# Volatility and Asymmetry of Small Firm Growth Rates Over Increasing Time Frames

by

**Rich Perline, Robert Axtell, and Daniel Teitelbaum**

for

SBA
Office of Advocacy
www.sba.gov/advo

under contract number SBAHQ-05-Q-0018

Release Date: December 2006

SBA Office of Advocacy
www.sba.gov/advo
Small Business Research Summary
Advocacy: the voice of small business in government

December 2006 No. 285

# Volatility and Asymmetry of Small Firm Growth Rates Over Increasing Time Frames

Rich Perline, Robert Axtell, and Daniel Teitelbaum
Arlington, VA. Under contract no. SBAHQ-05-Q-0018 [21] pages.

With the emergence of new longitudinal data sets, researchers are now able to better address questions about the dynamics of businesses. This study focuses on characterizing the large dispersion of business growth rates over increasing time frames and considers whether employment expansions and contractions occur symmetrically.

## Overall Findings

Previous studies found annual business growth rates having heavy tailed distributions (many large expansions and contractions). This study found that over longer time frames, the distributions slowly move towards lighter tails (fewer large growth changes). This slow change in the distributions indicates firms tend to maintain their employment change trend.

The study also found evidence to support the belief that a systematic difference between job creation and job destruction exists. There are more large employment swings among shrinking than expanding businesses.

## Highlights

- Business growth rates were found to be reasonably well approximated by the asymmetric Subbotin distribution, which is a flexible statistical distribution with parameters that allow it to vary from very light-tailed to very-heavy tailed. The observed business growth rates show a tendency over increasing time periods (from one to five years) towards slightly more asymmetry and slightly lighter tails. In general, negative rates appear more volatile than positive rates.
- The above highlight points out the importance of firms being able to weather difficult periods and survive negative growth shocks.
- Given that job creation and destruction result from different processes, accurately modeling positive and negative growth rates necessitates using a more flexible statistical distribution that can accommodate the observed lack of symmetry.
- Matching previous research, the variance of growth rates was fairly independent of business size.
- Truncating the data by eliminating very large expansions and contractions produced differing results. Researchers need to realize that excluding what they believe to be outliers or unimportant groups of businesses can have a big impact on their results.
- Survival rates over the five-year period were similar for the different size classes, except the smallest. The survival rate for establishments with 4–7 employees was 75.3 percent; this rate slowly rose to 84.2 percent for the 512–1,023 size class. Even small size classes had relatively high five year survival rates, 61.4 percent for one-person establishments and 70.1 percent establishments with two to three persons.
- The report shows the value in utilizing non-publicly available microdata by creating special tabulations to answer important industrial organization questions.

## Scope and Methodology

The researchers utilized special tabulations from the Census Bureau's Statistics of U.S. Businesses (SUSB). (The Office of Advocacy is a partial funder of SUSB.) SUSB includes nearly all employer establishments in the United States, except farms. Firm-establishment identifiers exist in the data to create firm data. But establishments were used as a proxy for firms because of the difficulties in following firm mergers, spin-offs, and ownership changes in the underlying microdata.

Establishments (or business locations) surviving from 1998 to 2003 were the basis for the study. Establishments that opened and closed during the period were not included in the analysis.

Establishments were divided into equally spaced log employment growth rate bins by start year, and establishment employment size for differing time frames. The distribution of the frequency of the logged growth bins were charted and compared.

Given that small firm establishments represent the bulk of the data, the overall results reflect small firm growth.

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 with the U.S. Small Business Administration's Office of Advocacy are available on the Internet at ***www.sba.gov/advo/research***. Copies are available for purchase from:

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
(800) 553-6847 or (703) 605-6000
TDD: (703) 487-4639
*www.ntis.gov*

Order number: PB2006-114720
Pricing Information:
Paper copy A03 ($26.50)
CD-ROM A00 ($20.00)
Download A0 ($15.95)

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 of the Report*

The availability of U.S. Census Department data on the entire universe of U.S. businesses has begun to have a profound impact on what is known about the population of firms operating in our country. Since these data are not a sample, but a comprehensive inventory of U.S. firms, they provide heretofore unknown depth and breadth to the field of empirical industrial organization. The data have already begun to reshape certain basic ideas about firm organization and dynamics.

In this report we have summarized our research into the nature of firm growth in the U.S. over the 5 year period 1998-2003 using comprehensive data on U.S. businesses extracted from the Census database. We have analyzed these data with a particular eye to the departures from standard assumptions and results in economics generally, and within the field of industrial organization specifically. It should be noted that our analysis is unusual in an important way compared to most analyses of firm growth rates: the typical size of establishments in the Census database is preponderantly very small - the observed modal size is a single employee. More commonly, growth rate studies use size thresholds that are well above this, and as a result, the conclusions drawn from such truncated data may be incomplete or distorted.

Our first main result concerns how heavy-tailed growth rates systematically change when the time period for tracking growth changes. Given the limited five year duration of annual firm cross-sectional data to work with, our results are consistent - though only weakly - with a reduction in the 'heavy-tailedness' of growth rate distributions over time. We analyzed this time trend by modeling the empirical distributions with the Subbotin family of distribution functions. The Subbotin family specializes from very heavy-tailed distributions to the light-tailed normal distribution for particular values of a shape parameter. We find that this parameter changes - but very slowly - as we move from annual to two year data, and keeps changing - but still slowly - up to five year growth rates, the limit of our data. The parameter moves in a monotone fashion over time in the direction of the normal distribution, yet it stops considerably short of the normal distribution even in our five year window. Why this trend towards normality should occur so slowly raises interesting questions.

Our second main result concerns the asymmetry between firm growth and firm shrinkage, i.e., between positive and negative firm growth. Earlier time series studies have suggested a systematic difference between job creation and job destruction processes. Here we quantify this independently by noting the asymmetry in the growth rate distribution. Once again, given that most of the firms in our data are small businesses, we are able to draw a conclusion for how small business firms are run. Specifically, given that relatively more small firms experience negative growth of any particular magnitude than those that experience comparable positive growth, it would seem that successful small firms—including those that are on their way to becoming mid-size or larger firms—must have a way to grow steadily and avoid negative growth episodes. Firm size shrinkage is *not* just an inverse event from firm size increase, and firms that regularly flirt with both sides of the distribution would appear to play a risky game. While size fluctuations are difficult to control, our conclusions on this issue would be more clear if we had access to growth autocorrelation data. Suffice it to say here that policies aimed at helping small (and perhaps shrinking) businesses through difficult times might go far to increase the pool of successful small businesses.

Overall, our study has shed new light on old topics of firm growth, particularly with respect to small businesses.

# I. Introduction

In the past few years the availability of data on the entire universe of U.S. businesses has begun to have a profound impact on what is known about the population of firms operating in our country. Since these data are not a sample, but a comprehensive inventory of U.S. firms, it provides heretofore unknown depth and breadth to the field of empirical industrial organization. These data have already begun to reshape certain basic ideas about firm organization and dynamics.

Specifically, since the time of Gibrat [1931] it was thought that the lognormal distribution described the size distribution of firms. That is, log firm sizes were normally distributed. It was well-known that the Pareto distribution closely approximated the upper tail of the (mostly publicly-held) corporations that make up the largest firms doing business in the U.S. (Simon and Ijiri [1977]). However, the relatively small number of such firms (< 10,000) made this an unrepresentative sample of U.S. businesses as a whole, especially small businesses. Indeed, there were nearly 6 million firms having at least 1 employee in the U.S. in the late 1990s, most of which were small businesses.

In order to better understand the relation between the small number of large, public firms and the large population of small firms, Axtell [2001] used the data compiled by the Census Department from the so-called Business Master File (BMF) for 1997 and showed that the distribution of firm sizes is well-approximated by the Pareto distribution with exponent near unity—the so-called Zipf distribution—over the entire range of firm sizes, from the million plus firms with single employee to the single firm with nearly one million employees.

Subsequently, using this same comprehensive database, Teitelbaum and Axtell [2005] investigated the fluctuations in firm sizes—firm 'growth rates'—over the period 1998-1999. The conventional view of such growth rates was that their logarithm should be normally distributed, with mean near zero. However, based on the paper of Stanley *et al.* [1996], there was some evidence that once one controls for size, log firm growth rates are Laplace distributed—also known as the double exponential distribution—instead of being normally distributed. This was a particularly interesting finding insofar as the Laplace distribution is relatively heavy-tailed in comparison to the Gaussian, and given that small firms dominate these data, most of the probability mass in the tail of the Laplace distribution would be due to small firms experiencing either rapid growth or rapid decline. The Laplace distribution of growth rates has been shown to be robust to how firm size, and thus growth, is defined, e.g., size based on sales volumes of corporations (e.g., Stanley *et al.* [1996]; Botazzi, Cefis and Dosi [2002]). Teitelbaum and Axtell [2005] showed that heavy-tailed growth rate distributions also obtain intra-industry. Sales volumes of newly developed pharmaceutical products have recently been shown to be Laplace distributed (Fu *et al*. [2005]). These broad findings suggest that there may exist a general explanation for the genesis of the Laplace distribution.

This report, which is based on U.S. Census Department data for businesses, aims to analyze the establishment growth rate data in relation to the following issues:

1. How do heavy-tailed growth rate distributions change with the period of observation? That is, following the same firms over times longer than a single year, do log firm growth rates retain their heavy-tailed character or, like finance data on returns, do they become more Gaussian over time?

2. Fu *et al*. [2005] have given empirical results indicating that the Laplace distribution proposed by Stanley et al [1996] as a model of log growth rates is often not an adequate fit to the data in the upper and lower tails. They observed that for many of their examples, the Laplace fits the central part of the data, but a distribution with still heavier tails is required to fit the ends. Similarly, Teitelbaum and Axtell [2005] found that the Subbotin distribution—a generalization of both the Laplace and the normal distribution—provided a better fit to some of the Census growth rate data than the Laplace.

3. A careful look at the data presented by various investigators often reveals some asymmetry in the distribution of log growth rates. For example, most of the histograms displayed in Figures 3 and 4 of Bottazzi *et al.* [2002] exhibit an obvious asymmetry, as does Figure 3c of Havemann *et al.* [2005]. We will show that this characteristic is also very evident in almost all of the Census establishment growth data we analyze here, with a particular kind of asymmetry present across firm types.

## II. Descriptive Statistics of Firm Growth

Our raw data give the number of continuing establishments broken down by establishment size, where size is measured by number of employees, starting in 1998 and still existing in the years 1999, 2000, 2001, 2002, or 2003. An establishment is a physical location where work is conducted. Typically a large company will have many establishments. Taking WalMart as an illustration, each of its thousands of stores is regarded as a separate establishment, and every one of them will have its own unique entry in the Census database.[1]

[1] For technical reasons, having to do with the difficulties of tracking establishment ownership, the establishment data are only aggregated into enterprise (firm) data annually, but without direct comparability of enterprises from year to year. That is, an establishment may change ownership and hence the enterprise to which it belongs, but this fact is not systematically accounted for in the data. Because we are interested in firm level data as well as establishment data, we investigated applying a correction to establishment level tabulations to build up statistical firms

Owing to confidentiality issues, the data have been given to us in summary form showing only binned counts. This imposes various limitations on our analyses. Our original data tables were structured as follows (we are using the 1998-2002 time span as an example for Table 1):

| **Log Growth Rate Interval** | | **1998 Employment Size of Establishment** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Start Interval (Inclusive)** | **End Interval (Exclusive)** | **Total** | **1** | **2-3** | **4-7** | **8-15** | **…** | **2,048-4,095** | **4,096-8,191** | **…** |
| **…** | **-2.0790** | **34,871** | **0** | **0** | **0** | **12,939** | **…** | **62** | **23** | **…** |
| **-2.0790** | **-2.0405** | **380** | **0** | **0** | **0** | **0** | **…** | **2** | **0** | **…** |
| **-2.0405** | **-2.0020** | **1,338** | **0** | **0** | **0** | **633** | **…** | **0** | **0** | **…** |
| **-2.0020** | **-1.9635** | **769** | **0** | **0** | **0** | **0** | **…** | **0** | **1** | **…** |
| **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** |
| **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** |
| **-0.0770** | **-0.0385** | **57,893** | **0** | **0** | **0** | **9,744** | **…** | **62** | **15** | **…** |
| **-0.0385** | **0.0000** | **22,395** | **0** | **0** | **0** | **0** | **…** | **65** | **21** | **…** |
| **0.0000** | **0.0385** | **1,050,254** | **367,826** | **327,335** | **208,781** | **86,090** | **…** | **126** | **35** | **…** |
| **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** |
| **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** | **…** |
| **2.0020** | **2.0405** | **1,416** | **0** | **846** | **291** | **162** | **…** | **0** | **0** | **…** |
| **2.0405** | **2.0790** | **441** | **0** | **0** | **195** | **143** | **…** | **0** | **0** | **…** |
| **2.0790** | **…** | **32,664** | **18,017** | **7,395** | **4,050** | **1,840** | **…** | **0** | **0** | **…** |
| **Total** | | **4,327,403** | **606,501** | **926,955** | **1,061,980** | **797,702** | **…** | **1,290** | **360** | **…** |
| **Mean of the Change Rate** | | **0.0214** | **0.4319** | **0.0844** | **-0.0316** | **-0.0849** | **…** | **-0.2884** | **-0.3953** | **…** |
| **Variance of the Change Rate** | | **0.4311** | **0.4237** | **0.3905** | **0.3614** | **0.3612** | **…** | **0.9851** | **1.5734** | **…** |

**Table 1.** Part of the binned Census data showing the log growth rates of establishments between 1998 and 2002. The data are binned by 1) size in 1998; and 2) log growth rate between 1998 and 2002.

The data of Table 1 are binned by equally spaced log growth rates in increments of .0385. All log growth rates <= -2.0790 and >= 2.0793 are grouped together in the first and last bins. The growth rate is defined as $G =$ (Number of Employees at Time 2 / Number of Employees in 1998). Time 2 can be any of the 5 years 1999, 2000, 2001, 2002 or 2003. Log growth rates are defined as $g = \ln(G)$ and we use the natural log throughout. <u>Note that in the cases where an establishment had 0 employees in Time 2, it is excluded from that table.</u> <u>It should be kept in mind that the calculations of means and variances of *g* do not take into account these most extreme cases of shrinking size.</u>

The endpoints of the intervals of the table ($g = \pm 2.079$) correspond to actual growth rates $G$ of $e^{-2.079} \approx 1/8$ and $e^{2.079} \approx 8.0$. Therefore, in this example, we see from the table that there were 12,939 establishments in the size range 8-15 in 1998 that by 2002 had decreased in size to no more than 1/8th their original size. (In

that would approximate their distributional properties. For instance, we know the distribution of establishments per enterprise unconditionally, as well as conditional on firm size. Not surprisingly, because these data are numerically dominated by small firms, our results turned out to be insensitive to this correction.

this case, it can be deduced that all of these establishments had to be exactly of size 1 in 2002 - they could not be of size 0 in 2002 or they would not appear in the table and they could not be of size > 1 in 2002 because than their growth rates would have to have exceeded the value of 1/8.) Similarly, we also see from Table 1 that 7,395 establishments of size 2-3 in 1998 exhibited at least an 8-fold increase by 2002. In this instance, we cannot conclude more than this -- some of these establishments could have had much more than an 8-fold increase, but we do not have the exact value.

Besides the binned log growth rates, establishments are binned by their 1998 sizes. These size bin intervals are geometrically spaced (sizes 1, 2-3, 4-7, 8-15, etc.), and therefore prevent some fine grained analyses that would be informative. Also, the data are very sparse and therefore not very useful extending beyond the size interval of 4,096-8,192 employees. Other limitations in the data will be pointed out below.

Tables 2(a) and (b) show the mean values of $g$ for establishment sizes in 1998 with size ranges up to the 2048-4095 for each of the 5 time intervals. We give two versions of these means. In the top Table 2(a), the means are calculated using all the non-zero observations, including those in the first growth rate bin $(g \leq -2.079)$ and the last $((g >= 2.079)$. In the bottom Table 2(b), we give the means after excluding these two end intervals. The reason for showing both sets of values is that we believe a small number of establishments - a subset of those that decreased in size so much that their values of g were $\leq -2.079$ - will have an excessive effect on the overall mean values. For example, consider establishments in the size range 4096-8195 in 1998, some of which might have shrunk as far as to size 10 or 20, say, in a subsequent year. These huge drops in size for just a few establishments could drag the overall mean down considerably. By excluding these outliers, we aim for a more robust estimate. The differences between the two estimates are often quite large. For example, the overall mean for the size range 2048-4095 in the 1998-2002 time interval was from Table 2(a) was -.288, but after excluding the extreme bins, the mean value of the truncated observations increases to -.103, as seen in Table 2(b).

| Size | 98-99 means | 98-00 means | 98-01 means | 98-02 means | 98-03 means |
|---|---|---|---|---|---|
| 1 | 0.222 | 0.307 | 0.357 | 0.432 | 0.475 |
| 2-3 | 0.013 | 0.043 | 0.057 | 0.084 | 0.101 |
| 4-7 | -0.035 | -0.025 | -0.028 | -0.032 | -0.026 |
| 8-15 | -0.050 | -0.047 | -0.060 | -0.085 | -0.081 |
| 16-31 | -0.054 | -0.054 | -0.074 | -0.116 | -0.114 |
| 32-63 | -0.059 | -0.069 | -0.095 | -0.159 | -0.155 |
| 64-127 | -0.066 | -0.085 | -0.124 | -0.205 | -0.201 |
| 128-255 | -0.083 | -0.114 | -0.166 | -0.251 | -0.261 |
| 256-511 | -0.103 | -0.147 | -0.205 | -0.309 | -0.328 |
| 512-1023 | -0.109 | -0.155 | -0.229 | -0.352 | -0.395 |
| 1024-2047 | -0.089 | -0.154 | -0.218 | -0.312 | -0.342 |
| 2048-4095 | -0.074 | -0.141 | -0.173 | -0.288 | -0.347 |
| All | 0.009 | 0.027 | 0.026 | 0.021 | 0.030 |

**Table 2(a) Means including all non-zero values.**

| | | | | | |
|---|---|---|---|---|---|
| 1 | 0.210 | 0.280 | 0.318 | 0.381 | 0.410 |
| 2-3 | 0.014 | 0.038 | 0.047 | 0.070 | 0.081 |
| 4-7 | -0.033 | -0.027 | -0.033 | -0.039 | -0.036 |
| 8-15 | -0.029 | -0.024 | -0.031 | -0.053 | -0.050 |
| 16-31 | -0.029 | -0.025 | -0.034 | -0.072 | -0.070 |
| 32-63 | -0.030 | -0.029 | -0.044 | -0.097 | -0.095 |
| 64-127 | -0.031 | -0.036 | -0.055 | -0.118 | -0.114 |
| 128-255 | -0.042 | -0.051 | -0.079 | -0.143 | -0.149 |
| 256-511 | -0.049 | -0.061 | -0.093 | -0.158 | -0.171 |
| 512-1023 | -0.052 | -0.066 | -0.098 | -0.168 | -0.184 |
| 1024-2047 | -0.034 | -0.044 | -0.060 | -0.115 | -0.137 |
| 2048-4095 | -0.021 | -0.023 | -0.042 | -0.103 | -0.124 |
| All | 0.018 | 0.034 | 0.033 | 0.029 | 0.034 |

**Table 2(b). Means excluding first and last growth rate bins.**

**Table 2**. Mean values of log growth rates ($g$) by 1998 establishment size and the 5 time intervals under study. Two versions of the calculations are given. The top table shows means calculated from all non-zero values in the data. The bottom table excludes the first bin where values of $g \le -2.079$ and the last bin where $g \ge 2.079$.

The most obvious trend visible in Table 2 in both versions of the calculated means is that for a given size range, <u>except for the smallest size ranges 1, 2-3 and 4-7,</u> "negative growth" (shrinking size) is observed as the time interval increases. It can also be seen that for a given time interval, growth becomes more negative as size increases up to the size ranges 1024-2047 and 2048-4095, at which point there is a slight reversal in the trend.

The most exceptional situation in Table 2 is where we see *increasing* average values of $g$ over time for the smallest size establishments (1, 2-3). Here there is actually positive growth, on average, over the longer time intervals. Recognize, however, that the estimates in both tables are generally upwardly biased because they have excluded establishments that vanish from the database, and this is most likely to affect the very smallest establishments. Although we do not have counts by specific size ranges of the number of establishments rolling off the database when their size decreased to 0, we can get some idea of their magnitudes from Table 3.

In 1998, there were 6,187,599 establishments on the Census database. Referring to Table 3, there were only 5,534,708 establishments that were in the database in both years 1998 and 1999 and were used for growth calculations. This is already a loss of 11.5% establishments in a single year. Looking over the entire time span of the study, the total number of establishments available for analysis dropped to 4,055,605, so that approximately 1/3rd of the 1998 establishments were lost in the 5 year period. Of course, these losses occur across all size groups, but the size groups of 1 and 2-3 are more likely to have dropped to size 0 in a one year interval than are the larger establishments. Therefore, it is clear that the estimates of the means of $g$ in Table 2 (for both tables 2(a) and 2(b)) are biased since both excluded the establishments that went to size 0.[2]

| Time Span | Total | 1 | 2-3 | 4-7 | 8-15 | 16-31 | 32-63 | 64-127 | 128-255 | 256-511 | 512-1,023 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 98-99 | 5,534,708 | 898,428 | 1,229,445 | 1,326,681 | 967,686 | 552,412 | 295,815 | 148,206 | 71,435 | 28,076 | 10,477 |
| 98-00 | 5,081,020 | 777,503 | 1,114,955 | 1,230,766 | 908,016 | 520,143 | 278,960 | 140,006 | 67,908 | 26,836 | 10,054 |
| 98-01 | 4,703,611 | 688,203 | 1,019,459 | 1,147,773 | 854,023 | 491,410 | 264,154 | 132,676 | 64,804 | 25,721 | 9,707 |
| 98-02 | 4,327,403 | 606,501 | 926,955 | 1,061,980 | 797,702 | 461,003 | 248,332 | 124,786 | 61,156 | 24,246 | 9,245 |
| 98-03 | 4,055,605 | 551,848 | 861,346 | 998,490 | 755,449 | 438,481 | 235,920 | 118,452 | 58,302 | 23,161 | 8,825 |
| % Survivors from 1999 | 73.3% | 61.4% | 70.1% | 75.3% | 78.1% | 79.4% | 79.8% | 79.9% | 81.6% | 82.5% | 84.2% |

**Table 3.** This table shows the total number of establishments in the growth rate database for the five time intervals being analyzed. The counts are further broken down for selected establishment sizes (as of 1998) from 1 employee through the size range 512-1023. The decreases in total counts over time occur as establishments fall off the database (i.e., go to size=0).

[2]Our survivor rates are remarkably consistent with an old study by Popkin [1991], who reported on survivorship in the 1976-1986 period. As just one example, for firms of size 20-49 in 1976, he reported 12% lost in the two year 1976-1978 interval. Compare that with our establishments of size 16-31 in 1999, where the non-survivor rate in the period 1999-2001 was 11%, as calculated from Table 3 (= 1- 491,410/552,412).

We next look at the variances of $g$, broken down as above by size ranges and time intervals as displayed in Tables 4(a) and (b).

| 1998 Size | 98-99 var | 98-00 var | 98-01 var | 98-02 var | 98-03 var |
|---|---|---|---|---|---|
| 1 | 0.232 | 0.322 | 0.373 | 0.424 | 0.472 |
| 2-3 | 0.222 | 0.300 | 0.334 | 0.391 | 0.433 |
| 4-7 | 0.185 | 0.261 | 0.367 | 0.361 | 0.398 |
| 8-15 | 0.178 | 0.249 | 0.313 | 0.361 | 0.391 |
| 16-31 | 0.187 | 0.263 | 0.334 | 0.386 | 0.412 |
| 32-63 | 0.201 | 0.290 | 0.367 | 0.440 | 0.458 |
| 64-127 | 0.222 | 0.326 | 0.434 | 0.526 | 0.548 |
| 128-255 | 0.242 | 0.378 | 0.505 | 0.613 | 0.645 |
| 256-511 | 0.302 | 0.492 | 0.635 | 0.805 | 0.840 |
| 512-1023 | 0.312 | 0.497 | 0.712 | 0.952 | 1.067 |
| 1024-2047 | 0.317 | 0.628 | 0.862 | 1.074 | 1.108 |
| 2048-4095 | 0.291 | 0.689 | 0.697 | 0.985 | 1.104 |
| All | 0.214 | 0.300 | 0.367 | 0.431 | 0.472 |

**Table 4(a) Untruncated Variances**

| | | | | | |
|---|---|---|---|---|---|
| 1 | 0.160 | 0.210 | 0.236 | 0.265 | 0.284 |
| 2-3 | 0.205 | 0.270 | 0.311 | 0.342 | 0.373 |
| 4-7 | 0.182 | 0.246 | 0.296 | 0.337 | 0.369 |
| 8-15 | 0.129 | 0.185 | 0.228 | 0.269 | 0.294 |
| 16-31 | 0.119 | 0.172 | 0.212 | 0.254 | 0.278 |
| 32-63 | 0.113 | 0.165 | 0.202 | 0.248 | 0.270 |
| 64-127 | 0.107 | 0.162 | 0.199 | 0.248 | 0.270 |
| 128-255 | 0.099 | 0.153 | 0.194 | 0.241 | 0.263 |
| 256-511 | 0.103 | 0.159 | 0.205 | 0.257 | 0.281 |
| 512-1023 | 0.100 | 0.144 | 0.190 | 0.240 | 0.261 |
| 1024-2047 | 0.079 | 0.128 | 0.164 | 0.217 | 0.238 |
| 2048-4095 | 0.076 | 0.125 | 0.147 | 0.205 | 0.226 |
| All | 0.168 | 0.230 | 0.274 | 0.320 | 0.349 |

**Table 4(b) Truncated Variances**

**Table 4.** Variances of log growth rates ($g$) by 1998 establishment size and the 5 time intervals under study. As in Table 2, two versions of the calculations are given. The top table shows variances calculated from all non-zero values in the data. The bottom table shows the variances after excluding the first bin where values of $g \leq -2.079$ and the last bin where $g \geq 2.079$.

The story in Table 4 is a bit more complicated than that in Table 2, but as with the calculations of means, we have displayed the variance calculations in two versions, one without excluding the bottom and top bins (Table 4(a) ) and the other that excludes them (Table 4(b)). The first clear trend, visible in both the (a) and (b) tables, is that for a given size range, there is an increase in the variance of $g$ over time. This is typical and to be expected of most random processes, since the longer the interval, the more opportunities there are for variation to occur.

However, there is another trend present in Table 4(a) but not in 4(b). In 4(a), for a given time interval and ignoring the first three size ranges (1,2-3,4-7), the variances generally increase with size. This pattern is quite consistent and holds true for each of the 5 time intervals (see also Figure 1(a)). The same pattern was noted by Teitelbaum and Axtell (2005) in their analysis of the 1998-99 Census growth data.

This result is the *opposite* of what was reported in Stanley et al [1996] in their study of the growth rates of publicly traded manufacturing firms. They observed a strong regularity in which the standard deviations of $g$ *decreased* systematically in relation to size ($S$) approximately conforming to the log-log relationship $\log(\sigma(g)) \approx \beta * \log(S)$, with $\beta \approx -.16$. In fact, this relationship held approximately true when $S$ was measured either in terms of sales volume or by the number of employees.

However, there have also been some reports *not* finding this effect. For example, Bottazzi Cefis and Dosi (2002), although aware of Stanley et al's result, did not find this pattern in their analysis of a large longitudinal sample of Italian manufacturing firms. In their study, the variances of $g$ were essentially flat with respect to firm sizes.

Our Table 4 gives variances $\sigma^2(g)$, not standard deviations $\sigma(g)$, but of course in terms of Stanley et al's log-log linear relationship, we would see $\log(\sigma^2(g)) = 2 * \log(\sigma(g)) = 2 * \beta * \log(S)$ with $\beta < 0$. However, from our Table 4(a) and Figure 1(a), we clearly see a positive relationship (again, ignoring the two smallest size groups of 1 and 2-3) between $\log(\sigma^2(g))$ and $\log(S)$.

How can we account for this difference between the studies? As we have remarked, one significant factor that distinguishes our growth analysis here from all the published reports we have seen is the distribution of the sizes of the entities being analyzed. Our analysis is quite unusual in that the modal size category of establishments is size =1. The published reports that have found the general log-log relationship between $\ln(\sigma^2(g))$ or $\ln(\sigma(g))$ and $\log(S)$ seem to have been based on distributions with a relatively high threshold such that the size in terms of number employees is never less than, say, 10-20. In our census data, roughly 80% of the establishments have sizes under 15 employees and 15%-20% have only a single employee.

This high concentration of small establishments bears significantly on the relationship between $\sigma^2(g)$ and size $S$. In particular, it further highlights the issue of establishments (or corporations in other studies) that exit from a longitudinal database because they drop to size 0. We pointed out in connection with Table 3 above that 1/3rd of the establishments existing in 1998 disappeared by 2003. Dropouts are especially high in our data because there are so many small sized establishments. We believe that this problem has been generally ignored in most studies. The obvious difficulty is that analyses using g=ln(*G)* as the growth metric, which seems quite natural for most purposes, have to exclude cases where *G*=0. That is, if the size $S_2$ at time 2 drops to 0, then $G = S_2 / S_1 = 0$, and so the observation cannot be included in the calculation of a mean or variance of *g*.

The published reports discussing the finding $\log(\sigma(g)) \approx \beta * \log(S)$ have not addressed this issue in any detail. For example, Stanley et al [1996] commented in the text of their Figure 1 that they ignored some data "because between 1992 and 1994 there are several companies with zero sales." Yet they probably encountered dropouts in other years, as well. Similarly, Sutton's (2002) analysis of the variances of growth rates ignores the issue. In any event, to try to understand why our own results do not exhibit this relationship, we have looked at the variance calculations both with and without including the lowest and highest bins, as displayed in Table 4. Figure 1, below, is a side-by-side plot of these two versions of variance estimates for ease of comparison.

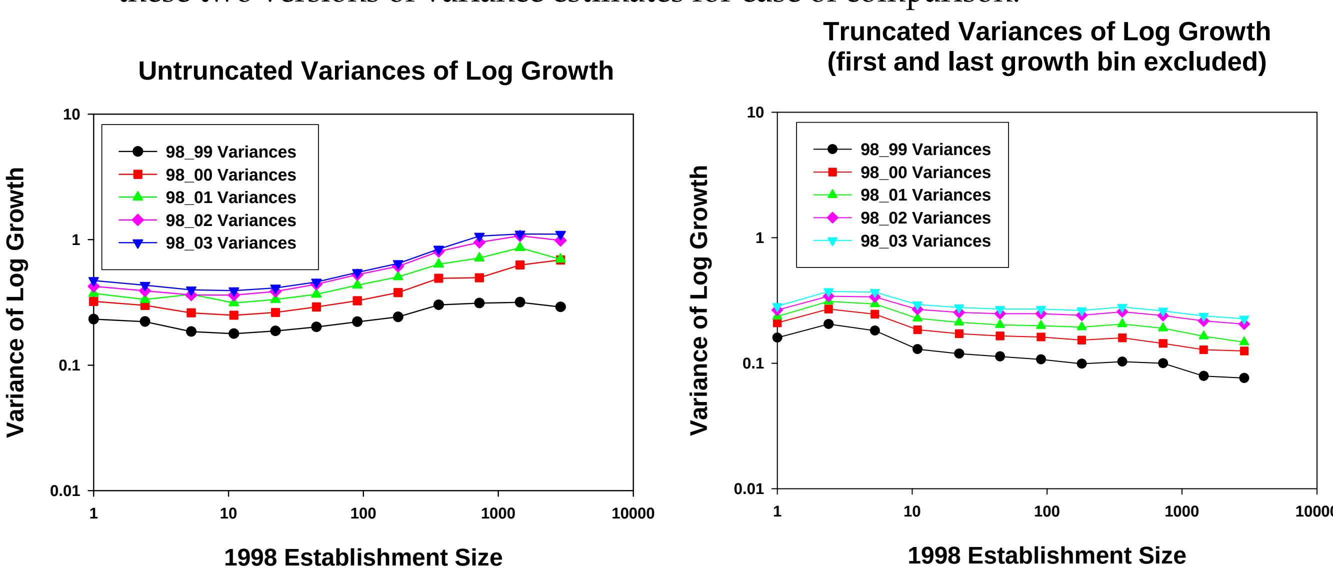



**Figure 1**. The points in the left graph are the values given in Table 4(a) and show calculated variance values that include the first and last growth bins. The graph on the right plots the values in Table 4(b), which excludes the two end bins in the calculation.

The variance values in the graph on the left, which are the "untruncated" calculations that include the lowest growth bin ($g \leq -2.079$) and the highest ($g >= 2.079$), after ignoring the smallest size groups, exhibit a generally increasing relationship between $\log(\sigma^2(g))$ and $\log(S)$ in our data. The variances in the right graph (the "truncated" versions) were calculated after excluding the two end growth bins. In this case, we *do* see the monotonically, roughly log-log linear *decreasing* association reported by Stanley et al. It may be, then, that Stanley et al's finding are connected with how outlier values are handled in the variance calculation.

It is worth emphasizing again, as was stated in Teitelbaum and Axtell [2005], that perhaps the main general finding is that the relationship between $\log(\sigma^2(g))$ and $\log(S)$ is quite *flat*, irrespective of whether there is a small positive or small negative association. This flatness is by itself quite surprising and deserves to be explained.

## III. The Distribution of Firm Growth Rates Over Time

In addition to reporting the relationship between $\log(\sigma^2(g))$ and $\log(S)$ discussed in Section II, Stanley et al [1996] found that the distributions of log growth rates for the firms in their database, which might for several reasons be expected to be approximately normally (Gaussian) distributed, have far heavier tails and a much more peaked shape than would occur if the Gaussian approximation held. They showed that their log growth rates could be better fit with a Laplace distribution. Specifically, for their empirical data, the histograms of $g = \ln(G)$ were better approximated by the Laplace distribution, with probability density function (PDF) $f(g) = \frac{1}{\sigma_g\sqrt{2}} e^{-\sqrt{2}|g-\mu_g|/\sigma_g}$, than the normal PDF $f(g) = \frac{1}{\sigma_g\sqrt{2\pi}} e^{-(g-\mu_g)^2/2\sigma_g^2}$, for parameters $-\infty < \mu_g < \infty$ and $\sigma_g > 0$. However, Teitelbaum and Axtell [2005], found that their log growth rates were in several cases better fit by a symmetric exponential power distribution (Ayebo and Kozubowski [2003]), which is a generalization of *both* the normal and Laplace distributions. The PDF for the symmetric version of this distribution, sometimes also known as the Subbotin distribution (Kotz, Kozubowski and Podgorski [2003] ), is given by

$$f(g) = \frac{\alpha}{2\sigma_g\Gamma(1/\alpha)} \exp\left(-\frac{|g-\mu_g|^\alpha}{\sigma_g{}^\alpha}\right),$$

where $\Gamma(x)$ is the standard gamma function, $\alpha > 0$, $\sigma_g > 0$, and $-\infty < \mu_g < \infty$.

A plot of the logarithm of the PDF, $\log(f(g))$, against $g = \log(G)$ will produce a distinct "tent shaped" function in the Laplace case and a parabolic function in the normal case. This is clear in the Laplace case because

$$\log(f(g)) = -\sqrt{2}\,|\,g - \mu_g\,|\,/\,\sigma_g - \log\left(\sigma_g\sqrt{2}\right),$$

is a linear function of $g$ with positive slope $\sqrt{2}/\sigma_g$ when $g < \mu_g$ and a linear function with negative slope $-\sqrt{2}/\sigma_g$ when $g > \mu_g$. In the normal case,

$$\log(f(g)) = -(g - \mu_g)^2 / 2\sigma_g^2 - \log(\sigma_g\sqrt{2\pi}),$$

which is a quadratic relation between $\log(f(g))$ and $g$ producing an upside-down parabola. For the general Subbotin case,

$$\log(f(g)) = \log(\alpha) - \log(2\sigma_g\Gamma(1/\alpha)) - \frac{|\,g - \mu_g\,|^\alpha}{\sigma_g^\alpha},$$

so that the shape of a plot of $\log(f(g))$ against $\log(g)$ will depend critically on the value of the parameter $\alpha$. For $\alpha > 1$, the plot will have a roughly upside-down, parabolic appearance; for $\alpha = 2$, we obtain a Gaussian distribution; and for $\alpha > 2$, as in curve 4 of Figure 2, the shape becomes increasingly "uniform-like" as $\alpha \to \infty$. When $\alpha = 1$, the Subbotin reduces to the Laplace and when $\alpha < 1$, the PDF looks Laplace-like in its center, but has heavier tails, as in curve 3 of Figure 2.

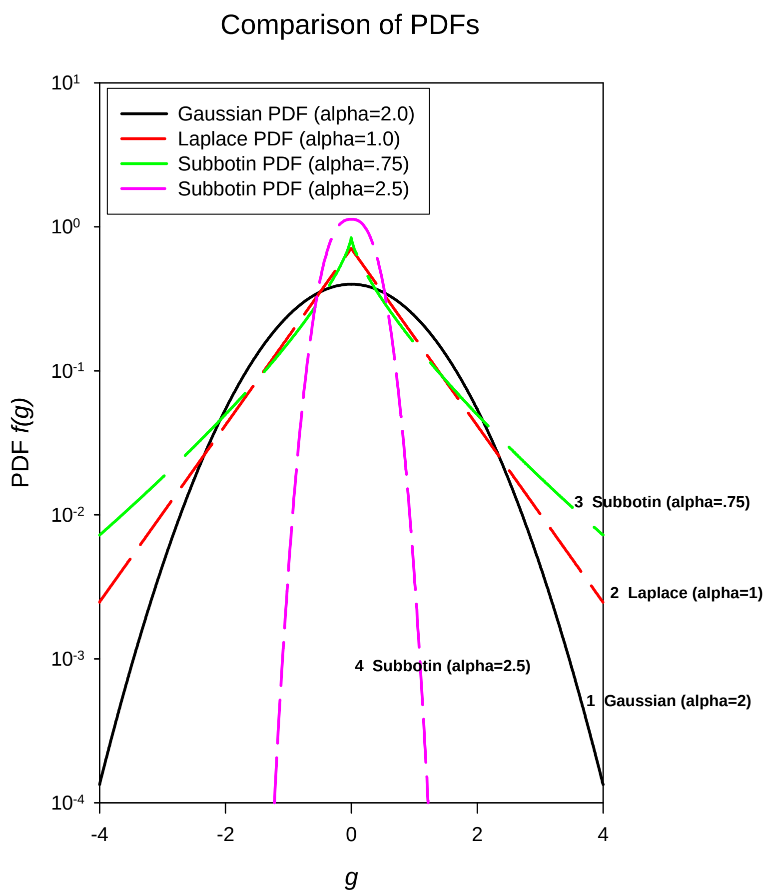



**Figure 2.** Comparing the theoretical PDFs of the Laplace, Gaussian and general Subbotin distributions. All three distributions fall into the same family, with each specific one characterized by a different value of the parameter $\alpha$ as shown.

Figure 2 clearly shows that how the tails of the Laplace, Gaussian and general Subbotin compare. The tails of the general Subbotin distribution depend critically on the parameter $\alpha$ - the distribution can be very heavy-tailed, as in curve 3 of Figure 2, or very light-tailed, as in curve 4.

The analyses of growth rates of business establishments in Teitelbaum and Axtell [2005] looked at distributions by different industry sectors using the U.S. Census Bureau database of approximately 6 million business establishments from the time period 1998-1999.[3] They examined 20 different industries, as identified by their NAICS Sector codes, and found that in every case the Laplace model gave a

---

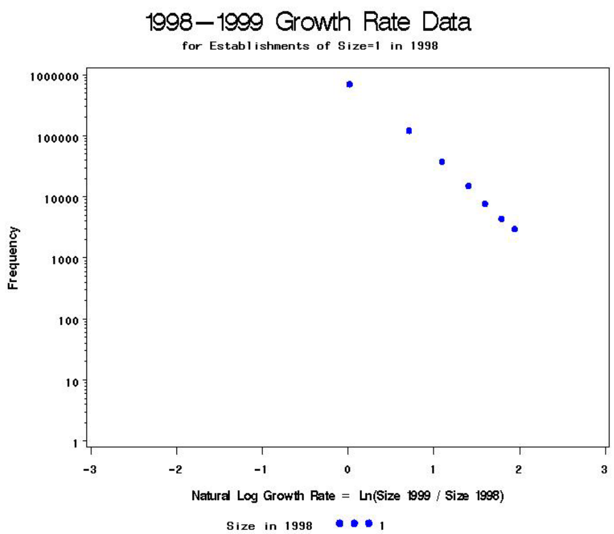
[3] The data used by Teitelbaum and Axtell [2005] was analyzed with a correction to account for multi-establishment enterprises as was done for the data analyzed here and described in our earlier footnote #1. Again, because the Teitelbaum and Axtell data distribution was so dominated by small establishments, just as with the data here, there were no distinguishable differences between the corrected and uncorrected results.

better fit than the Gaussian model. However, they also observed departures from the Laplace form and suggested that in some cases a still better fit to the data could be obtained from a more heavy-tailed Subbotin distribution.

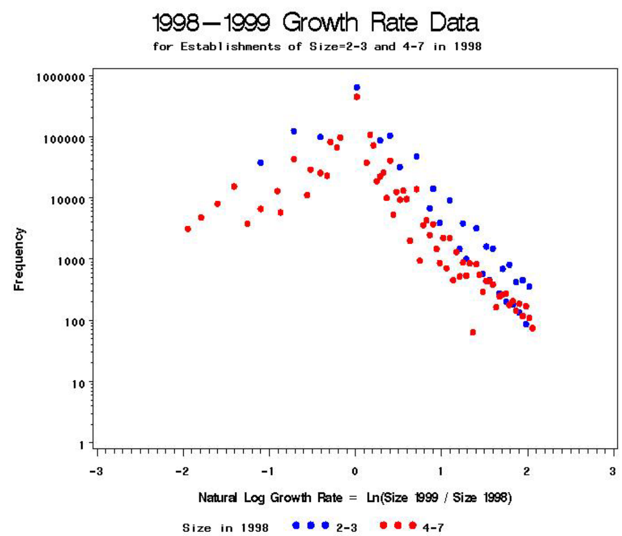
Recently, Stanley and his colleagues (Fu et al [2005]) have examined growth rate data from a number of sources and have also concluded that the Laplace model often does not accurately represent the distribution of growth rates over the full ranges of the data. They proposed a proportional growth model that leads to Laplace-like behavior in the center of the distribution and substantially heavier tails.

Several of the plots in Teitelbaum and Axtell's report (for example, Figures 9, 11 and 12), exhibited an obvious asymmetry in the log growth rate distributions. This asymmetry is again apparent in our Figure 3 below. In the three figures of Figure 3, we show the histograms of the log growth rates for establishments in the 1998-99 period. Three separate graphs are presented: Figure 3(a) is for establishments of size 1 in 1998; Figure 3(b) is for establishments of sizes 2-3, and 4-7; and Figure 3(c) is for the establishment size bins of 8-15, 16-31, 32-63, 64-127, 128-255, 256-511 and 512-1023. In Figure 3(a), since this is for establishments of size=1 in 1998, we can only observe the right arm of the histogram, which shows positive log growth values. Establishments of size 1 can only get smaller by going to 0 and exiting the database. In this case we have an obvious and explainable asymmetry in the curve. There is a similar bias operating to some extent for the left (negative log growth) side of the histograms for the small establishments shown in Figure 3(b). However, Figure 3(c), which also exhibits asymmetry between the left and right sides of the histogram for *all* the size intervals, cannot be explained away in this same way.

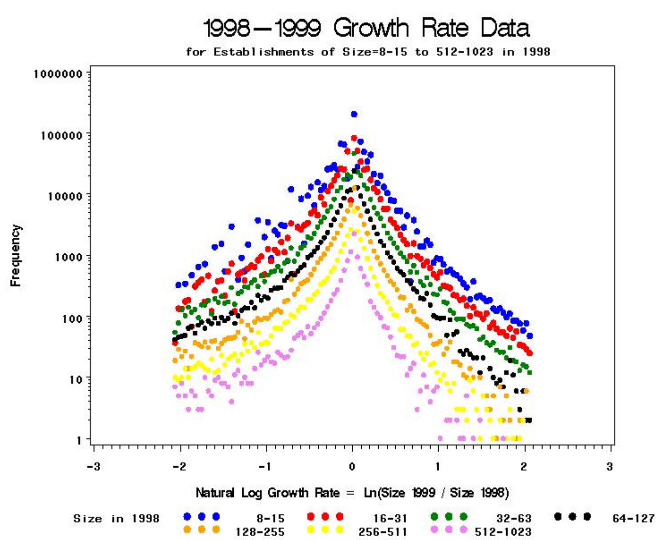
The nonlinear, concave and asymmetric shape of the histogram plots in Figure 3(c) underlines the need for a more flexible distribution than the symmetric Laplace or even the symmetric Subbotin distribution. (The nested character of the 7 curves in Figure 3(c) merely reflects the fact that there are many more business establishments of smaller size than larger.)




(a)




(b)




(c)

**Figure 3**. Fig. 3(a) shows the histogram for 98-99 growth rates for the establishments of size=1 in 1998. The only way size=1 establishments can get smaller is to shrink to size 0, in which case $g$ cannot be computed and so only a non-negative right arm is visible in the graph. A similar effect leads to shortened left arms in the histograms for establishments of sizes 2-3 and 4-7 as displayed in Figure 3(b). Figure 3(c) shows the histograms of growth rates for 98-99 for the seven size groups 8-15 through 512-1023.

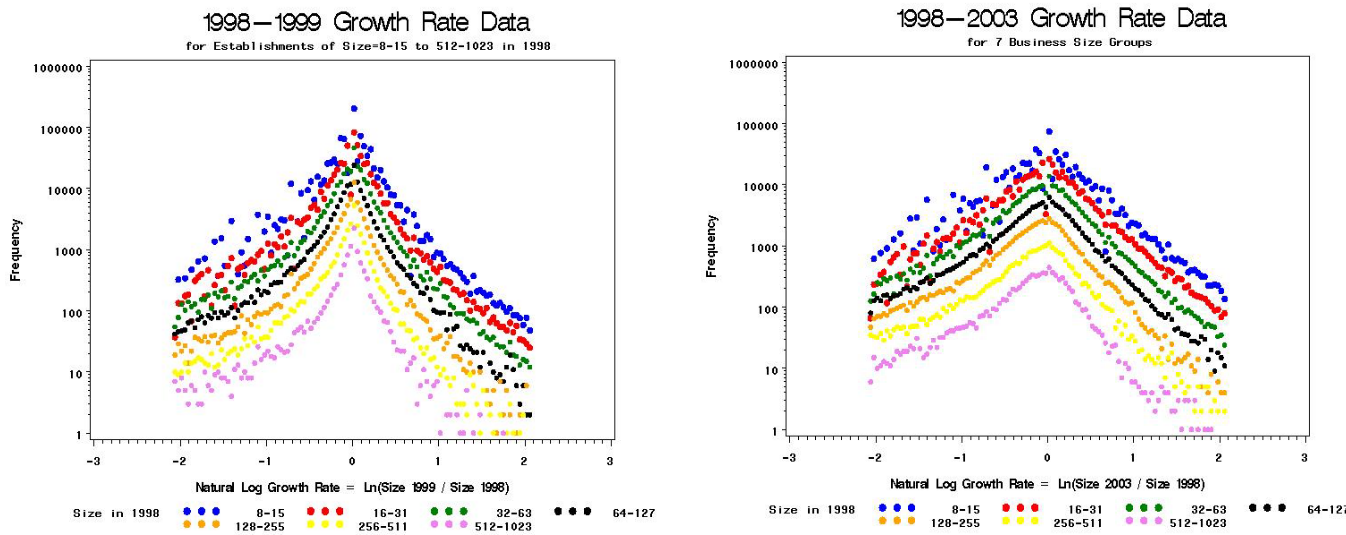



**Figure 4.** The histogram plot on the left is the same as that in Figure 3(c) - it is shown here again in order to compare it directly to the 1998-2003 histograms in the right plot. The plots on the right for the longer time interval exhibit wider arms and a somewhat less peaked center.

The plots on the right side in Figure 4 (1998-2003) have essentially the same shape as the plots on the left (1998-1999) except that the curves are less peaked and more spread out. This greater spreading out is a consequence of larger variances of log growth rates. This was seen in Table 4, which gives the variances of the log growth rates for the 5 time periods for 12 different group sizes.

Figure 5 helps to reveal this pattern still more clearly. The histogram plots on the left of Figure 5 show stacked histograms for log growth rates for the size group 256-511 that reveals the changing shape of the curves as the time intervals over which log growth rates is calculated increase. The histogram plots on the right are for the size group 512-1023. These are "stacked" histograms - the frequencies are correct for the bottom curve, but each curve above has its counts multiplied times a factor of 10 in order to see the shapes of the curves most easily.

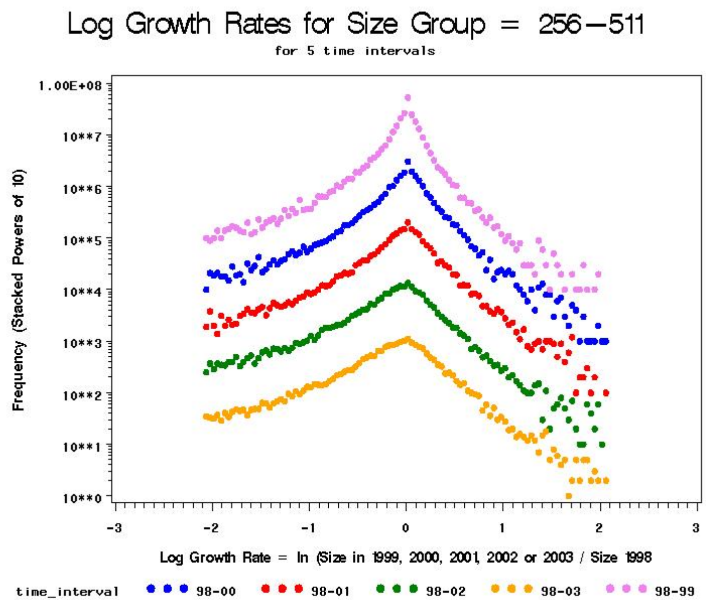



**Figure 5(a)**

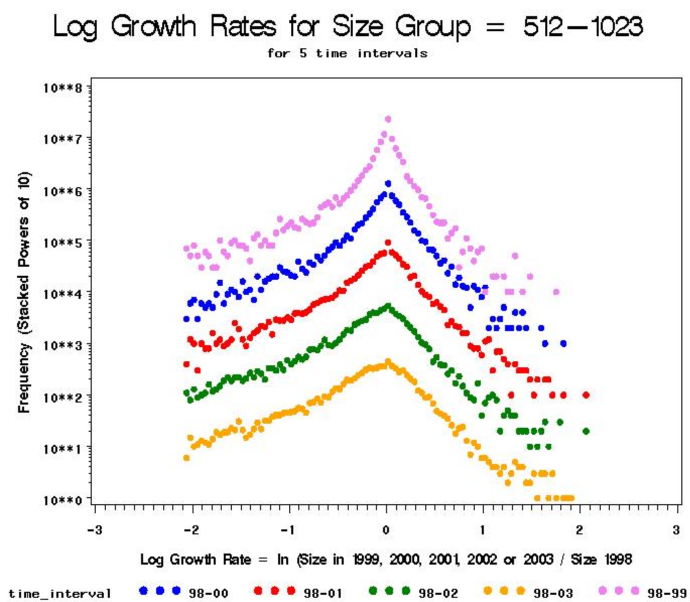



**Figure 5(b)**

**Figure 5.** The graph in Figure 5(a) shows "stacked" histograms for the size group 256-511 for each of the 5 time intervals. The graph in Figure 5(b) shows the same for the size group 512-1023. In both cases, the frequencies in the vertical axis are exact for the bottom curve, but each curve above the bottom one has frequencies multiplied by increasing powers of 10.

Because of the evident asymmetry of these log growth rate distributions, we have used the *asymmetric* form of the Subbotin distribution as a data model. The asymmetric Subbotin has the PDF (Ayebo and Kozubowski [2003]):

$$f(g)=\begin{cases}\dfrac{\alpha\kappa}{\sigma\Gamma(1/\alpha)(1+\kappa^2)}\exp(-\dfrac{\kappa^\alpha}{\sigma^\alpha}(g-\theta)^\alpha) & \text{if } g\geq\theta,\\ \dfrac{\alpha\kappa}{\sigma\Gamma(1/\alpha)(1+\kappa^2)}\exp(-\dfrac{1}{\sigma^\alpha\kappa^\alpha}|g-\theta|^\alpha) & \text{if } g\prec\theta,\end{cases}$$

which differs from the symmetric version of this PDF give above with the presence of a skewness parameter $\kappa>0$. (We are using $\theta$ instead of $\mu_g$, as we did earlier with the symmetric Subbotin, to be consistent with the notation in Ayebo and Kozubowski.) For $\kappa=1$, this reduces to the symmetric Subbotin; for $\kappa=1$ and $\alpha=1$ this is the symmetric Laplace distribution and for $\kappa=1$ and $\alpha=2$ it becomes a Gaussian distribution.

We fit this 4-parameter Subbotin distribution to various subsets of our data. The parameter estimates were obtained by a grid search that found the parameters minimizing the "distance" between the empirical data and the theoretical model as assessed by the Kolmogorov-Smirnov (K-S) statistic.

The main results of these estimations can be illustrated using the data for the establishments of size 32-63 in 1998. Figure 6 (a)-(e) shows the fitted densities superimposed on the empirical distributions for the log growth rates in the time periods 1998-99 through 1998-03.

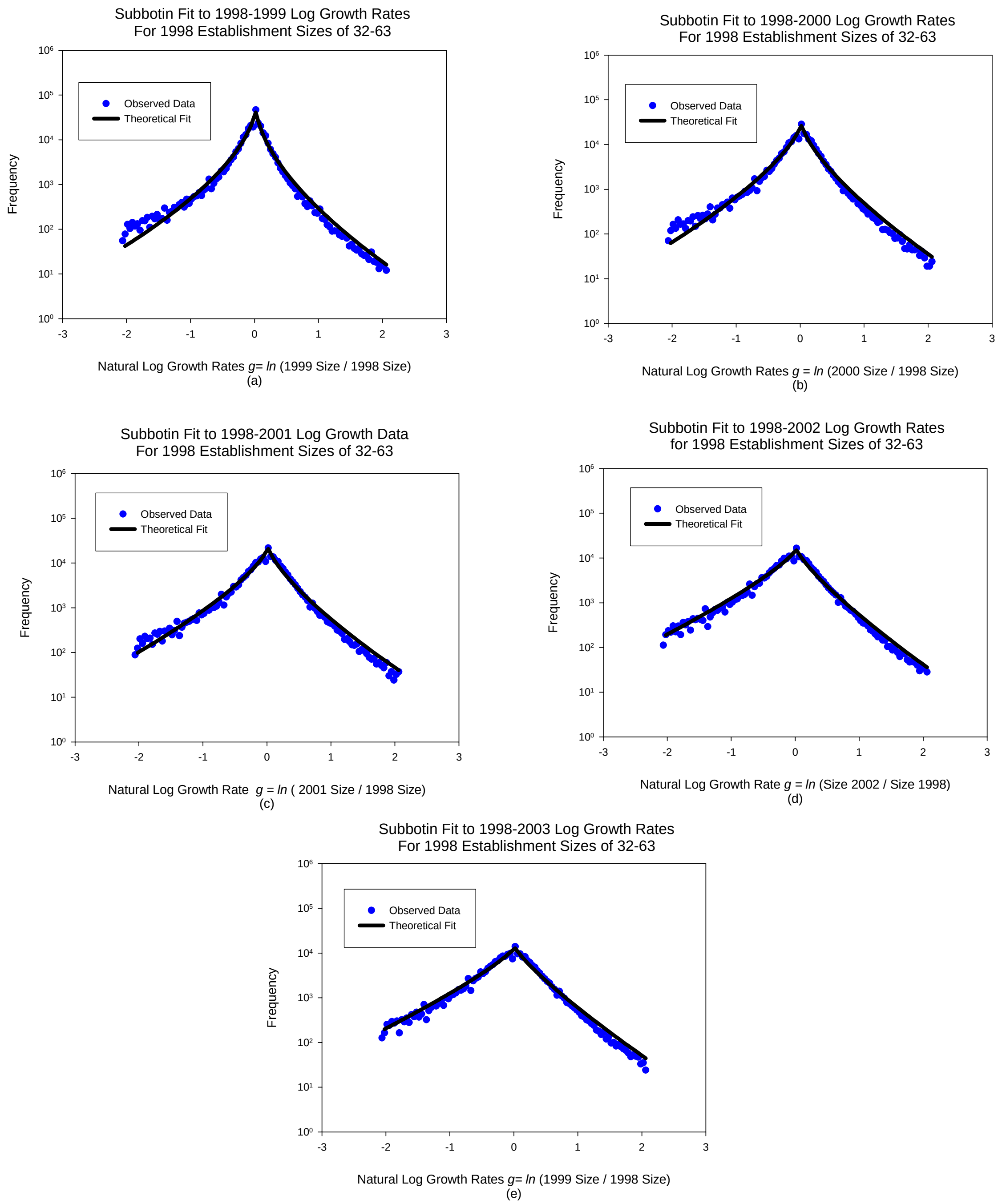



**Figure 6**. Fitting the 4-parameter Subbotin distribution to the log growth rate data for business establishments of size 32-63 in 1998. The five plots (a)-(e) show the empirical growth rate data for the periods 1998-99 through 1998-03 together with the fitted distributions. The fits become progressively better as the time interval increases. The graph in (e) exhibits the best fit relative to all the others.

The parameter estimates and the K-S and Chi Square statistics for the five fits are given in Table 5 below.

| Time Interval | Alpha | Sigma | Kappa | Theta | K_S | Chi Sq |
|---|---|---|---|---|---|---|
| 1998 - 99 | 0.60 | 0.07 | 1.11 | 0 | 0.017 | 440.8 |
| 1998 - 00 | 0.70 | 0.14 | 1.08 | 0 | 0.017 | 242.1 |
| 1998 - 01 | 0.73 | 0.18 | 1.11 | 0 | 0.017 | 187.4 |
| 1998 - 02 | 0.80 | 0.26 | 1.22 | 0 | 0.013 | 138.3 |
| 1998 - 03 | 0.83 | 0.30 | 1.20 | 0 | 0.012 | 107.4 |

**Table 5**. Parameter estimates and fit statistics for the Subbotin PDFs shown in the 5 histograms of Figure 6. Note the general trends over longer time intervals: the values of alpha and sigma increase and the fit statistics, k_s and Chi Square, improve (decrease).

Inspection of Table 5 reveals a definite trend in the Subbotin distribution parameters over longer time intervals: both $\alpha$ and $\sigma$ gradually increase as the time increases. Another trend is the gradual improvement of the quality of the fits with longer time intervals, as indicated by gradual decreases in both the K-S statistic and the Chi Square values. In fact, using the Chi Square test on 103 degrees freedom, the only time interval where the Subbotin PDF formally fits the empirical data is the 1998-03 interval. This is also confirmed from the visual impression of the graphs (a)-(e). Nevertheless, even for the four cases (a)-(d) where the Subbotin fit is formally rejected, this distribution provides a useful approximation to the empirical data.

Although we included $\theta$ in our grid search of the parameter estimates, it can be seen from Table 5 that its value was always estimated at 0 for these data. $\theta$ is a location parameter, but it is not the mean of the distribution. If the log growth rates $g$ conform to a 4-parameter Subbotin distribution, then the expected value, $E(g)$, is given by (Ayebo and Kozubowski [2003]) $E(g) = \theta + \sigma\left(\frac{1}{\kappa} - \kappa\right)\frac{\Gamma(2/\alpha)}{\Gamma(1/\alpha)}$. As an example, using the parameter estimates for the 1998-03 time interval from Table 5 ($\alpha = .83$, $\sigma = .30$, $\kappa = 1.20$, and $\theta = 0$), we get $E(g) = -.15$. This is quite close to the observed mean of -.16.

## IV. Conclusions: The Extreme Character of Firm Growth

In this report we have summarized our research into the nature of firm growth in the U.S., using comprehensive data on U.S. businesses derived from Census Department data. We have analyzed these data with a particular eye to the departures from standard assumptions and results in economics generally and within the field of industrial organization specifically. We have first confirmed the previous findings of Teitelbaum and Axtell [2005], extending Stanley *et al*. [1996], that log firm growth rate distributions, conditional on firm size at the beginning of a period, are very heavy-tailed relative to the normal distribution. The empirical distributions are reasonably well-fit by a Subbotin distribution, with the key exponential parameter universally estimated to be less than unity (with unity representing the Laplace distribution and 2 being the normal). Beyond this, we have two new main findings.

Our first main result concerns how heavy-tailed growth rates, as modeled by the Subbotin distribution, systematically change when the time period of aggregation changes. We find that longer time periods correspond to a slight 'thinning out' of the heavy tails. Although the relatively brief period under scrutiny here — 5 years — does not allow for a high precision estimate of the duration required to transit to a truly light tailed distribution such as the normal, the trend in this direction appears definite.

Our second main result concerns the asymmetry between firm growth and firm shrinkage, i.e., between positive and negative firm growth. From earlier time series studies (e.g., Davis, Haltiwanger and Schuh [1996]) it has been perceived that there is a systematic difference between job creation and job destruction processes. Here we quantify this independently, by noting the asymmetry in the growth rate distribution and the fact that there is more variance on the negative growth side of the distribution, albeit essentially equal mass. Once again, given that most of the firms in our data are small businesses, we are able to draw a conclusion for how small business firms are run. Specifically, given that relatively more small firms experience negative growth of any particular magnitude than those that experience comparable positive growth, it would seem that successful small firms—including those that are on their way to becoming mid-size or larger firms—must have a way to grow steadily and avoid negative growth episodes. Firm size shrinkage is *not* just an inverse event from firm size increase, and firms that regularly flirt with both sides of the distribution would appear to play a risky game. While size fluctuations are difficult to control, and our conclusions on this issue would be more clear if we had access to growth autocorrelation data, suffice it to say here that policies aimed at helping small (and perhaps shrinking) businesses through difficult times might go far to increase the pool of ultimately successful small businesses.

Overall, our study has shed new light on old topics of firm growth, particularly with respect to small businesses, since the vast majority of firms in our sample are classified as small. We hope that a better understanding of basic firm dynamics will result from these investigations.

## Acknowledgments

The authors thank M.E. Cole III from the U.S. Census Bureau for providing the data extract used in the analyses.

## References

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

Ayebo, A. and T.J. Kozubowski. "An Asymmetric Generalization of Gaussian and Laplace Laws", *Journal of Probability and Statistical Science,* 1(2), 187-210, 2003.

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

Fu, D., F. Pammolli, S.V. Buldyrev, M. Riccaboni, K. Matia, K. Yamasaki, H.E. Stanley. *Proceedings of the National Academy of Science,*102(52): 18801-18806., 2005.

Gibrat, R. *Les Inégalités Économiques; Applications: Aux Inégalités des Richesses, à la Concentration des Entreprises, Aux Populations des Villes, Aux Statistiques des Families, etc., d'une Loi Nouvelle, La Loi de l'Effet Proportionnel*. Paris, Librarie du Recueil Sirey, 1931.

Havemann, F., M. Heinz, R. Wagner-Doebler. "Firm-Like Behavior of Journals? Scaling Properties of Their Output and Impact Growth Dynamics", *Journal of the Society of Information Science and Technology,* 56(1): 3-12, 2005.

Ijiri, Y. and H. Simon. *Skew Distributions and the Sizes of Business Firms,* Amsterdam, North-Holland Publishing Company, 1977.

Kotz, S., T.J. Kozubowski, K Podgorski. *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance.* Birkhauser: Boston, Mass., 2005.

Popkin, J. "Business Survivor Rates by Age Cohort of Business", Small Business Administration Research Summary 122. Published under SBA contract SBA 4118-OA-89, 1991.

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

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

Teitelbaum, D. and R. Axtell. "Firm Size Dynamics of Industries: Stochastic Growth Processes, Large Fluctuations and the Population of Firms as a Complex System" Small Business Administration Research Summary #247. Published under SBA contract SBAHQ-03-Q-0015, 2005.