Are Big Firms Born or Grown?
—*Preliminary and Incomplete*—

Rebecca Hellerstein and Miklós Koren*

September 15, 2006

Abstract

The distribution of firm size is extremely skewed. To understand the aggregate economy, we need to understand the behavior of those few big firms that account for most of aggregate sales, employment, business cycle fluctuations, and trade. In this paper we ask why big firms are big; are they born big or do they grow big over time? In contrast to the existing literature, we show empirically that firm-size dynamics play a small role in shaping the dispersion of the overall firm-size distribution. We then build a model that explains both firm-size heterogeneity upon entry (as a product of an idea-generating process together with resource allocation across those ideas) and subsequent firm-size dynamics (via Jovanovic-style learning) that is consistent with the stylized facts for the dispersion of firm-size distributions.

This paper quantifies the effects of firm size at entry on the distribution of firm sizes over time. The distribution of firm sizes is extremely skewed: A small number of firms account for a large share of output. This has implications for our understanding of a range of macroeconomic and international-trade issues, such as the magnitude of business cycles (Gabaix, 2005) and the patterns of export-market participation (Bernard, Redding and Schott, 2005). Much of the literature on firm-size distributions has focused

*Federal Reserve Bank of New York. Corresponding author: Miklós Koren, International Research Group, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045; E-mail: `Miklos.Koren@ny.frb.org`. We thank Aaron Cooper for valuable research assistance. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or of the Federal Reserve System.

on the dynamics of individual firm sizes over time, deriving ergodic size distributions from models of firm entry, growth, and exit. This paper argues that pre-entry scale decisions may be more important than subsequent firm dynamics in causing the observed skewness in firm size distributions.

We begin by showing that firm sizes are very persistent over time using *Compustat data*. We find the correlation of a firm's current size with its size 25 years ago to be 0.71, which means that size depends more on initial conditions than on subsequent firm dynamics over 25 years. We then demonstrate that this finding means firms fail to converge to an ergodic size distribution over even quite long horizons. The existing literature on the causes of skewed firm size distributions may be mistaken in its (*de facto*) assumption of a Markovian model of firm growth (a la Hopenhayn (1992)).

We quantify the importance of firm size at entry (initial sales) relative to the dynamics of firm size over time (variance in sales over time) for the cross-sectional distribution of firm sizes. We conclude with a simple non-Markovian model of the distribution of firm sizes in which an entrepreneur's pre-entry scale decision (which allows for shocks, both temporary and persistent, to profitability and entrepreneurial learning a la Jovanovic) determines the majority of their sales variance over time: Where firms start is where they end up. Our results are similar to those of Abbring and Campbell (2005), who show that Texas bars demonstrate non-Markovian history dependence – initial sales significantly contribute to forecasts of sales going forward, and those of Pakes and Ericson (1998) who find that retail firms exhibit non-Markovian growth in sales over time.

1 Literature

The widespread appearance of skew distributions in various settings, particularly in the biological sciences, has received increased attention in recent years in economics.[1] A number of recent papers examine the extreme skewness of the size distribution of firms.

Axtell (2001) finds that the log of the right tail probabilities of the log-size distribution, with firm size measured by employment, is on a virtual straight line with a slope of -1.059. This suggests that the distribution of firm size is well approximated by a Pareto

[1] "Kapteyn argued that underlying such distributions was a simple Gaussian process: a large number of small additive influences, operating independently, would generate a normally distributed variate z. The observed skew distribution of some variate x could be modelled by positing that some underlying function of x was normally distributed. Gibrat postulated that the log of x developed in such a fashion, that is, that growth was scale independent."

distribution with right tail probabilities in the form of $S^{-\theta}$ where S is the size of the firm and $\theta = 1.05$. A number of authors argue that it is remarkable how well this distribution fits the log-size distribution of firms and some efforts have been made to connect it to data on firm growth, exit, and entry.[2]

A number of theoretical papers seek to explain observed firm-size distributions by appealing to models of firm dynamics. For example, Klette and Kortum (2004) develop a model of R&D that produces a stable skewed firm-size distribution with firm growth rates independent of size (with growth dynamics that follow Gibrat's law). [3]

We ask whether the recent focus on firm dynamics as the source of the observed skewness in firm size distributions has been misplaced. We suggest that pre-entry scale decisions may dominate firm dynamics in determining the distribution of firm sizes at any point in time.

2 The Markovian Model of the Firm Size Distribution

This section presents a reduced-form model of Markovian firm dynamics, in the flavor of Hopenhayn (1992). The model is general enough to serve as a benchmark that encompasses most existing Markovian models.

Firm-size dynamics are first-order Markov: the current size of the firm is a sufficient statistic to describe the stochastic paths in the future. Firms are born with an exogenous size. They are hit by exogenous productivity shocks, which induces endogenous response in firm growth. Firms may choose to exit if they are unproductive. There is no aggregate uncertainty.

There are four key ingredients of the model: a function describing growth dynamics, a function describing exit rules, the entry size(s) of firms, and the mass of new entrants. We take the model as a reduced form since we do not discuss how the dynamics is derived from the primitives of the model (technologies, stochastic process of productivity, etc.).

Let $f(\Delta y_t|y_{t-1})$ denote the probability (density) that a firm will grow by Δy_t conditional on being size y_{t-1} in year $t-1$. This distribution is conditional on the firm surviving until year t. Let $h(y_{t-1})$ denote the probability that a firm with size y_{t-1} in

[2]Gabaix (1999) discusses more generally how probability models give rise to power laws and the implications for economics.

[3]See also Luttmer (2005).

year $t-1$ will choose to exit in year t. The firm starts from size y_0 at time $t=0$, which is drawn from a distribution $G(y_0)$. Each year there are a measure μ of new, ex ante identical entrants. For ease of exposition, we assume that all distributions are continuous.

Let $\phi(y,t)$ denote the measure of firms of size y in year t. This firm size distribution evolves as follows:

$$\phi(y,t) = \mu g(y) - h(y)\phi(y,t-1) + \int_x f(y-x|x)\phi(x,t-1)dx. \tag{1}$$

There is a mass $\mu g(y)$ of firms that enter at exactly size y. A mass $h(y)\phi(y,t-1)$ of firms exits between $t-1$ and t. Each firm with size x in year $t-1$ becomes size y with probability $f(y-x|x)$, hence the measure of firms growing (or shrinking) to size y is $f(y-x|x)\phi(x,t-1)$.

In year 0, the distribution of firms is just

$$\phi(y,0) = g(y). \tag{2}$$

Given the functions f, g, and h, one can derive the evolution of the firm size distribution, $\{\phi(y,1),...,\phi(y,t)\}$, and an ergodic distribution (if exists),

$$\phi_\infty(y) = \lim_{t\to\infty} \phi(y,t).$$

An invariant distribution $\tilde{\phi}$ satisfies the functional equation

$$\tilde{\phi}(y) = \mu g(y) - h(y)\tilde{\phi}(y) + \int_x f(y-x|x)\tilde{\phi}(x)dx,$$

and under some conditions, it is unique and coincides with the ergodic distribution.

To illustrate how the properties of f, g, and h affect the evolution of the firm size distribution, let us consider some simple examples.

Example 1 (Random growth) *Suppose there is no entry or exit, $\mu = h = 0$, and growth is independent of current size, $f(x|y) = f(x)$. Then big firms are just as likely to grow as small firms, and small firms are just as likely to shrink as big firms. The size dispersion increases without bound, and there is no limiting distribution.*

More concretely, if $f(x)$ has mean μ and variance σ^2, then, by the law of large numbers, $(y_t - \mu t)/(\sqrt{t}\sigma)$ converges to a standard normal distribution as $t \to \infty$.

Example 2 (Mean reverting growth) *Suppose there is no entry or exit, $\mu = h = 0$, and growth is negatively related to current size. That is, if $x' < x$, then $f(y|x')$ stochastically dominates $f(y|x)$ in the first order. Then there exists a steady-state size distribution.*

More concretely, suppose y_t is $AR(1)$ with iid Gaussian disturbances,

$$\Delta y_t = -\rho y_{t-1} + u_t,$$

and y_0 is normally distributed with variance σ_0^2. Then y_t is normally distributed with variance

$$\mathrm{Var}(y_t) = \sigma_0^2(1-\rho)^{2t} + \sigma_u^2 \frac{1-(1-\rho)^t}{\rho}.$$

The steady-state variance is σ_u^2/ρ, independent of the initial size distribution.

Example 3 (No entry) *Suppose there is no entry, and f and h are such that an ergodic distribution exists. Then the initial size distribution of firms does not affect the ergodic distribution.*

We can summarize existing Markovian models of firm dynamics with what they state (or assume) about f, g, and h.

Model	Entry size	Growth rate	Exit rule
Gabaix	none	size independent for large firms	none
Luttmer	drawn from incumbents	size independent	small size
Klette–Kortum	small	size independent	size = 0
Rossi-Hansberg–Wright	same as incumbents	mean reverting	size independent

Table 1: Markovian Models of Firm Dynamics

3 Data

Our data comes from Standard & Poor's Compustat North America Industrial Annual (NAIA) database. Compustat provides financial data covering publicly traded companies in the U.S. and Canada. Companies are added to the Compustat database when their stock begins trading on the New York Stock Exchange, American Stock Exchange

or NASDAQ. Companies must meet additional Securities and Exchange Commission requirements in order to be included.

Compustat's NAIA database presents annual data from 1950 through 2004. This is a particularly appealing characteristic of the data, since it allows us to thoroughly study the long-term dynamics of the firm size distribution. Our analysis uses firm sales data (as opposed to the popular employment figures) as a measure for firm size. The correlation of employees with sales is .63. Since we are interested in the output share of large firms in the distribution, sales data makes intuitive sense. Because of the extreme skewness of the size distribution, we also find it convenient to examine the data using logs.

The sample changes dramatically over the course of its 50 year span. Compustat reports sales for 652 firms in 1950, increasing to over 7100 by 2004. Overall, there are 24600 firms in the sample. Most firms are added to the data in three waves, one group in 1950, another in 1966 and a third in 1985. The sample of firms reporting sales grows much more smoothly over time, with noticeable jumps in 1960, 1974 and 1995 (see Figure 1). The newly added firms are typically smaller, which induces downward jumps in the average size of firms over time (Figure 2). Grouping firms into cohorts based on entry year allows us to control for the volatility of the data due to such influxes.

The average lifespan of a firm is just over 12 years. There are, however, 4400 firms with at least 20 years of data (the number is 2200 for 30, 920 for 40, and 380 for 50 years). This enables us to study the long-run dynamics of a large set of firms.

While firms may enter the sample in a consistent manner, the nature of the Compustat database prevents us from observing the true 'birth' of a firm; consequently, we proxy a firm's entry by its first year of reported sales. On the other hand, there are a multitude of reasons for firm exit. In Compustat's NAIA dataset, firms may exit via bankruptcy (the most relevant for standard models of firm dynamics), merger and acquisition, going private, liquidation, or leveraged buyout. We control for exits other than bankruptcy in most cases by dropping those firms from the data entirely.

Over 20 2-digit NAICS industries are represented, as well as many more sub-categories of these industries. Manufacturing is the most highly represented industry in the data, with Information, Financial Services, Mining and Utilities as other large industries. In 1990, for example, 25 percent of the firms were in Manufacturing, while 11 percent were in Financial Services and 7 percent were in Information.

One final limitation of the data is that we can only observe the size dynamics of the

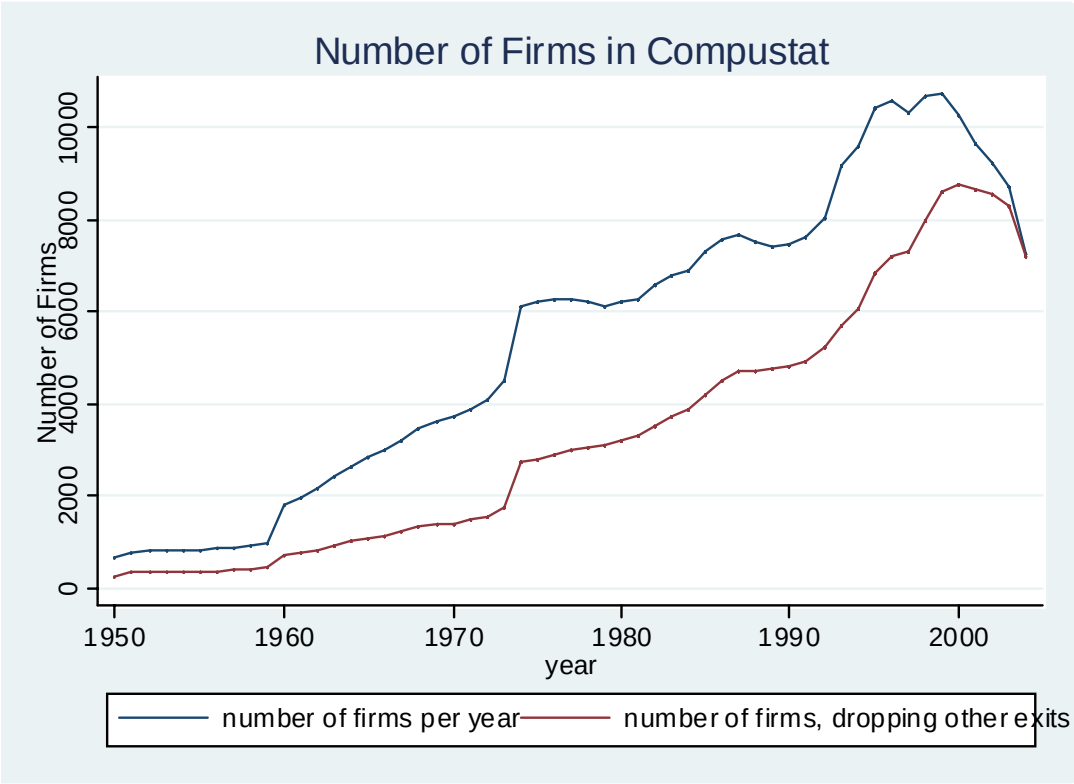

Figure 1:

largest firms in the economy. However, we have reason to believe that this will not limit our analysis, particularly as it pertains to the largest firms in the size distribution.

4 Stylized Facts on the Evolution of the Firm Size Distribution

Fact 1 *The size distribution of firms is extremely skewed.*

The top 5 percent of firms sell about 50 percent of output. The distribution of firm size is well approximated by a lognormal distribution. The cross-sectional standard deviation of log sales in 1999 is 2.69. This means that the 75th percentile firm sells about 30 times as much as the 25th percentile. (The 90-10 range is about 760.) Figures 3–5 display the histogram of firm sales, log firm sales, and the Lorenz curve of firms sales, respectively.

Fact 2 *The size dispersion is much higher within industries than across industries.*

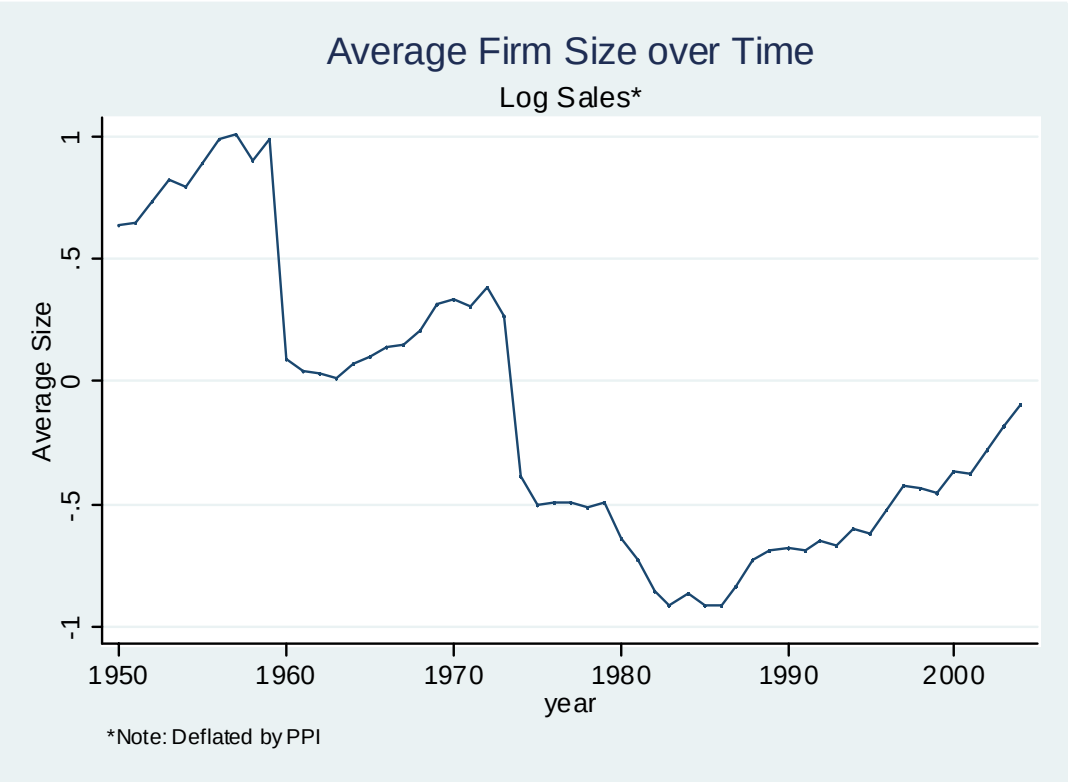

Figure 2:

Industry fixed effects (4-digit NAICS) explain only 25 percent of the variation in firm size. Broader sectors (23 2-digit NAICS categories) account for only 9 percent. This indicates that the vast heterogeneity of firms sizes is not primarily due to technological or market structure differences across industries, as all industries exhibit huge size variation.

Fact 3 *For a given cohort, the size distribution is relatively stable over time for earlier cohorts, and increases for later cohorts.*

We proxy firm entry by using the first recorded sales data, and then group firms in 5 year cohorts according to firm entry year. We analyze cohorts to make sure that the dynamics of the size distribution is not driven by new firms being added to the dataset. We use the standard deviation of log sales to examine the dispersion of the size distribution over time. Figure 6 shows the evolution of dispersion for all 5-year cohorts between 1950 and 1970. The 1950 and 55 cohorts are relatively stable, whereas later cohorts exhibit increasing dispersion.

At least for the older, more established cohorts, the data contradicts models assuming or implying Gibrat's law that states that growth rates are independent of size (Gabaix,

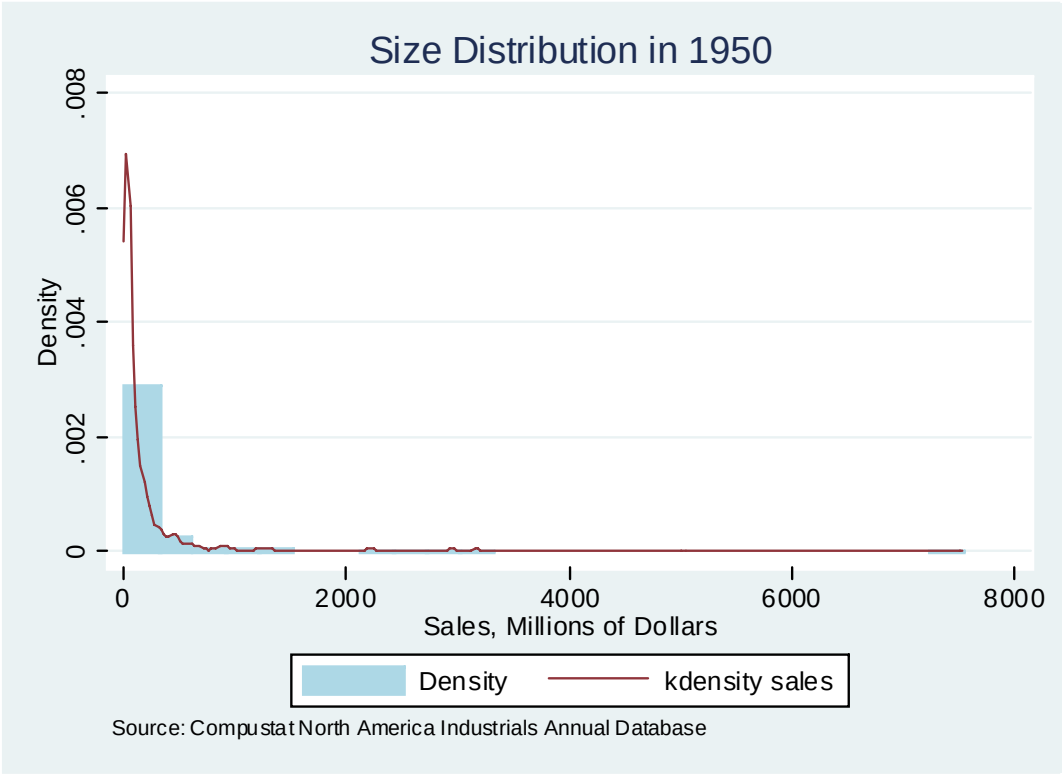

Figure 3:

1999; Luttmer, 2004; Klette and Kortum, 2004). These models would imply a continuous increase in the size dispersion of a given cohort of firms.

Fact 4 *Differences in firm growth contribute little to changes in the firm size distribution.*

Figure 7 presents the results of the following decomposition exercise:

$$\begin{aligned}\mathrm{Var}(y_t) &= \mathrm{Var}(y_{t-1} + \Delta y_t | \text{survival}) \\ &= \mathrm{Var}(y_{t-1}) + [\mathrm{Var}(y_{t-1}|\text{survival}) - \mathrm{Var}(y_{t-1})] \\ &\qquad + \mathrm{Var}(\Delta y_t|\text{survival}) + 2\,\mathrm{Cov}(\Delta y_t, y_{t-1}|\text{survival})\end{aligned}$$

The cross-section variance (dispersion) of firm sizes depends on past dispersion, non-random exit (if smaller firms are more likely to exit, this compresses the distribution), the dispersion of growth rates, and their relationship to the firms' past size. In particular, if Gibrat's law holds (growth rates are independent of size), the covariance term is zero

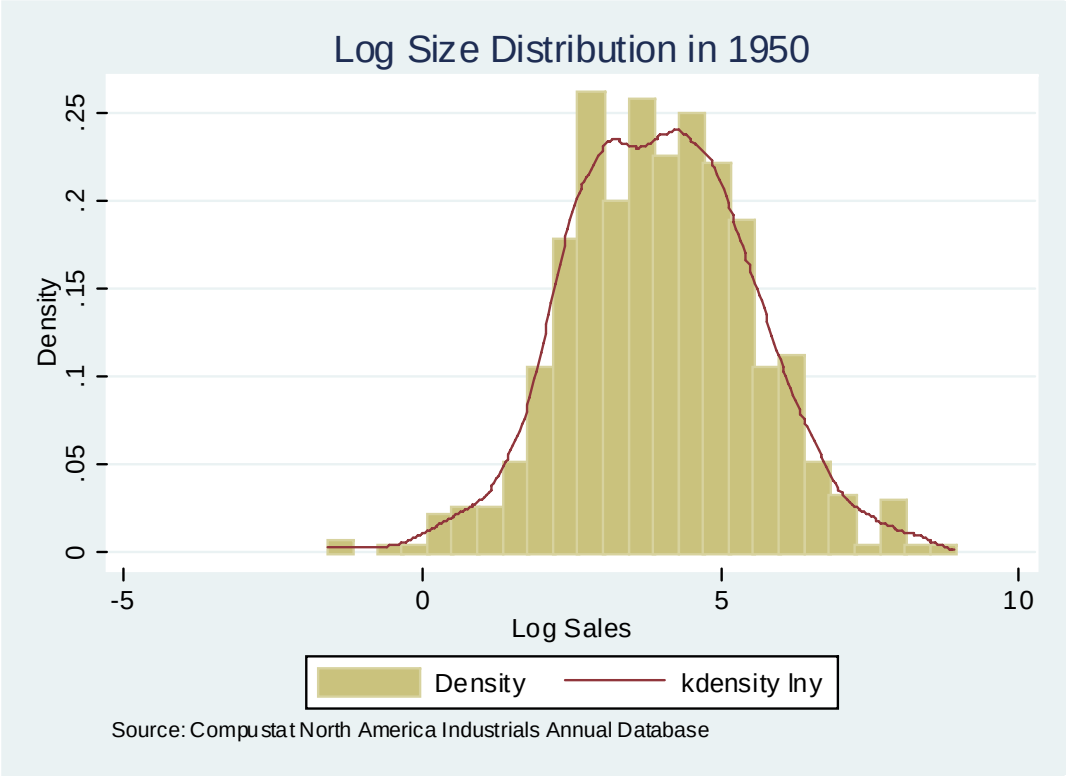

Figure 4:

and idiosyncratic growth dynamics increases size dispersion. Mean reverting models imply a negative covariance term.

Note that the use of Compustat may bias the results of this exercise somewhat, as publicly listed firms tend to be less volatile than privately held firms. The cross sectional dispersion of the growth rate of firms in the Longitudinal Business Database of the Census (Davis, Haltiwanger, Jarmin and Miranda, 2006) is about 0.30, about 50% higher than in Compustat. However, the dispersion of firm *size* is also higher in the universe of firms than in Compustat, which suggests that this bias is not quantitatively important.

This finding demonstrates that the size dispersion of a cohort is primarily determined by its 10-year-ago size dispersion, and is little affected by growth and exit throughout the past 10 years. This is in contrast with models that derive the ergodic size distribution from mean-reverting growth dynamics (Rossi-Hansberg and Wright, 2004).

What can we learn about the largest firms in the sample? Figure 8 tracks the shares of GDP of the biggest 5% of firms in the data over time. We select firms based on their

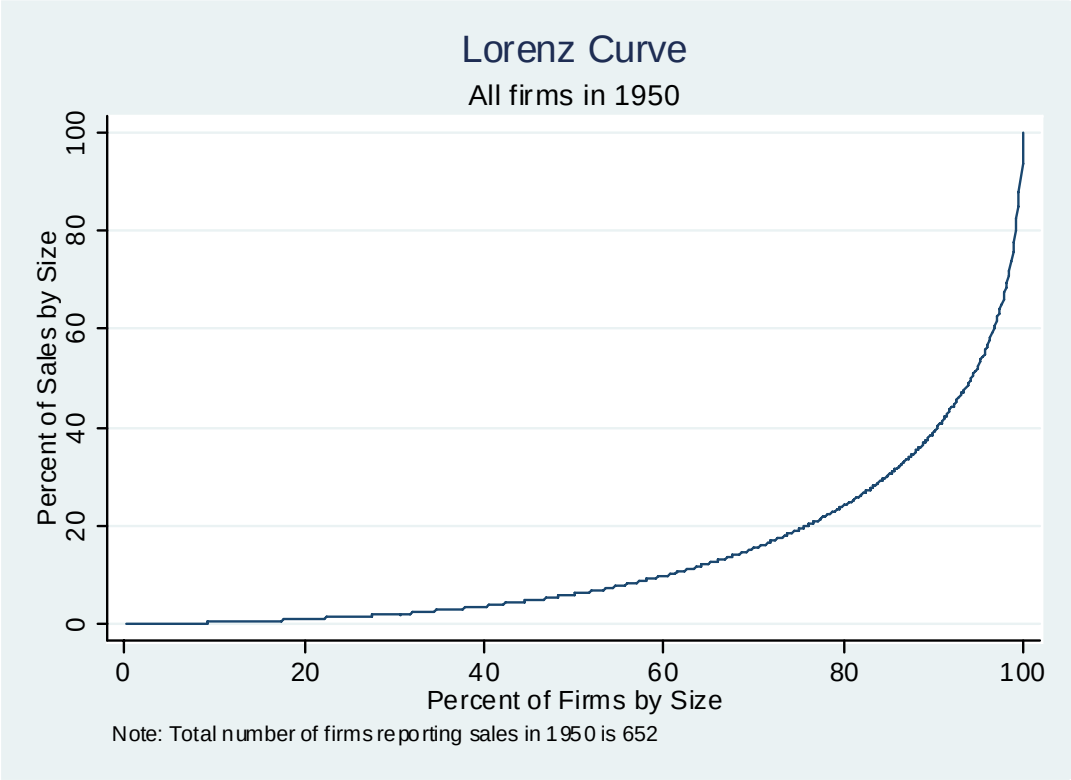

Figure 5:

size in their seventh year. These shares are very steady over time, although they vary between different cohorts, suggesting that big firms remain big once they are established.

Fact 5 *The position of top firms is remarkably stable over time.*

The total sales of the top 5 percent of firms in 1950 remained 18–30 percent of total GDP throughout the subsequent 40 years. For all entry cohorts, there is very little tendency of the big firms to decline. Big firms are remarkably stable and are able to keep their relative positions in the economy for very long periods. Again, this indicates that year-to-year growth dynamics is relatively unimportant for the understanding of these firms. If there existed substantial mean reversion in size, one would expect big firms to shrink rapidly.

Fact 6 *Throughout the distribution, there is little turnover in firm sizes.*

To examine the persistence of size throughout the distribution, we estimate transition matrices for firm size. Firms are most likely to remain in their size bin in the next decade

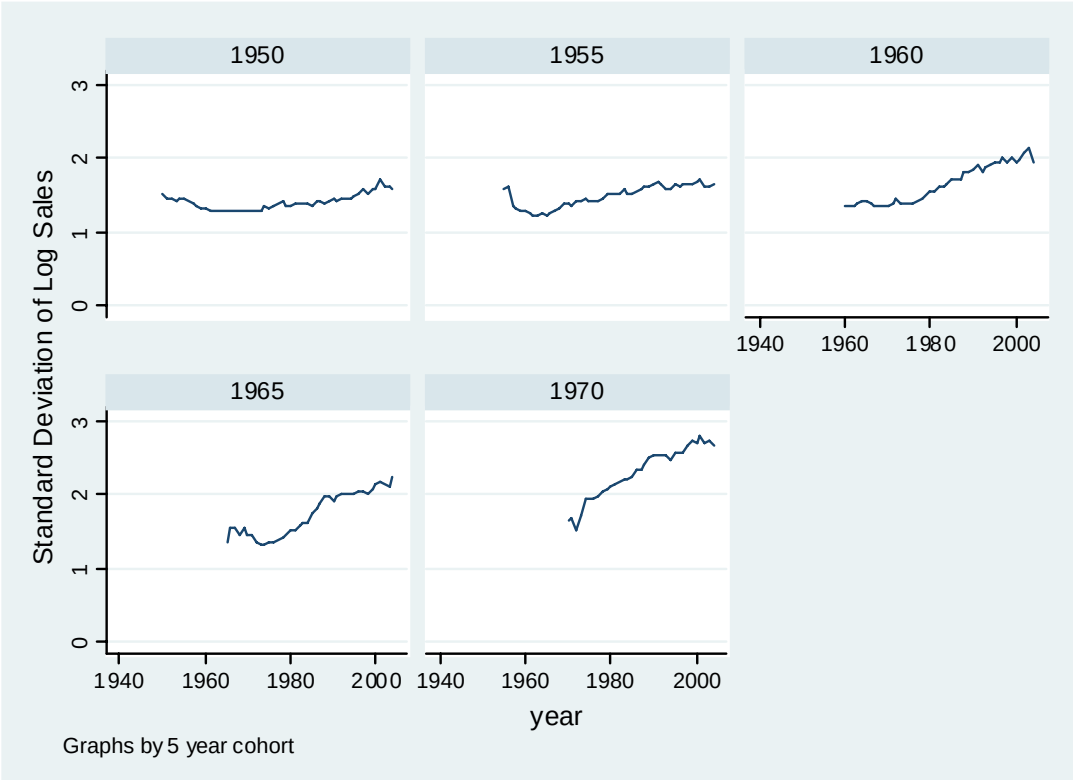

Figure 6: The Evolution of Within-Cohort Size Dispersion Over Time

Table 2: Transition Probabilities from 1980 to 1990

		1990					
		<10	**<100**	**<1000**	**>1000**	**Exit**	***Total***
	<10	0.7054	0.1752	0.0181	0.0000	0.1012	*662*
1980	**<100**	0.1861	0.6227	0.1207	0.0030	0.0674	*994*
	<1000	0.0098	0.1540	0.7269	0.0850	0.0242	*1,117*
	>1000	0.0039	0.0019	0.1965	0.7881	0.0096	*519*
	Total	*665*	*908*	*1,046*	*507*	*166*	***3,292***

Notes: Bins are based on sales levels in millions of dollars. Firms that exit for reasons other than bankruptcy are not included.

(with a probability of $0.62 - 0.79$). Based on the estimated transition matrix, we can calculate long-run transition probabilities to examine the long-run dynamics of firms. Firms starting in the lowest bin have only a 17 percent probability of ever moving to the top bin. Those that do reach the top bin do so in 152 years on average. Firms starting

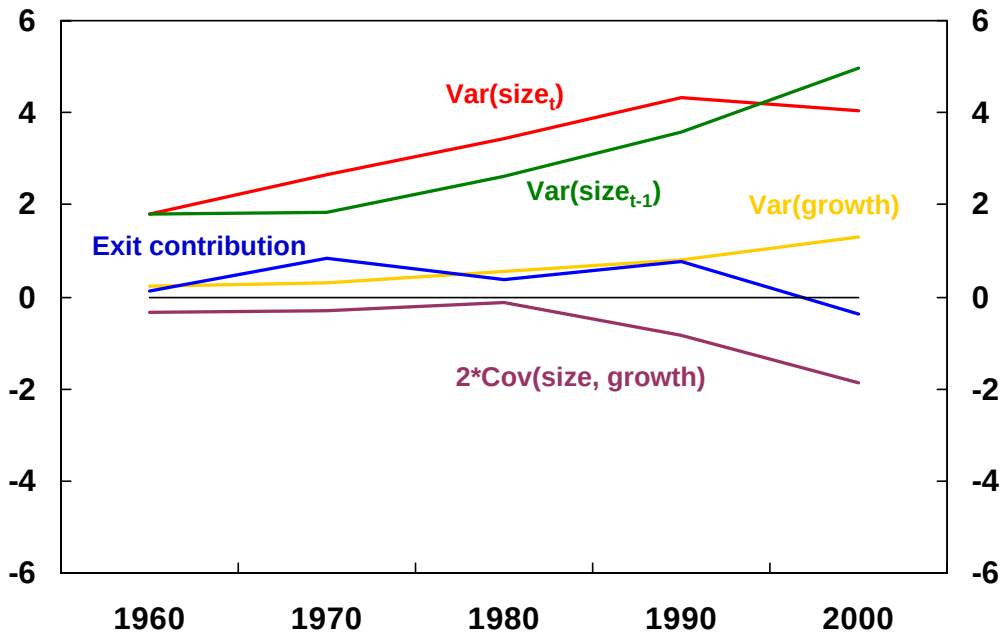

Figure 7: Decomposition of Size Dispersion

Notes: Firms exiting for reasons other than bankruptcy are not included.

in the top bin have an expected lifetime of 247 years. However, 41 percent of these firms will go bankrupt before shrinking to the smallest bin.

This finding demonstrates that the lack of turnover carries over to other parts of the distribution. The model that firms enter small, grow, shrink and then die (e.g., Klette and Kortum, 2004) is not consistent with the data. Small firms are very unlikely to ever make it to the top. A big firm has a large chance of going bankrupt even before shrinking substantially.

Fact 7 *Firm size dynamics is non-Markovian.*

We regress current size of the firm on lagged sizes and its size upon entering Compustat. Figure 9 reports the estimated coefficients and the 95 percent confidence intervals. Even at age 20, the initial size matters significantly.

In a more structural exercise, we estimate a fixed-effects $AR(p)$ model ($p = 1, 2, 3$),

$$y_{it} = \alpha_i + \sum_{k=1}^{p} \beta_k (y_{i,t-k} - \alpha_i) + u_{it}$$

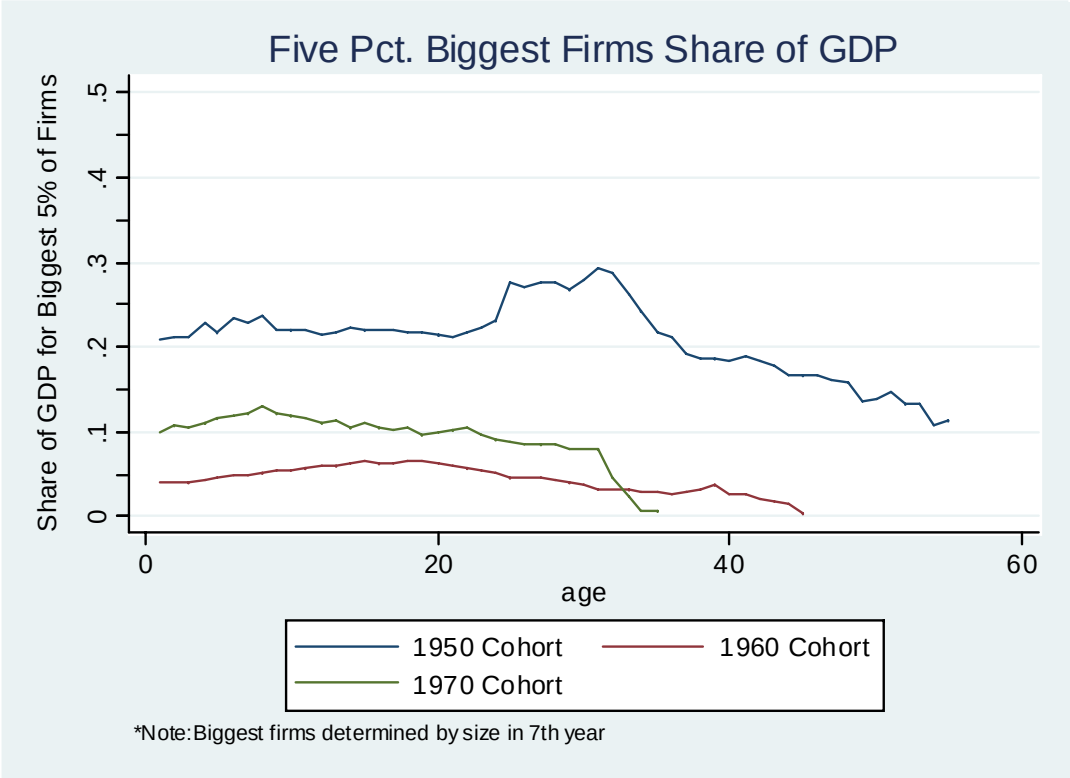

Figure 8: The Share of Large Firms over Time

using the Arellano and Bond (1991) estimator. In this model, the expected steady-state size of firm i is

$$\lim_{T\to\infty} E_t(y_{iT}) = \alpha_i.$$

As long as α_i varies across firms, their dynamics conditional on past sizes is different. That is, dynamics is not Markovian.

We find that the variance of fixed effects is 5.93, compared to the variance of firm size in 2004, 6.71. Overall, α_i explains 88 percent of the variation in firm sizes.

These exercises provide some evidence that firms may fail to converge to an ergodic size distribution at all. That is, not only is firm size persistent, but different firms tend to converge to different long-run sizes. Though admittedly indirect (Compustat only contains publicly listed firms), our evidence also points to a substantial heterogeneity of firm sizes upon entry.

The next set of facts consider the dynamics of individual firms.

Fact 8 *Bigger firms are more stable.*

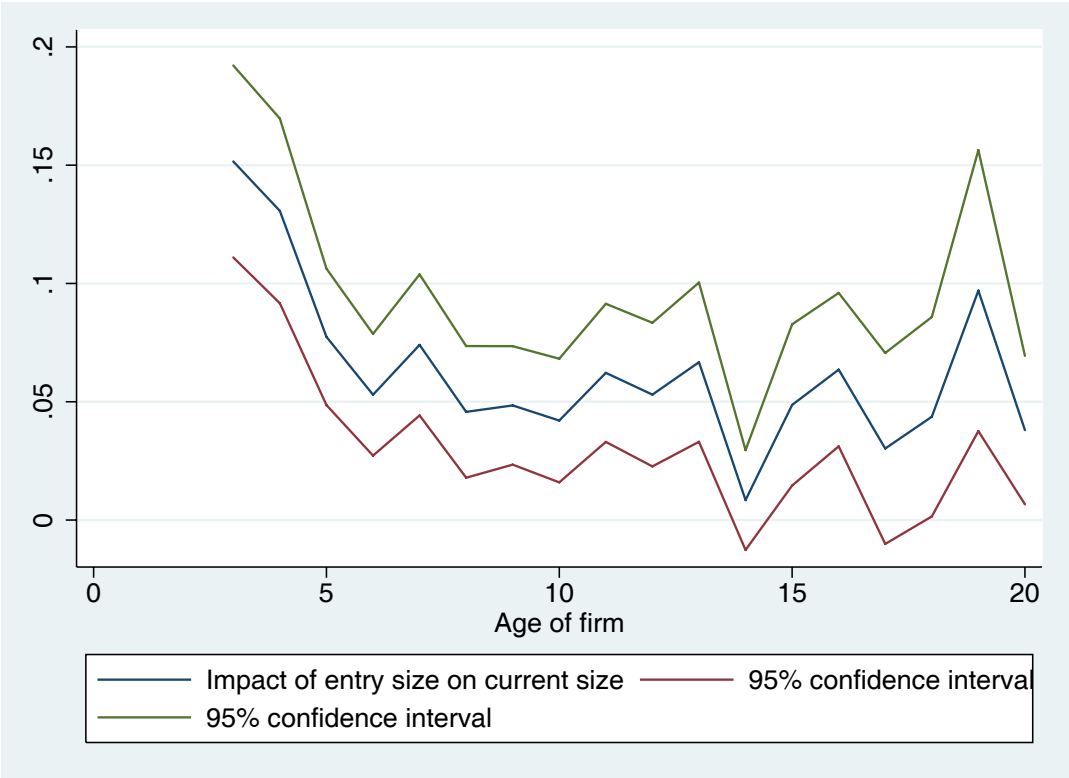

Figure 9: The Impact of Entry Size on Current Size

Notes: Reports regression coefficient of entry size and its 95% confidence interval based on heteroskedasticity-consistent standard errors. Regression controls include lagged size and year fixed effects.

Fact 9 *Older firms are more stable.*

Table 3 presents the results of a regression of sales growth volatility on firm size and firm age. Volatility is measured as the standard deviation of sales growth between 1989 and 1998. The first column displays the individual impact of size on volatility. Bigger firms tend to be more stable. The second column shows the individual impact of age. Older firms tend to be more stable. Note, however, that older firms also tend to be bigger, so the two results are not independent. Column 3 displays the partial impacts of the two variables. Both size and age have an independent negative impact on volatility, although the impact of age conditional on size is much smaller than its unconditional correlation with volatility.

The above stylized facts indicate that there are persistent firm-specific factors that

Table 3: Sales Volatility, Size and Age

	Standard deviation of growth rates (log)		
Sales (log)	–0.199***		–0.178***
	(0.007)		(0.008)
Age (log)		–0.187***	–0.062***
		(0.012)	(0.013)
Observations	4319	4086	4086
R–squared	0.17	0.07	0.18

Notes: Standard deviation of sales growth is taken between 1989 and 1998. Size and age are measured in 1989. Robust standard errors in parentheses.

determine the firm's long-run size. This is in sharp contrast with models where firms are ex ante homogeneous and all ex post heterogeneity in size is the outcome of idiosyncratic size dynamics ("luck").

Firm size is partly predetermined upon entry, and there are ex ante heterogeneities across firms. Ex ante heterogeneity is an important feature of models such as the span of control model of Lucas (1978) (also see Atkeson and Kehoe, 2005) or the learning model of Jovanovic (1982). However, these models tend to assume an exogenous distribution of firm types ("managerial ability," productivity etc.) Heterogeneity in the entry sizes of firms is entirely driven by the exogenous type heterogeneity.

To better understand what determines a firm's pre-entry scale decision, we build a model where potential entrants' optimal search, selection and matching affects the size distribution in an endogenous manner.

5 A Model of Pre-Entry Size Heterogeneity and Post-Entry Learning

Firms require a blueprint and some resources (say, capital) to produce. Blueprints may correspond to firm-level strategic decisions about management style, product lines, production techniques, among others. One firm has exactly one blueprint and blueprints

are not marketable across firms.

5.1 Technology

Output of a firm depends on the quality of the blueprint (productivity), A and the quantity of inputs it uses, K. We assume that the production function is multiplicatively separable in A and K

$$Q = g(A)f(K). \tag{3}$$

This separable formulation emphasizes the scale economies of blueprints. A blueprint that improves the productivity of a small firm would improve the productivity of a large firm by the same proportion.[4]

By choosing units appropriately, we normalize $g(A) = A$. That is, A is a measure of total factor productivity. With this technology, the output of firm i is

$$Q_i = A_i f(K_i).$$

Suppose firms are ordered by productivity such that A_i is an increasing (and, as we will prove later, a continuously differentiable) function of i.

The production function $f()$ satisfies the usual Inada conditions, so that all firms with a positive productivity find it optimal to produce at a positive (but finite) scale. In future analysis, we will study the role of fixed operating costs and selection by productivity.

To measure the size dispersion of firms, we look at how sensitively firm size varies with firm rank. In particular, we measure the proportional change in output with an improvement in rank:

$$\frac{d \ln Q_i}{di} = \frac{d \ln A_i}{di} + \varepsilon_f \frac{d \ln K_i}{di}. \tag{4}$$

More productive firms will typically also use more resources (more on this later), which contributes to additional size dispersion across firms. To quantify the importance of this channel, we need to understand how the use of resources varies with the productivity rank ($i \in [0, 1]$) of the firm.

[4]Here we emphasize that the model is about the size distribution of *firms*, not *plants*. Strategic decisions likely affect all plants of a firm in a similar manner. One could introduce some small diseconomies of scale (due to limits on the span of control, say) for multiplant firms without affecting the results very much.

5.2 Market Environment

A question of key importance for the firm size distribution is how resources are allocated across firms with different productivity. Here we consider two allocation mechanisms.

5.2.1 Assignment Problem

Suppose there is a given distribution of indivisible chunks of resources. A given resource can only contribute to the production of one firm. A firm can only use one resource at a time. Resources are indivisible and their sizes cannot be changed.

The size distribution of resources is given by the continuous distribution function $G(K)$.

Because of the superadditivity (?) of the production function, the efficient assignment of resources to productivities involves positive assortative matching: bigger resources are assigned to more productive firms. In this case, the productivity rank of a firm coincides with its rank in terms of input use, $i = G(K)$.

Firm of rank $i \in [0, 1]$ will use $K = G_i^{-1}$ resources. The overall increase in firm size with firm rank is

$$\frac{d \ln Q_i}{di} = \frac{d \ln A_i}{di} + \frac{\varepsilon_f}{g[K_i]K_i}$$

If the distribution of resources is skewed, it adds to the skewness of the size distribution of firms. Note, however, that the allocation of resources only depends on the productivity *ranking* of firms, not on the actual productivity differentials.[5]

5.2.2 Competitive Markets

An alternative allocation mechanism is when firms can buy freely divisible units of K in perfectly competitive markets for a price r. In this case, they are adjusting the use of resources on the intensive margin.

$$\max_K Q_i - rK, \tag{5}$$

$$\text{s.t. } Q_i = A_i f(K). \tag{6}$$

Profit maximization ensures that the marginal product of capital equals the rental cost (assumed to be the same across firms),

$$A_i f'(K) = r.$$

[5] This approach is used by Gabaix and Landier (2006) to explain the skewness and the evolution of the distribution of CEO pay.

The capital use of firm i depends on its productivity and the curvature of the production function. More productive (higher ranked) firms use more capital because they have a higher marginal product:

$$\frac{d\ln K_i}{di} = -\frac{1}{\varepsilon_{f'}}\frac{d\ln A_i}{di} > 0. \tag{7}$$

This magnifies the effect of productivity dispersion. The size-rank relationship will be steeper, as even small productivity differences get magnified by the increased use of inputs,

$$\frac{d\ln Q_i}{di} = \left(1 - \frac{\varepsilon_f}{\varepsilon_{f'}}\right)\frac{d\ln A_i}{di}. \tag{8}$$

This mechanism is similar to Rosen's (1981) "superstar" phenomenon. In his model, talented performers earn especially high incomes because they both charge higher prices and attract larger audiences. In our model, more productive firms both make a better use of available resources and use more resources. The outcome is a skewed firm size distribution with a few "superstar" firms.

The main driving force behind this mechanism is the scale economy of blueprints. Implementing a better blueprint at a huge firm is no more costly than implementing it at a small firm. This brings about the complementarity of firm size and blueprint quality.

5.3 The Distribution of Productivity

Entrepreneurs come up with blueprints, each blueprint z drawn independently from the uniform distribution over $[0, 1]$. A blueprint z corresponds to TFP $\alpha(z)$, where $\alpha()$ is a twice continuously differentiable strictly increasing function. That is, a blueprint with a higher index corresponds to higher productivity.[6]

Entrepreneurs differ in the number of blueprints they have before setting up a firm. An entrepreneur of type N has N blueprints. When the entrepreneur starts a firm, she implements the best blueprint she has.

The best blueprint is the highest z from a set of N blueprints drawn independently (the Nth order statistic). Its distribution function is given by

$$F(z|N) = \Pr(Z \leq z) = \Pr(Z_1, ..., Z_N \leq z) = z^N.$$

[6]The function need not be bounded. An example of $\alpha(z)$ is shown in Figure 10.

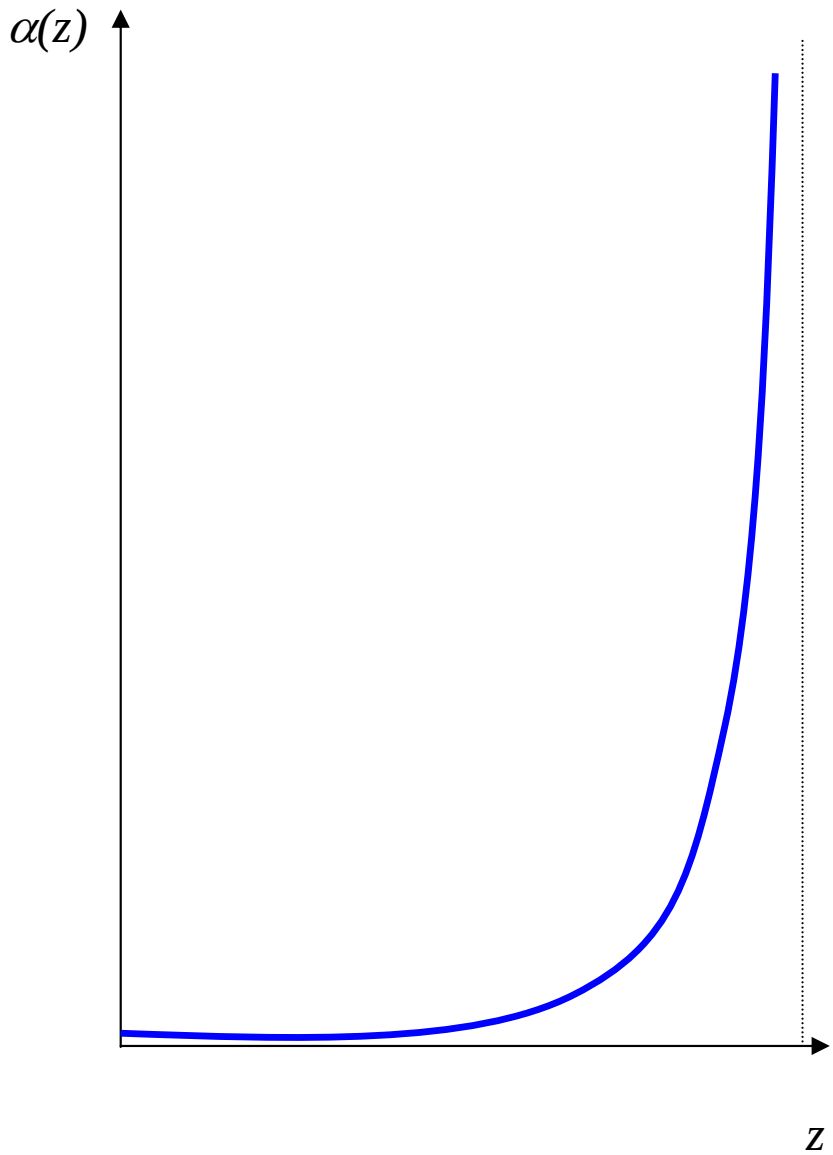

Figure 10: Productivity as a Function of Blueprint Quality

Note that $F(z|N)$ stochastically dominates $F(z|N')$ in the first order if and only if $N > N'$. This implies that entrepreneurs with more blueprints have higher expected productivity $E[\alpha(z)|N]$, because α is increasing in z.

The distribution of entrepreneurial talent (N) is Poisson with parameter λ,

$$\Pr(N = k) = \frac{e^{-\lambda}\lambda^k}{k!}.$$

This implies that both the expected number of blueprints and the cross-sectional variance of the number of blueprints is λ.

The unconditional distribution of blueprint quality is a mixture of the conditional distributions,

$$F(z) = \sum_{k=0}^{\infty} \Pr(N = k)F(z|k) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} z^k = e^{\lambda(z-1)} \sum_{k=0}^{\infty} \frac{e^{-\lambda z}(\lambda z)^k}{k!} = e^{\lambda(z-1)}.$$

This characterizes the initial distribution of z in a cohort of new firms.[7]

5.4 Dynamics

An entrepreneur knows its N and is able to select its best idea but does not know its actual quality, z. She will hence base her initial scale decision on her expected productivity.

The prior of a type-N firm before entry is $F(z|N) = z^N$. The entry scale of firm i is hence defined by

$$E[\alpha(z_i)|N_i]f'(K_{i0}) = r.$$

Entrepreneurs with more blueprints start at a bigger scale.

Similarly to Jovanovic (1982), entrepreneurs learn about their true productivity every year. More specifically, the entrepreneur receives a binary signal of the quality of her blueprint, $w_t \in \{\text{success}, \text{failure}\}$. The distribution of w_t is independent Bernoulli with the probability of success being z, the true quality. The firm then updates its prior, obtaining a Beta distribution as its posterior.

[7] In general, any power series distribution (PSD) for N would provide a closed-form solution for the initial distribution of blueprints. A PSD is of the form

$$\Pr(N = k) = \frac{a_k}{\alpha(\theta)}\theta^k,$$

where $\alpha(\theta) = \sum_{k=0}^{\infty} a_k\theta^k$. Mixing $F(z|N)$ with this PSD,

$$F(z) = \sum_{k=0}^{\infty} \Pr(N = k)F(z|k) = \frac{\alpha(\theta z)}{\alpha(\theta)}.$$

If after T years, a type-N firm has had k successes and $T-k$ failures, its posterior for its true quality is Beta

$$\tilde{z}|N,T,k \sim B(N+k, T-k+1).$$

The expected quality, $E\tilde{z}_T = (N+k)/(N+T+1)$, is increasing in the fraction of successes and N. The precision of the entrepreneur's estimate of the true quality is

$$\frac{1}{Var\tilde{z}_T} = \frac{(N+T+1)^2(N+T+2)}{(N+k)(T-k+1)},$$

increasing in N and T. That is, volatility decreases with both size at entry and the age of the firm. Eventually, as $T \to \infty$, $Var \to 0$ and the firms gets to learn its true quality, z, with probability one. Then its scale decision will be based on the *actual productivity*,

$$\alpha(z_i)f'(K_{i\infty}) = r.$$

This feature of the model is consistent with two of the stylized facts about firm-size dynamics identified in our data. First, that older firms are more stable; And second, that bigger firms are more stable.

6 Conclusion

Why are big firms big? Are they born that big or do they grow big over time? The empirical analysis of Compustat data indicates that firm dynamics explain only a small portion of differences in firm sizes over time. Instead, there are persistent firm-specific factors that determine a firm's long-run size.

The distribution of firm sizes displays the following regularities: First, it is extremely skewed. Second, the size dispersion is much higher within industries than across industries. Third, for a given cohort, the size distribution is relatively stable over time for earlier cohorts, and increases for later cohorts. Fourth, differences in firm growth contribute little to changes in the firm-size distribution. Fifth, the position of very large firms is remarkably stable over time. Sixth, older firms are more stable. And seventh, bigger firms are more stable.

We argue that the above stylized facts indicate that there are persistent firm-specific factors that determine a firm's long-run size. Many theories explaining skewed firm-size distributions are at odds with some of these findings. In particular, an important body

of the theoretical literature predicts that such skewness can reasonably be explained with models where firms are ex ante homogeneous. This theoretical prediction results from the focus on firm dynamics. We present an alternative theory that emphasizes the role of firms' pre-entry size heterogeneity. The key idea is that firm-size heterogeneity upon entry, which is a product of a blueprint-generating process together with resource allocation across those blueprints, and subsequent firm-size dynamics, via Jovanovic-style learning, is more consistent with the empirical stylized facts for firm-size distributions. The model matches, first, the persistently skewed firm-size distribution over time; and second, the characteristics of firm size dynamics, in particular, that older firms are more stable, and second, that larger firms are more stable.

References

[1] Abbring, Jaap H. and Jeffrey R. Campbell. 2005. "A Firm's First Year," Working Paper, December.

[2] Axtell, Robert L. 2001. "Zipf Distribution of U.S. Firm Sizes," *Science*, vol. 293 (September), pp. 1818-20.

[3] Bernard, Andrew B., Stephen Redding, and Peter K. Schott. 2005. "Importers, Exporters, and Multinationals: A Portrait of Firms in the U.S. that Trade Goods," Working paper. Dartmouth.

[4] Cabral, Luis M.B. and Jose Mata. 2003. "On the Evolution of the Firm Size Distribution: Facts and Theory," *The American Economic Review*, September, pp. 1075-1090.

[5] Davis, Steven J., John Haltiwanger, Ron Jarmin, and Javier Miranda. 2006. "Volatility and Dispersion in Business Growth Rates: Publicly Trade versus Privately Held Firms," *NBER Macroeconomics Annual*.

[6] Gabaix, Xavier. 1999. "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, vol. 114, pp. 739-67.

[7] Gabaix, Xavier. 2005. "The Granular Origins of Aggregate Fluctuations," Working paper. MIT.

[8] Gabaix, Xavier and Augustin Landier. 2006. "Why Has CEO Pay Increased So Much?" Working paper. MIT.

[9] Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, vol. 60, pp. 1127-50.

[10] Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry," *Econometrica*, vol. 50, pp. 649-70.

[11] Jovanovic, Boyan and Peter L. Rousseau. 2001. "Why Wait? A Century of Life before IPO," *American Economic Review*, vol. 91, no. 2.

[12] Luttmer, Erzo G. J. 2005. "Growth and the Size Distribution of Firms," Working paper. Minneapolis Fed.

[13] Pakes, Ariel and R. Ericson. 1998. "Empirical Implications of Alternative Models of Firm Dynamics," *Journal of Economic Theory*, vol. 79, pp. 1-45.

[14] Rosen, Sherwin. 1981. "The Economics of Superstars," *American Economic Review*, vol. 71, no. 5.

[15] Rossi-Hansberg, Esteban and Mark L.J. Wright. 2004. "Firm Size Dynamics in the Aggregate Economy," Working Paper.

[16] Sutton, John. 1997. "Gibrat's Legacy," *Journal of Economic Literature*, vol. 35 (March), pp. 40-59.