data from the Census Bureau's "Statistics of U.S. Businesses," which includes all tax-paying U.S. business firms with employees. These data have many strengths and certain weaknesses. Since nearly all businesses have to pay some kind of tax in the U.S., these data represent what is essentially the universe of all business firms. These data are available on an annual basis and are increasingly available online.

Specifically, in order to compute growth rates we utilize Census data for all U.S. establishments that had employees in 1998 and which continued in business into 1999.[5] While all establishments in this data are tied to enterprises in a given year, linkage of enterprises across years is problematical for various reasons, and this is the major limitation of these data. We have attempted to correct for this deficiency by converting establishment data into enterprise (firm) data through various distributions of the number of establishments per firm across firm size and industry/sector classifications.[6] Of course, our corrections are imperfect. We believe this to be a minor problem for small businesses, which typically have one or at most a few establishments, and a potentially bigger problem for large businesses.[7, 8] Throughout what follows we equate a firm's size with the number

---

[5] Note that this excludes all establishments that operated in but one of these years, i.e., all firm births and deaths that occurred in these years are explicitly excluded.

[6] Establishment growth and decline through mergers, acquisitions and spin-offs enter our analysis only through changes in the firm size distribution, which was minor over this period.

[7] In assessing other databases to utilize for this study, the non-comprehensiveness of each weighed so strongly against them that in the end we felt the lack of longitudinal enterprise linkages in the Census data presented the fewest problems.

[8] A further drawback of the use of these data by researchers is that they are subject to disclosure analysis in order to address privacy concerns. That is, the sensitive nature of the data held by Census limits the kinds of analyses that can be performed. For instance, whenever data is requested on an aggregation of firms that

of its employees. We conjecture that our analysis is robust to alternative specifications of size, but this is a topic for future work.[9]

IV   Analysis and Results

This section presents a variety of analyses of firm data, primarily growth rates as functions of size and industry. But we begin with what is in essence a check that these data are similar to those analyzed previously for 1997 (Axtell [2001]).

Establishment and Enterprise Sizes

We obtained from Census data on the size distributions of enterprises and establishments for 1998 and 1999. These in general have a strong 'power law' character. To see this, we plotted the $\log_{10}$ of the size of the enterprise on the horizontal axis, and the $\log_{10}$ of frequency on the vertical axis. If the data arrange themselves in a straightline then we have a Pareto distribution or 'power law'. The results are shown in Figure 4, below.

$$y = -1.0055x + 6.5667$$
$$R^2 = 0.9898$$

Log(freq)

log(Enterprise Size) All Industries, 1998
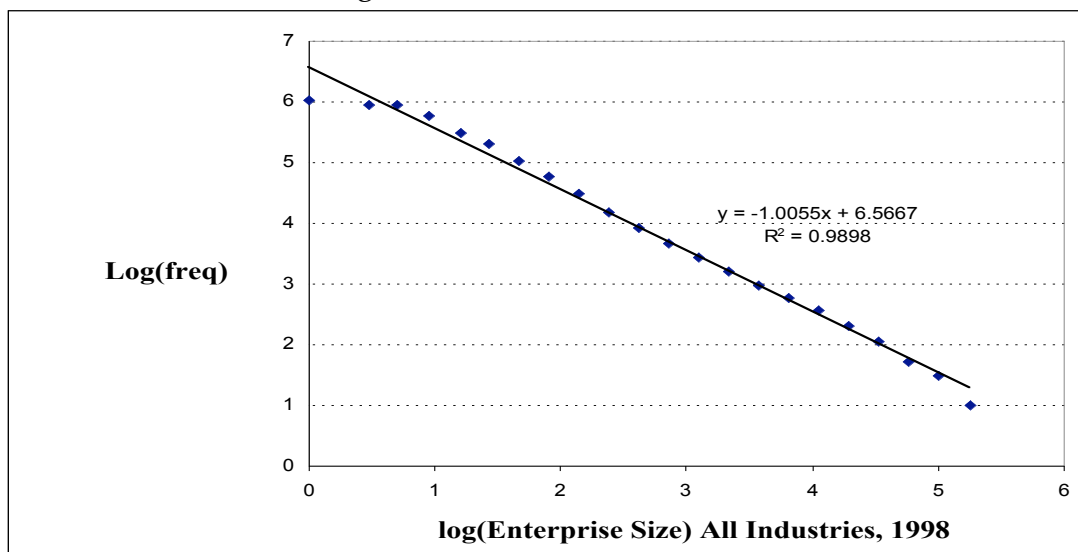
Figure 4: Size distribution of U.S. business enterprises in 1998 (Census data), log-log scale; size in terms of number of employees

U.S. firm sizes are approximately Pareto distributed with exponent extremely close to 1.0, a distribution is also known as the Zipf distribution [Perline 1996].

---

includes only one or at most a few firms then usually the data are suppressed. This can lead to important difficulties, such as when large tabulations have significant portions of them blocked for disclosure reasons.
[9] An argument against this conjecture is that declines in employment—such as occur when labor is replaced by capital or jobs are out-sourced—would not generally be reflected in sales or profit declines. The extent to which this is significant is an empirical question.
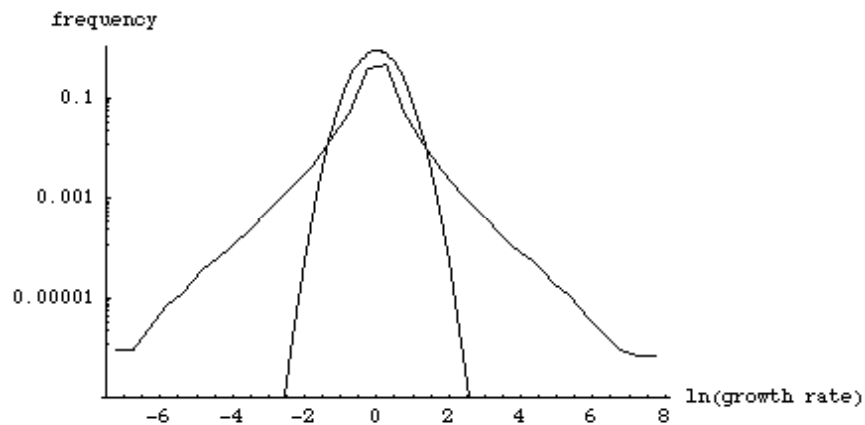
Distribution of Growth Rates

Next, we computed the distribution of establishment growth rates from data on the number of employees in continuing firms over 1998-99. We have also attempted to convert these into firm growth rates, through measures of average size. The mean growth rate over this period was approximately unity—no growth—making the mean log growth very close to 0. The variance in the log growth rate is about 0.20. The entire distribution of nearly 6 million entities is shown in Figure 5 in semi-log coordinates (better to emphasize tail events).
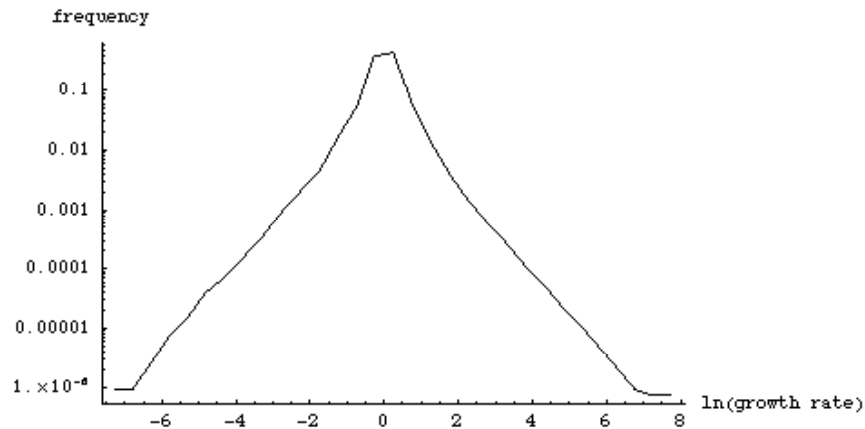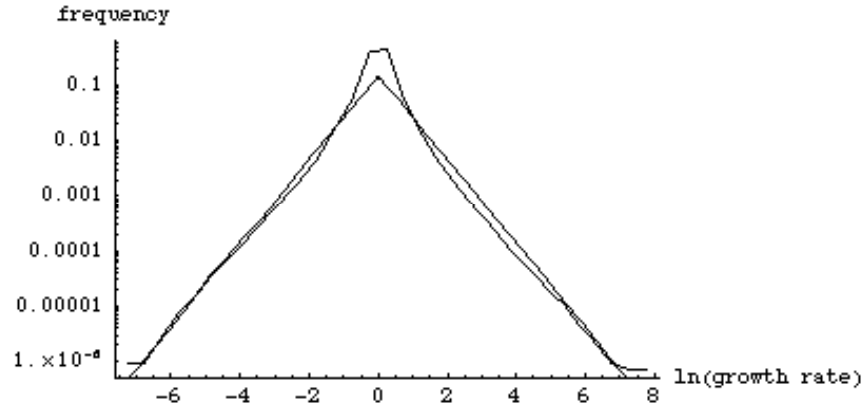

Figure 5: Distribution of log growth rates of U.S. business enterprises in 1999 (Census data), semi-log scale; growth rate in terms of changes in number of employees

When we fit these data to a normal distribution having the same mean and variance, Figure 6 results.
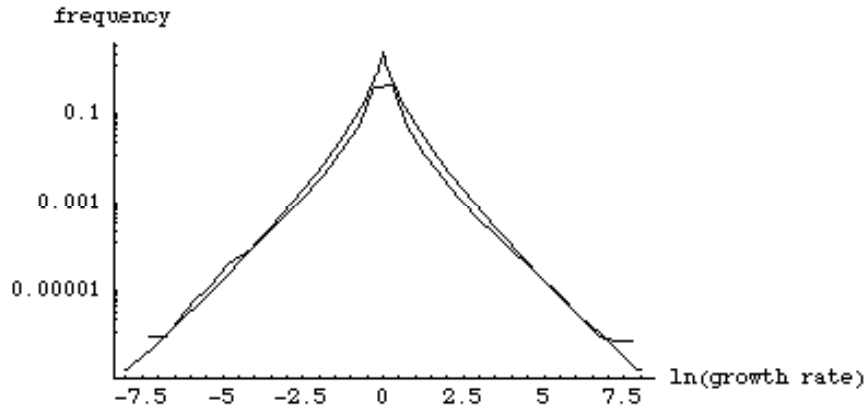

Figure 6: Comparison of empirical distribution of log growth rates of U.S. business enterprises in 1998 (Census data) with normal distribution having the same mean and variance, semi-log scale

9

Clearly, this is not a good fit as the normal misses nearly all of the mass in the tails. When a Laplace distribution is estimated, a much closer fit results, in accord with the results of Stanley *et al.* [1996], as can be seen in Figure 7.

Figure 7: Comparison of empirical distribution of log growth rates of U.S. business enterprises in 1998 (Census data) with Laplace distribution, semi-log scale

Note that while this distribution does an extremely good job through the tails, it misses much of the mass of the empirical distribution near the no growth mode, median and mean. So we estimate the Subbotin distribution and find that an exponent parameter $b \sim 0.7$ produces approximately the best fit, as shown in Figure 8 below.


Figure 8: Comparison of empirical distribution of log growth rates of U.S. business enterprises in 1998 (Census data) with Subbotin distribution ($b \sim 0.7$) having the same mean and variance, semi-log scale

This simple characterization of U.S. firm growth rate dynamics is new, we believe, and will serve as the basis for much of what follows.

Distribution of Growth Rates by Size

Census also provided data on growth rates tabulated by size, using some 19 equal-in-logs size bins that we specified. Growth rate frequencies are shown in Figure 9 below for three of these size bins along with all the data (uppermost curve).
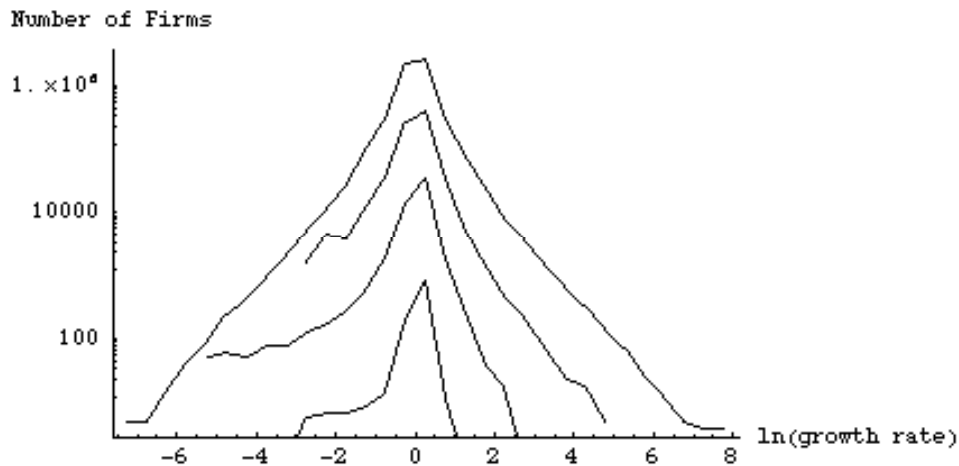
Figure 9: Histograms of growth rates of U.S. business firms in 1998-99 (Census data) for three distinct size classes and for the data overall; larger size classes have smaller support.

The lowermost curve is for one of the largest size bins, the next higher curve is for a medium sized bin, while the curve nearest the uppermost one is for one of the smallest bin sizes. The 'nesting' of the curves results from smaller numbers of firms in the larger bin sizes. These figures all show qualitative agreement with the Laplace distribution found by Stanley *et al.* [1996]. Notice that there is some asymmetry present, with the size-specific curves, in particular, displaying somewhat fatter left tails with their curvature suggestive of the Subbotin distribution. Overall, its reasonable to say that Laplace distributions of growth rates hold both for the economy as a whole, and when controlling for size.

At first blush, Figure 9 seems to also suggest that log growth rate variance is also declining with size, as indicated by the progressive narrowing of the distributions as one moves to higher size bins. However, this is not a correct inference. Explicit computation of the variance of these distributions reveals that it is relatively invariant over these size bins. Indeed, it even rises slightly with size, as shown in the following histogram, Figure 10.
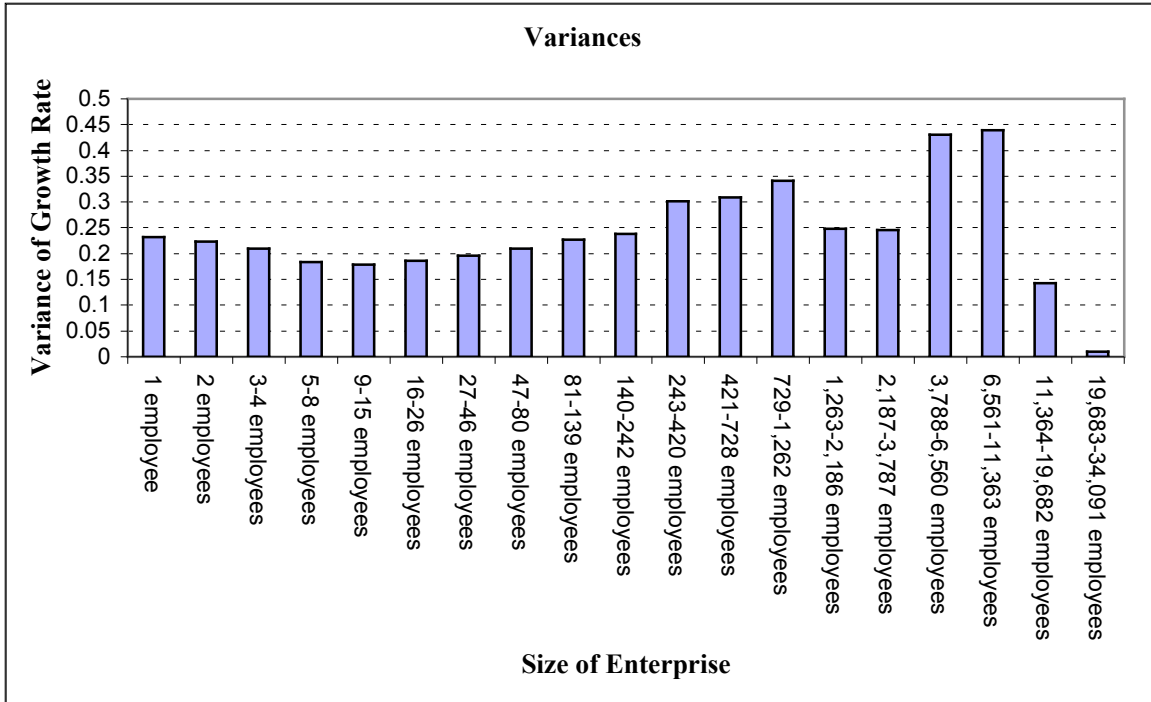
Figure 10: Variances of Growth Rates of U.S. business firms in 1998-99 (Census data)

We do not see the pattern we expected from Stanley *et al* [1996], i.e., decreasing variance in log growth rates with size. This result is surprising and requires further explanation. To see what the data tell us, graphically, we have to renormalize Figure 9 for the number of firms in the size cohort. Doing this gives us Figure 11.
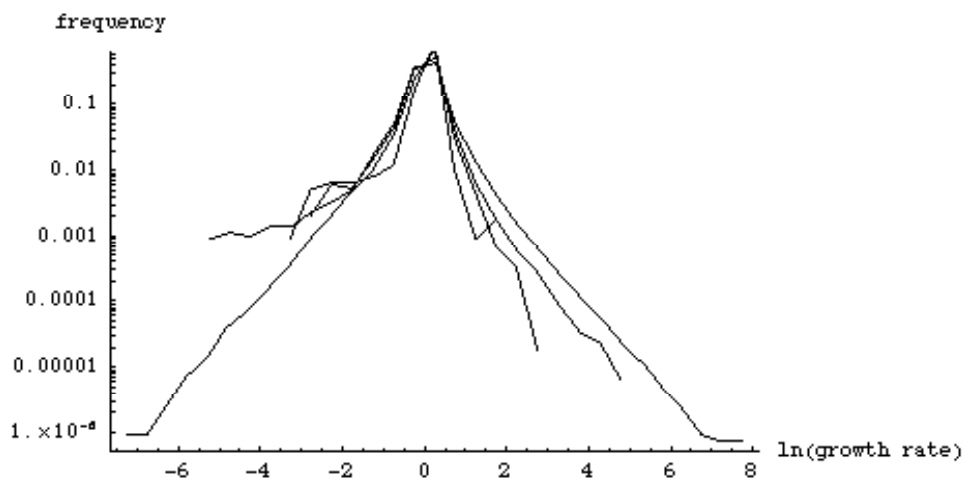
Figure 11: Histograms of growth rates of U.S. business firms in 1998-99 (Census data) for three distinct size classes and for the data overall; variance is essentially independent of firm size class

12

From this it is relatively clear that each of these curves has more or less the same variance, although there may exist some fine structure reflecting decreasing variance if one looks only at the right half—positive growth—of these distributions.

Distribution of Growth Rates by Industry

Census also provided us with growth rate data disaggregated by industry, which we used to analyze industry-specific growth rates at the coarsest classification.[10] We first computed the first two moments of these distributions, as shown in Table 2.

NAICS Sector	Code	Mean growth	Variance
Agriculture, Forestry, Fishing and Hunting	11	-0.0395	0.22
Mining	21	0.0419	0.16
Utilities	22	0.0792	0.27
Construction	23	0.0729	0.16
Manufacturing	31-33	0.0517	0.16
Wholesale Trade	42	0.0640	0.16
Retail Trade	44-45	0.0592	0.21
Transportation and Warehousing	48-49	0.0673	0.25
Information	51	0.0503	0.17
Finance and Insurance	52	0.0108	0.18
Real Estate and Rental and Leasing	53	0.0373	0.19
Professional, Scientific, and Tech. Services	54	0.0529	0.25
Management	55	0.0673	0.27
Admin, Support, Waste Mgmt., Rem. Svcs	56	0.0879	0.18
Education Services	61	0.0664	0.16
Health Care and Social Assistance	62	0.0563	0.24
Arts, Entertainment, and Recreation	71	0.0621	0.18
Accommodation and Food Services	72	0.0360	0.15
Other Services	81	0.0831	0.23
Auxiliaries, Exc Corp, Subsid., Reg. Offices	95	-0.0056	0.27
Total All Industries	N/A	0.0033	0.20

Table 2: Mean and variance in log growth rates for 20 sectors in 1998-99 (Census data)

While the mean growth rates are somewhat variable in this period, the variances are within a factor of 2, ranging from 0.15 to 0.27.

The Kolmogorov-Smirnov (KS) goodness-of-fit test was used to see how well the sector specific growth rate cross-sections were fit by a normal distribution versus a Laplace distribution. The KS test statistic measures the largest distance between the CDF of the data and the CDF of the distribution to which we are comparing.

---

[10] This was necessitated by data disclosure issues, i.e., finer classifications admit smaller numbers of firms in each growth rate-size bin and therefore require higher levels of disclosure analysis, typically leading to fewer data being made available. Potentially interesting growth patterns within industries are thus masked.

This test is non-parametric. For each industry, if the data are better fit by a Laplace distribution, we concluded that that industry is characterized by heavy-tailed growth distributions. The results are summarized in Table 3 below.

NAICS Sector	Code	KS Laplace	KS Normal
Agriculture, Forestry, Fishing and Hunting	11	0.148	0.378
Mining	21	0.205	0.379
Utilities	22	0.177	0.269
Construction	23	0.151	0.368
Manufacturing	31-33	0.129	0.316
Wholesale Trade	42	0.140	0.336
Retail Trade	44-45	0.170	0.330
Transportation and Warehousing	48-49	0.148	0.368
Information	51	0.147	0.338
Finance and Insurance	52	0.192	0.340
Real Estate and Rental and Leasing	53	0.167	0.429
Professional, Scientific, and Tech. Services	54	0.165	0.412
Management	55	0.120	0.290
Admin, Support, Waste Mgmt., Rem. Svcs	56	0.151	0.349
Education Services	61	0.132	0.402
Health Care and Social Assistance	62	0.175	0.319
Arts, Entertainment, and Recreation	71	0.148	0.370
Accommodation and Food Services	72	0.138	0.333
Other Services	81	0.178	0.378
Auxiliaries, Exc Corp, Subsid., Reg. Offices	95	0.144	0.298

Table 3: Comparison of Kolmogorov-Smirnov (KS) Tests for the normal and Laplace distributions for 21 Industries in 1998-99 (Census data)

We see in Table 3 that in each industry the data fit a Laplace distribution better than a normal distribution (i.e. the KS test statistic is lower for a Laplace than for a normal in every case.)

The question remains as to just how well the Laplace distribution fits these sector-specific growth rate distributions. In lieu of further analytics, we turn to graphical depiction of the growth rate distributions that served us earlier in this research. Specifically, for each sector we have fit Laplace distributions to the data using the first moments of the data to estimate the Laplace. Eight representative fits are shown in Figure 12 below, where each is a semi-log plot of frequency vs. log growth rate. These are for the following sectors, arranged from left to right, top to bottom: agriculture (NAICS 11), utilities (22), construction (23), manufacturing (31-33), retailing (44-45), transportation (48-49), information (51) and finance (52).[11]

---

[11] A plot of cumulative distribution function (CDF) values is perhaps somewhat more conventional but we have found that these probability density function (PDF) plots are more revealing of where the fits are good and not so good.
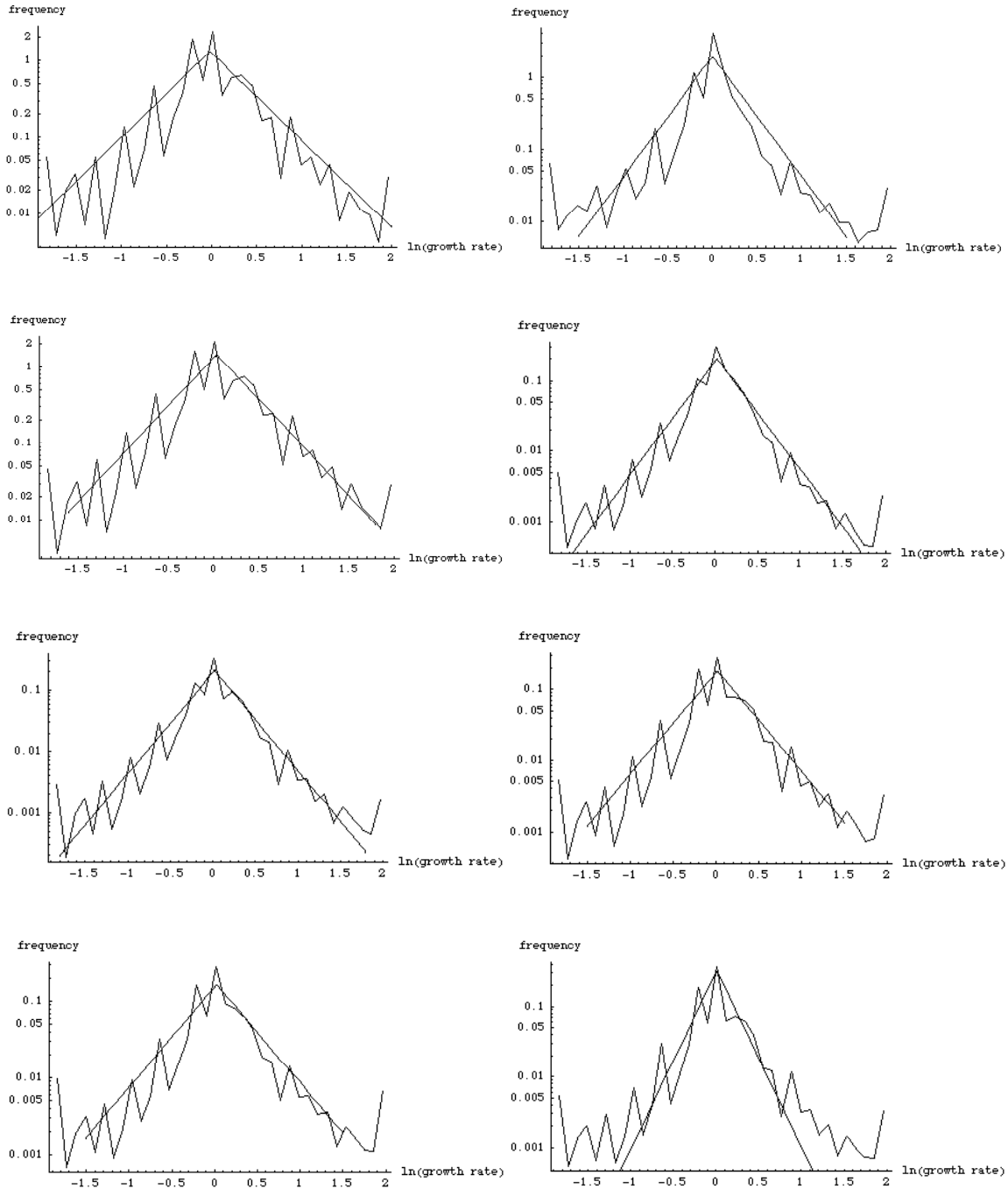
Figure 12: Laplace distribution of growth rates for 8 two digit sectors; from top left to bottom right, agriculture, utilities, construction, manufacturing, retailing, transportation, information, and finance

While these data are much more irregular than for the entire population of establishments and enterprises, each is reasonably well-fit by the Laplace distribution, with the exception of the finance sector. Note that in the case of finance there is even more mass in the tails than the Laplace can pick up, and so

it is natural to turn to the Subbotin with shape parameter, b, less than unity. Such an estimation is shown in Figure 13.
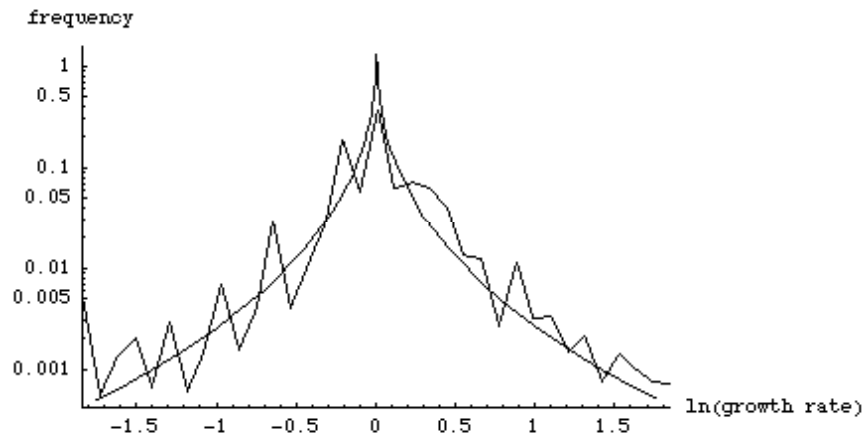
Figure 13: Subbotin distribution of log growth rates for the finance sector (two digit), $b \sim 0.4$

Here we see that the Subbotin is a better fit than the Laplace, a fact that is born out by a better K-S statistic in this case. It is unclear what the origin of this extremely heavy-tailed distribution is, and why it seems isolated to finance, although a reasonable speculation is that large fluctuations in financial markets are somehow responsible.

V   Discussion

While the importance of small businesses is well understood (Acs [1999], Acs and Audretsch [1990]), our results place a new emphasis on small firms. The modal firm size in the U.S. is a single employee, while the median is as few as 2-3 employees (depending on whether or not self-employed people are taken into account). The mean is only somewhat higher at about 20, and all these statistics point to the fact that the typical business entity in the U.S. is very small.

Is it the case that small firms face larger growth rate fluctuations than comparable large firms? Our answer, surprisingly, is both 'yes' and 'no.' It is 'no' in the sense that both small and large operations have essentially the same mean and variance—the same distribution—in growth rates. However, it is 'yes' in the sense that realizations of extreme fluctuations require large numbers of business entities, and that there are so few large firms (relatively speaking) that few ever experience giant fluctuations. However, the population of small firms is so large that some of them are essentially always experiencing very large fluctuations.

The years we examined—1998 and 1999—are the very peak of the so-called 'dot-com' boom. Perhaps this has affected our results. Start-up firms were rampant.

Venture capital was pouring into high-tech issues. Stocks were experiencing the a bull market. Perhaps those years were characterized by abnormal growth. Only by analyzing data from other years will we be able to tell. But our hypothesis is that, while certain parameters may have changed from that period to now, the qualitative character of our findings is likely to be unchanged.

Future Work

Conventional econometrics of firm growth has been dominated by specifications involving normal distributions in one form or another. The main contribution of the present work has been to show that the conventional approach is deeply problematical. Actual growth fluctuations of business entitites are much more heavy-tailed than the normal distribution permits, and it is the action in the tails that we are primarily interested in—negative growth events as a risk factor to firm survival and a harbinger of firm exit, positive growth as the key to economic growth.

We have uncovered in the Census data on firms the near ubiquity of heavy-tailed growth rate distributions in general and Laplace and Subbotin distributions in particular. Firms operating in economic environments characterized by the kinds of large fluctuations made possible by these heavy-tailed distributions will, of necessity, evolve into highly skew size distributions, with a tiny mode (unity in the U.S. data), a slighty larger median (2-3 for the U.S.) and a modest average size (especially in comparison to the largest firms).

Rather than providing exhaustive treatment of this subject, the present work provides an opening to future research endeavors along these lines. Many open questions remain:

1. As one 'drills down' into higher digit industries/sectors, does the Laplace characterization of growth rates break down at some point?
2. Are there better ways to relate establishments to enterprises than we have done here, and do such alternatives alter the basic picture in any substantive way?
3. To what extent can a 'declining growth rate variance with size' story be recovered for the largest firms, for which it seems to be true (Bottazzi and Secchi [2004]), perhaps as a consequence of the multi-divisional structure of coroporations and multinationals?
4. Is there an abrupt transition from no dependence of variance on size to power law dependence (as in Stanley *et al*. [1996]), or is it gradual?
5. What is the mechanism by which heavy-tailed annual and sub-annual growth rate fluctuations are aggregated into long term growth rates that are more Gaussian? That is, does large employment growth/decline persist for some firms or within some industries?
6. Similar analyses of firm growth by firm age, either in addition or in place of firm size, may also prove interesting.
7. Finally, given the significance of firm birth and death in the overall dynamics of the U.S. economy, if there was some way of naturally incorporating such firms into analyses such as these it would be important to do so. The problem as we see it is that to go from non-existent to having

employees involves essentially infinite growth on the part of a new firm, while to go from some number of employees to being out of business involves a log growth rate of essentially negative infinity. Perhaps certain robust statistical procedures can be developed for handling these extreme events, but this is opaque to us at present.

Some of these will require better data for their resolution, others simply more extensive analysis.

References

Acs, Z., ed. 1999. *Are Small Firms Different?* Kluwer Academic: Boston, Mass.

Acs, Z., C. Armington, and A. Robb. 1999. "Measures of Job Flow Dynamics in the U.S. Economy." Report in Partial Fulfillment of SBA Contract No. SBAHQ-98-0612.

Acs, Z. and D. Audretsch. 1990. *Innovation and Small Firms*. MIT Press: Cambridge, Mass.

Amaral, L.A.N., P. Gopikrishnan, V. Plerou, and H.E. Stanley. 2001. "A Model for the Growth Dynamics of Economic Organizations." *Physica A, 299*: 127-136.

Amaral, L.A.N., S.V. Buldyrev, S. Havlin, H.Leschhorn, P. Maass, M.A. Salinger, H.E. Stanley, and M.H.R. Stanley. 1997. "Scaling Behavior in Economics: I. Empirical Results for Company Growth." *Journal of Physique I France*, 7: 621-633.

Armington, C. 1998. "Statistics of U.S. businesses—Microdata and Tables." Report of SBA Contract #7.3112.0188.

Axtell, R.L. 2001. "Zipf Distribution of U.S. Firm Sizes." *Science, 293*: 1818-1820.

Axtell, R.L. 2004. "Firm Sizes: Facts, Formulae, Fables and Fantasies." Working paper. Center on Social and Economic Dynamics, The Brookings Institution: Washington ,D.C.

Baum, J.A.C. and A.M McGaham, eds. Forthcoming. *Business Strategy over the Industry Lifecycle*. Advances in Strategic Management, volume 21. Elsevier: New York.

Bottazzi, G. 2001. "Firm Diversification and the Law of Proportionate Effect." LEM Working Paper 2001/01. S. Anna School for Advanced Studies, Pisa, Italy.

Bottazzi, G. and A. Secchi. 2004. "Explaining the Distribution of Firms Growth Rates." Working paper. S. Anna School for Advanced Studies, Pisa, Italy.

———. 2003. "Sectoral Specifities [sic] in the Dynamics of U.S. Manufacturing Firms." LEM Working Paper 2003/18. S. Anna School for Advanced Studies, Pisa, Italy.

———. 2002. "On the Laplace Distribution of Firms [sic] Growth Rates." LEM Working Paper 2002/20. S. Anna School of Advanced Studies, Pisa Italy

Brown, C. and J. Medoff. 1989. "The Employer Size-Wage Effect." *Journal of Political Economy*, 97 (5): 1027-1059.

Buldyrev, S.V., L.A.N. Amaral, S. Havlin, H. Leschhorn, P. Maass, M.A. Salinger, H.E. Stanley and M.H.R. Stanley. 1997. "Scaling Behavior in Economics: II. Modeling of Company Growth." *Journal of Physique I France*, 7: 635-650.

Davis, S.J., J. Haltiwanger and S. Schuh. 1996. *Job Creation and Destruction*. MIT Press: Cambridge, Mass.

Dunne, P. and A. Hughes. 1994. "Age, Size, Growth, and Survival: UK Companies in the 1980s." *Journal of Industrial Economics, 42* (2): 115-140.

Evans, D. 1987a. "Tests of Alternative Theories of Growth." *Journal of Political Economy, 95* (4): 657-674.

Evans, D. 1987b. "The Relationship between Firm Growth, Size and Age: Estimates for 100 Manufacturing Industries. *Journal of Industrial Economics, 35* (4): 567-581.

Gabaix, X. 2003. "Power Laws and the Origins of Macroeconomic Fluctuations." Working paper. MIT: Cambridge, Mass.

Gibrat, R. 1931. *Les Inégalitiés Économiques; Applications: aux inégalitiés des recesses, á la concentration des enterprises, aux population des villes, aux statistiques des families, etc., d'une loi nouvelle, lo loi de 'leffet proportionnel.* Librarie du Recuel Sirey: Paris.

Goldenfeld, N. and L.P. Kadanoff. 1999. "Simple Lessons from Complexity." *Science, 284* (2 April): 87-89.

Hall, B. 1987. "The Relationship between Firm Size and firm Growth in the U.S. Manufacturing Sector." *Journal of Industrial Economics, 35* (4): 583-606.

Hart, P. 1962. "The Size and Growth of Firms." *Economics, N.S., 29* (113): 29-39.

Hart, P.E., S.J. Paris. 1956. "The Analysis of Business Concentration: A Statistical Approach." *Journal of the Royal Statistical Society, Series A, 119*: 150-191.

Hart, P.E. and N. Oulton. 1996. "Growth and Size of Firms." *Economic Journal, 106*: 1242-1252.

Hymer, S. and P. Pashigan. 1962. "Firm Size and Rate of Growth." *Journal of Political Economy, 70* (6): 556-569.

Ijiri and Simon, 1977. *Skew Distributions and the Sizes of Business Firms.* North-Holland: New York.

Johnson, N.L., S. Kotz and N. Balakrishnan. 1994. *Continuous Univariate Distributions.* John Wiley & Sons: New York, N.Y.

Kesten, H. 1973. "Random Difference Equations and Renewal Theory for Products of Random Matrices." *Acta Mathematica, CXXXI*: 207-248.

Kotz, S., T. J. Kozubowski and K. Podgorski. 2003. "An Asymmetric Multivariate Laplace Distribution." Working paper. George Washington University: Washington, D.C.

Perline, R., 1996. "Zipf's Law, the Central Limit Theorem, and the Random Division of the Unit Interval", *Physical Review* E, 54(1):220-223.

Quandt, R.E. 1966. "On the Size Distribution of Firms." *American Economic Review, 56* (3): 416-432.

Reichstein, T. and M.B. Jensen. 2003. "Analyzing the Distributions of the Stochastic Firm Growth Approach." Working paper 03-012. Danish Research Unit for Industrial Dynamics, Aalborg University, Denmark.

Schmalansee, R. 1989. "Inter-Industry Studies of Structure and Performance." In Schmalansee, R. and R.D. Willig, eds., *Handbook of Industrial Economics,* volume 2. North-Holland: New York.

Simon, H.A. and C.P. Bonini. 1958. "The Size Distribution of Business Firms." *American Economic Review, 46* (4): 607-677.

Stanley, H.E. 1995. "Power Laws and Universality." *Nature, 378*: 554-555.

Stanley, M.H.R., L.A.N. Amaral, S. Buldyrev, S. Havlin, H. Leschorn, P. Maass, M.A. Salinger and H.E. Stanley 1996. "Scaling Behavior in the Growth of Companies." *Nature, 379*: 804-806.

Stanley, M.H.R., S.V. Buldyrev, R. Mantegna, S. Havlin, M.A. Salinger, and H.E. Stanley. 1995. "Zipf Plots and the Size Distribution of Firms." *Economics Letters, 49*: 453-457.

Steindl, J. 1965. *Random Processes and the Growth of Firms.* Hafner: New York.

Subbotin, M.T. 1923. "On the Law of Frequency of Errors." *Matematicheskii Sbornik.*

Sutton, J. 2002. "The Variance of Firm Growth Rates: The 'Scaling' Puzzle," *Physica A, 312*: 577.

Sutton, J. 2001. *Technology and Market Structure*. MIT Press: Cambridge, Mass.

Wyart, M. and J-P. Bouchaud. 2002. "Statistical Models for Company Growth." Working paper. Available at arXiv: cond-mat/0210479 v2.