The Growth and Failure of U. S. Manufacturing Plants

Timothy Dunne; Mark J. Roberts; Larry Samuelson

THE GROWTH AND FAILURE OF U. S. MANUFACTURING PLANTS*

TIMOTHY DUNNE
MARK J. ROBERTS
LARRY SAMUELSON

This paper examines the patterns of postentry employment growth and failure for over 200,000 plants that entered the U. S. manufacturing sector in the 1967–1977 period. The postentry patterns of growth and failure vary significantly with observable employer characteristics. Plant failure rates decline with size and age as do the growth rates of nonfailing plants. The expected growth rate of a plant, which depends on the net effect of these two forces, declines with size for plants owned by single-plant firms but increases with size for plants owned by multiplant firms.

I. INTRODUCTION

It is well recognized that a given change in total employment is composed of substantially larger shifts of individual workers between employment and unemployment as well as movements in and out of the labor force. This has given rise to substantial research on the decisions of workers to move between labor market states and the characteristics of workers that affect these choices. From this research has grown a greater appreciation of the important role that worker heterogeneity and the worker-employer matching process play in generating employment turnover.[1]

While much of the existing research on labor market turnover has focused on the supply side of the labor market, several recent empirical studies have begun to examine the turnover arising from fluctuations in labor demand. Leonard [1987] and Dunne, Roberts, and Samuelson [1989] examine the gross employment flows arising from the opening, expansion, contraction, and closing of individual

1. The distinction between net and gross employment flows of individuals is analyzed in Hall [1972] and Clark and Summers [1979]. The theoretical and empirical models of individual labor supply are reviewed in Pencavel [1986] and Killingsworth and Heckman [1986].

plants. A uniform finding of these studies is that the gross employ-
ment flows generated by employer turnover are substantially larger
than the resulting net change in employment. What is missing from
this research is the study of the employer's decision to enter, change
its size, or exit from the market and the characteristics of the
employer that affect this decision. These patterns of turnover
among heterogeneous employers are a necessary element in under-
standing the patterns of employment turnover.

    In this paper we examine the patterns of plant growth and exit
that underlie the gross employment flows for the U. S. manufactur-
ing sector for the 1967–1982 period. Attention is directed at iden-
tifying a set of plant characteristics that are systematically related
to the patterns of employer turnover. The findings are then used to
construct stylized employment histories for employers with dif-
ferent characteristics.

    The process of employer growth has been examined empiri-
cally by Hart and Prais [1956], Mansfield [1962], Hymer and
Pashigian [1962], Singh and Whittington [1975], Ijiri and Simon
[1977], Leonard [1985, 1987], Evans [1987a, 1987b], and Hall
[1987]. These papers focus primarily on the relationship between
the firm or plant's size, generally measured by employment, and
rate of growth. With the exception of Evans and Hall, they are
generally restricted by their data to analyzing the growth rates of
successful plants or firms and thus implicitly treat the growth
process as independent of plant or firm failure.

    Rather than being treated independently, growth and failure
are better viewed as outcomes of a single economic process of
industry or market development. We use the theoretical model of
Jovanovic [1982], which relies on employer heterogeneity and
market selection to generate patterns of employer growth and
failure, as the basis for the empirical work in this paper. The
empirical model is used to analyze the employment growth and
failure patterns of over 200,000 plants that entered the U. S.
manufacturing sector in the 1967, 1972, or 1977 Census of Manufac-
tures. The empirical work examines plant failure rates as well as the
mean and variance of the distribution of realized growth rates for
two groups of plants: all plants in operation at the beginning of each
time period and all plants in operation that do not fail.

    The empirical results indicate that plant size, age, and owner-
ship type, as well as interaction terms, are determinants of plant
growth and failure. Failure rates decline with increases in plant size
and age and differ with ownership type. If attention is limited to
successful plants, the mean growth rate declines with size. When

the failure probability is integrated into the analysis, however, the relationship between plant growth and size is negative for plants owned by single-plant firms but positive for plants owned by multiplant firms. These findings reflect a tradeoff between the rate of growth of nonfailing plants and probability of failing. For single-plant firms the decline of the growth rate of successful plants with size overwhelms the reduction in the failure rate. Expected growth rates then decline with size. For plants owned by multiplant firms, the decline in the failure rate with size is the dominant effect, and expected growth increases with size.

The remainder of this paper is divided into seven sections. Section II provides summary measures of plant growth rates that distinguish the growth of successful plants from the growth of all plants. Section III outlines the theoretical model of growth and failure and derives a set of testable predictions. Sections IV and V describe the data set, empirical model, and estimation method. The sixth and seventh sections report the empirical results and derive stylized plant histories for plants with different characteristics. The final section discusses some labor market implications of these results.

II. The Measurement of Plant Employment Growth

The analysis of plant employment growth begins with the discrete-time growth rate for individual plants.[2] Between period t and $t + 1$ this is defined as $g'_t = (S_{t+1} - S_t)/S_t$, where S_t is the plant's size in period t and S_{t+1} is the plant's optimal size in $t + 1$ if exit was not possible. Let $j(g'|x)$ be the probability density function for g' for plants with a given set of characteristics x. It will be referred to as the distribution of *potential* growth rates.

The theoretical model developed in the next section predicts that if a plant's growth rate falls below a critical level, denoted g^*, the plant will exit. If the plant exits, then its period $t + 1$ size equals zero, and the plant's observed or realized growth rate is -1. The *realized* growth rate g_t is then defined as

$$g_t = g'_t \quad \text{if} \quad g'_t \geq g^* \qquad \text{and} \qquad g_t = -1 \quad \text{if} \quad g'_t < g^*.$$

Let $f(g|x)$ be the density of realized plant growth rates for all plants with characteristics x in operation in period t *including* ones that

2. Many studies use continuous-time growth rates, which are measured as the difference in the logarithm of the firm or plant's size over time. Plants that fail have growth rates of $-\infty$. This convention makes it impossible to calculate and analyze the mean or variance of the distribution of realized growth rates across all plants.

exit before $t + 1$. The density $f(g|x)$ is defined over the interval $[-1, \infty)$ and will have a mass point at -1 representing the proportion of plants that fail between the two periods. The density $f(g|x)$ reflects the net effect of the interaction between the distribution of potential growth rates $j(g'|x)$ and the failure condition. Finally, let $h(g|x)$ be the density of realized growth rates for all nonfailing plants. The density $h(g|x)$ is $f(g|x)$ except with zero probability attached to the growth rate -1.

There is an explicit link between the means of $f(g|x)$ and $h(g|x)$. The mean growth rate of all successful plants, written as a function of x, is

$$(1) \quad \mu_h(x) \equiv \frac{\left(\int_{-1}^{\infty} gh(g|x)\,dg \right)}{P_s(x)}, \quad \text{where } P_s(x) \equiv \int_{-1}^{\infty} h(g|x)\,dg.$$

The denominator of (1) is the probability that a randomly drawn plant does not fail. Recognizing that $f(g|x)$ has a mass point at -1, the mean growth of all plants including failures is

$$(2) \quad \mu_f(x) \equiv \int_{-1}^{\infty} gf(g|x)\,dg = \mu_h(x)P_s(x) - (1 - P_s(x)).$$

The mean growth rate of all plants equals the mean growth rate of all successful plants, weighted by the probability of success, minus the probability of failure. $\mu_f(x)$ will always be less than or equal to $\mu_h(x)$, and the two will be equal only if there is no failure.[3]

Several earlier empirical studies have utilized data on successful plant growth rates to analyze how the mean and variance of $h(g|x)$ vary with plant or firm characteristics.[4] Equation (2) indicates that variation in μ_h with x is an accurate indicator of how μ_f varies with x if and only if plant failure rates do not vary with x. To see how they might differ, suppose that $\mu_h(x)$ diminishes with plant

---

3. A similar relationship can be developed between the variances of $f(g|x)$ and $h(g|x)$. The variance of the observed growth rates of all plants in operation in period t is

$$\sigma_f^2(x) \equiv \int_{-1}^{\infty} (g - \mu_f(x))^2 f(g|x)\,dg = \sigma_h^2(x)P_s(x) + (1 + \mu_h(x))^2 \, (P_s(x))(1 - P_s(x)).$$

4. Evans [1987a, 1987b] and Hall [1987] examine how the mean and variance of $j(g'|x)$ vary with plant or firm characteristics. They recognize and control for the fact that plant failure introduces sample selection bias in the distribution $h(g|x)$. They find that controlling for the sample selection bias in their data does not greatly affect their inferences about the relationship between plant characteristics and the growth process.

size, as virtually all studies have found. If the survival probability P_s (x) rises with an increase in size, it can offset the reduction in $\mu_h(x)$; and then $\mu_f(x)$ may increase, rather than decrease, with size. Large plants, on average, may thus grow faster than small plants even though large nonfailing plants grow more slowly than small nonfailing plants.

To illustrate the difference, Table I presents estimates of μ_h, μ_f, and plant failure rates for U. S. manufacturing plants.[5] Table Ia presents the average growth rate for all plants that do not fail, disaggregated by size and age. The values for most size and age categories are positive. They decline with plant size, for each of the three age categories, and with plant age, for every size class. Plant failure rates are reported in Table Ib. Small plants fail more often than large plants, and young plants fail more often than old plants. Table Ic reports the average realized growth rates for all plants, including failures. These mean growth rates are negative for all size and age categories, reflecting the fact that the total employment of each cohort of plants diminishes as the plants age. More importantly, the relationship between growth rates and size is U-shaped, indicating that growth rates initially decline and then increase with size. This holds in the aggregate for each age category and sharply contrasts with the pattern of successful plants. A similar U-shaped pattern appears for the effect of age on growth in total, although for most size classes growth increases with age.

The implication of the table is that failure rates are correlated with factors such as size and age which are also systematically correlated with the growth rates of successful plants. The interaction of these two factors produces the realized growth rate patterns for all plants.

III. A MODEL OF PLANT GROWTH

Many previous empirical studies follow Ijiri and Simon [1977] and begin with assumptions about the stochastic behavior of growth rates. The empirical model in this paper is based on the theoretical model of industry evolution developed by Jovanovic [1982]. This

---

5. The data pertain to all manufacturing plants with five or more employees which first appear in either the 1967, 1972, or 1977 Census of Manufactures. Each plant is classified according to its employment size and age category in year $t(t = 1967, 1972,$ and $1977)$ and then its employment growth rate from period t to $t + 1$ $(t + 1 = 1972, 1977,$ and $1982)$ is calculated.

TABLE I
PLANT GROWTH AND EXIT RATES

Age (years)	Size (number of employees)					
	5–19	20–49	50–99	100–249	>250	Total
a. Mean employment growth rate of successful plants						
1–5	0.606	0.299	0.187	0.132	0.067	0.446
6–10	0.338	0.136	0.066	0.011	−0.011	0.202
11–15	0.310	0.055	−0.006	−0.015	−0.018	0.153
Total	0.519	0.226	0.130	0.077	0.026	0.353
b. Plant exit rates						
1–5	0.412	0.396	0.390	0.327	0.229	0.397
6–10	0.347	0.268	0.281	0.245	0.158	0.303
11–15	0.304	0.206	0.234	0.212	0.131	0.255
Total	0.391	0.347	0.346	0.291	0.191	0.363
c. Mean employment growth rate of all plants						
1–5	−0.056	−0.216	−0.276	−0.238	−0.178	−0.129
6–10	−0.127	−0.169	−0.234	−0.236	−0.167	−0.162
11–15	−0.089	−0.163	−0.239	−0.224	−0.147	−0.141
Total	−0.074	−0.199	−0.261	−0.236	−0.170	−0.138
d. Number of plant-year observations on successful plants/failing plants						
1–5	75,959/53,325	29,938/19,649	13,758/8,794	9,472/4,601	3,281/977	132,408/87,346
6–10	27,409/14,569	15,268/5,584	7,577/2,961	5,829/1,889	2,630/494	58,713/25,947
11–15	7,773/3,400	4,675/1,216	2,198/673	1,568/421	911/137	17,125/5,847
Total	111,141/71,294	49,881/26,449	23,533/12,428	16,869/6,911	6,822/1,608	208,246/118,690

treats stochastic forces as determining plant characteristics and then exploits market behavior to generate plant growth rates.[6]
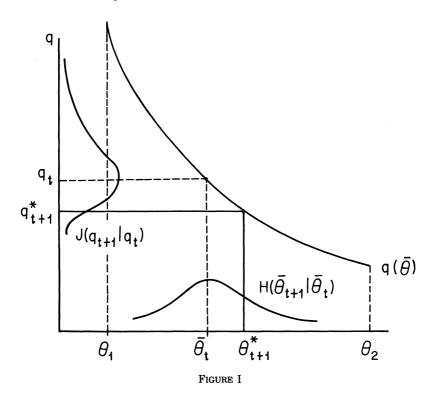
The key to the model is the presumption that plants have different efficiencies and hence different cost levels. A plant cannot observe its efficiency level and learns its cost only gradually through production. To capture these factors, let the cost of period t output level q_t for an arbitrarily chosen plant be given by $\gamma(q_t)\ \theta(c + \epsilon_t)$, where $\gamma(q_t)$ is a strictly convex cost function and $\theta(c + \epsilon_t)$ is a "cost multiplier" reflecting the fact that plants may be of differing efficiencies. Within θ, ϵ_t is a random variable with zero mean and c is an unknown parameter. In order to preclude the possibility of negative or arbitrarily high costs, θ is assumed to be increasing with $\lim_{z \to 0} \theta(z) = \theta_1 > 0$ and $\lim_{z \to \infty} \theta(z) = \theta_2 < \infty$. The parameter c is drawn randomly from a prior distribution $g_0(c)$. This prior is updated each period using Bayes rule given the observed value of $c + \epsilon_t$.

The industry is perfectly competitive, with a known time-path of output prices $\{p_t\}$. Plants are assumed to be risk neutral and to maximize expected profits. In each period t, plants begin with an expected value of the cost multiplier $\bar{\theta}_t$. They then choose output levels, draw values of the independent random variables ϵ_t, observe their profits, and update their expectations. Existing plants can then exit, or new plants can enter, by paying a one-time entry cost, and the process is repeated.

Jovanovic [1982] shows that in this model there exists a decreasing function $q_t(\bar{\theta}_t)$ expressing the plant's profit-maximizing output, if it produces in period t, given an expected cost multiplier of $\bar{\theta}_t$. One possible shape for this function, which may be either concave or convex, is illustrated in Figure I.[7] For each period t expectation $\bar{\theta}_t$, the information-updating process gives a density $H(\bar{\theta}_{t+1} | \bar{\theta}_t)$, which identifies the likely values of the period $t + 1$ cost multiplier. The density $H(\bar{\theta}_{t+1} | \bar{\theta}_t)$, which is represented in Figure I, has the property that $E(\bar{\theta}_{t+1} | \bar{\theta}_t) = \bar{\theta}_t$ [Feller, 1966]. $H(\bar{\theta}_{t+1} | \bar{\theta}_t)$ and $q_{t+1}(\bar{\theta}_{t+1})$ can then be used to derive a density $J(q_{t+1} | q_t)$ which

---

6. Simon's model is based on Gibrat's Law, which states that growth is independent of size. Leonard [1985] develops a model in which the evolution of plant size over time is governed by a process of partial adjustment toward a fixed optimal size with this process continually subjected to a random shock. The empirical studies by Evans [1987a, 1987b] are based on Jovanovic's model. Pakes and Ericson [1988] also develop growth models that rely on producer heterogeneity.

7. The concavity or convexity of $q_t(\bar{\theta}_t)$ depends on the first three derivatives of γ. See Jovanovic [1982, p. 652]. If the cost function is quadratic, then $q_t(\bar{\theta})$ is convex.

FIGURE I

summarizes the probabilities of different period $t + 1$ sizes given
that period t size equals q_t. This density is pictured on the vertical
axis of Figure I. Because the plant's discrete growth rate, given
period t size, is a linear transformation of q_{t+1}, the change of
variable theorem [Kmenta, 1986, p. 220] ensures that the density of
growth rates is identical to the density J with a relabeling of the
axes. In general, the shape of $J(q_{t+1}|q_t)$ depends on the curvature of
$q_{t+1}(\bar{\theta}_{t+1})$ as well as on the shape of $H(\bar{\theta}_{t+1}|\bar{\theta}_t)$.[8]

Figure I can be used to illustrate the way that period t size and
age affect the period $t + 1$ size of the plant. First, period t size
affects the location of $H(\bar{\theta}_{t+1}|\bar{\theta}_t)$ along the horizontal axis. Plants are
large in period t if $\bar{\theta}_t$ is small. This in turn affects the location of
$J(q_{t+1}|q_t)$ on the vertical axis and hence the likely period $t + 1$ size
of the plant. Second, age alters the variance of $H(\bar{\theta}_{t+1}|\bar{\theta}_t)$. Because
older plants have more previous observations on their cost level,

---

8. The shape of J determines the shape of the population distribution of
potential growth rates. In general, J will not be symmetric both because H is not
symmetric and because $q_{t+1}(\bar{\theta}_{t+1})$ is not linear.

new information causes smaller revisions in cost expectations. This reduces the variance of $H(\bar{\theta}_{t+1}|\bar{\theta}_t)$ and in turn reduces the variance of $J(q_{t+1}|q_t)$.

Jovanovic also shows that for each plant age there exists a failure boundary θ^*_{t+1}, which is illustrated in Figure I. Plants with cost expectations exceeding θ^*_{t+1}, and therefore optimal sizes less than q^*_{t+1}, will exit.[9] This critical size can be expressed as a critical growth rate, conditional on q_t, $g^*_t = (q^*_{t+1} - q_t)/q_t$. If the firm's optimal growth rate is less than g^*_t, it will exit and generate an observed growth rate of -1.

In general terms, the selection model predicts that both a plant's age and size will affect the failure rate and growth rate distribution. Some sign predictions can also be derived. First, failure rates should be a decreasing function of plant size. This occurs because the larger is $q_t(\bar{\theta}_t)$, the smaller is $\bar{\theta}_t$, and hence the less likely is $\bar{\theta}_{t+1}$ to exceed the failure boundary θ^*_{t+1}. Intuitively, a large plant is one that has received favorable cost information, and subsequent information is less likely to be unfavorable enough to induce exit. Second, failure rates should be a decreasing function of plant age. Older plants have relatively precise estimates $\bar{\theta}_t$, and this reduces the likelihood that $\bar{\theta}_{t+1}$ is substantially different from $\bar{\theta}_t$ and thus reduces the likelihood that $\bar{\theta}_{t+1}$ lies above the failure boundary.[10]

Some sign predictions concerning the effect of size and age on the mean and variance of $f(g|x)$ and $h(g|x)$ can also be derived from the selection model. Consider the mean and variance for nonfailing plants μ_h and σ^2_h. The mean growth rate of nonfailing plants should be an eventually decreasing function of current size, holding age fixed. This prediction relies on the fact that the efficiency level of nonfailing plants is bounded in the interval $[\theta_1, \theta^*]$. This implies that the plant's size is bounded and that as a plant increases in size

---

9. In general, the failure boundary θ^*_{t+1} will decrease with plant age. This occurs because expected profit is a convex function of the expected cost multiplier. With no change in market conditions, expected future profits are higher than current profits. As the plant ages, the difference between expected future profit and current profit diminishes because the increased precision of the plant's information causes the distribution of expected future profit to collapse around current profit. As a result, a given cost multiplier induces a lower future profit expectation as the plant ages. This would cause θ^*_{t+1} to decrease, and thus increase the failure rate, as plants age. Pakes and Ericson [1988] discuss conditions needed to sign this effect.

10. This is only true if the failure boundary does not decrease too rapidly with age. In the empirical results reported below, failure rates always decline with age which suggests that the former effect is dominant. To simplify the remaining discussion and predictions, we shall treat the failure boundary as if it were fixed with variations in plant age.

there is less (more) room for further increases (decreases).[11] A second prediction is that plant age should have a negative effect on σ_h^2 holding size fixed. This arises because, as a plant ages, the variance of $H(\bar{\theta}_{t+1}|\bar{\theta}_t)$ falls. This leads to smaller changes in plant size as the plant ages and thus a smaller variance in growth rates.[12] Predictions concerning the effect of size on σ_h^2 or age on μ_h depend on the convexity or concavity of $q_t(\bar{\theta}_t)$ and are ambiguous. Finally, implications of plant size and age for the mean and variance of $f(g|x)$ can be derived from equation (2) and the equation in footnote 3. Because they generally depend upon P_s, μ_h, and σ_h^2, they involve conflicting forces and thus yield ambiguous predictions.

IV. Data

This empirical analysis of plant growth and failure employs a new, comprehensive data set on individual plants in the U. S. manufacturing sector. The Census of Manufactures collects detailed data on each of the more than 300,000 U. S. manufacturing establishments in operation in each census year. The records of the individual manufacturing plants have recently been matched across the last five censuses, which cover the years 1963, 1967, 1972, 1977, and 1982, to provide a longitudinal data set that is useful for analyzing plant entry, growth, and exit.[13]

There are several unique aspects of this data set. First, the data provide complete coverage of the manufacturing sector, including small manufacturing plants. Because of data availability, most earlier growth rate studies were restricted to large firms. Second, the matching of plants across census years is based on plant identification numbers that allow us to identify many of the plants which undergo ownership changes. For these plants, ownership changes that simply change the name of the firm owning the plant will not be incorrectly measured as plant entry and exit. Finally, because of the multiple observations on a plant over time, it is possible to observe plants whose size in one census differs from their size in the initial census in which they appear. This makes it

11. Because this relies on the fact that plant size is bounded, it does not guarantee that the predicted negative relationship between size and the mean growth of nonfailing plants is monotonic.

12. A reduction in the variance of $H(\bar{\theta}_{t+1}|\bar{\theta}_t)$ ensures that the variance of $J(q_{t+1}|q_t)$ falls as long as $q_{t+1}(\bar{\theta}_{t+1})$ is not too nonlinear.

13. The Census of Manufactures covers all manufacturing establishments that have registered with the IRS to pay Social Security tax for their employees. Dunne, Roberts, and Samuelson [1988] discuss the construction of the data set in detail and provide summary measures.

possible to separately control for the effects of a plant's current size and its initial or entry size on its growth and failure rates.

Because the Census of Manufactures is only taken at five-year intervals, it is not possible to observe plants that enter and exit between census years. Plant failure rates constructed across adjoining census years will then underestimate the actual amount of plant failure. It is also not possible to measure the exact year of entry for any plant in the data set, but plants can be classified into age categories based on the census in which they first appear. In the empirical work age is treated as a categorical variable. Finally, it is not possible to identify all plant ownership changes that occur over time. This can result in a failure to identify some plants that continue in operation across census years, and this will lead to an overstatement of plant exit rates and the size of subsequent entry cohorts. For reasons discussed in detail in Dunne, Roberts, and Samuelson [1988], these matching errors will be concentrated among small plants owned by single-plant firms.

In order to minimize the effects of this potential measurement error, the smallest plants in the manufacturing sector are deleted from the analysis. The sample of data used in this paper is limited to manufacturing plants that entered in the 1967, 1972, or 1977 census and that have at least five employees in at least one census year.[14] There are a total of 219,754 different plants included. Because of the multiple time periods there are a total of 326,936 plant/year observations. In each census year plants with at least five employees account for approximately 64 percent of the total number of plants but over 98 percent of the total census year employment.[15]

---

14. Plants that have fewer than five employees are also handled differently in the data collection process. The data for most of these plants are not based on actual Census canvassing but rather are imputed from information reported by the Internal Revenue Service. This exclusion does not affect the implications of our analysis for gross employment flows because the included plants account for virtually all of a cohort's employment.

15. Once a plant enters the sample, it remains in the sample even if it falls below five employees. This is necessary to correctly measure plant exit rates. Otherwise, a plant whose size fluctuates around five employees over time would be recounted as entry or exit each time it shifted from one side of the five-employee cutoff to the other. Instead, to be counted as an exit, a plant's employment level should drop to zero. In general, small plants are imputed a number of employees equal to the ratio of plant payroll they report to the IRS and the average wage reported in the industry. While the reported figures should only apply to hired employees, they may include the plant owner. For this reason, any plant that falls to either zero or one employee is treated as an exit. A small number of plants, representing approximately 0.8 percent of all census establishments, are missing in one census year but reappear in later years. This can occur because of an error in matching plants across time or because the plant is closed in the missing years. These plants are included as entrants in the year in which they reappear if they have more than five employees.

V. Empirical Model

In this section an empirical model is developed to examine the patterns of employment growth and plant failure in U. S. manufacturing establishments.

1. Measurement of Plant Growth and Exit Rates

Observations on individual plants are grouped into cells based on the plant's current size, age, two-digit industry, year of observation, establishment ownership status, and initial size. The mean growth rate of plant employment, the variance of the growth rate, and the exit rate are measured for each cell using the observations on the numerous plants within the cell. The empirical model then examines how these summary measures of the growth rate distribution vary across cells.

Each plant in census year t is assigned to a cell based on the following set of characteristics:

(1) Three age categories: 1–5 years, 6–10 years, and 11–15 years. These are denoted as age categories 1, 2, and 3.

(2) Five current-size classes that measure a plant's employment in period t: 5–19 employees, 20–49 employees, 50–99 employees, 100–249 employees, and more than 250 employees.

(3) Twenty two-digit SIC manufacturing industries.

(4) Two ownership categories: single-unit plants (SU) and multiunit plants (MU). The former are plants owned by single-plant firms, while the latter are plants owned by multiplant firms.[16]

(5) Three initial-size classes that compare the plant's period t employment size class with its size class in its first year of observation. These identify the plant's initial size class as less than, equal to, or greater than its current size class.

For each data cell five sample statistics are constructed: the failure rate, the mean growth rate for all plants and for nonfailing plants, and the variance of growth rates for both all plants and nonfailing plants.[17] The failure rate for plants in cell c is constructed as $F_{ct} \equiv 1 - M_{ct}/N_{ct}$, where M_{ct} is the number of plants in the cell in

16. Plants which switch from single-unit to multiunit status over time are classified as multiunit plants throughout the entire time period. Across the last five Censuses of Manufactures there are 819,631 different manufacturing plants. Of these, 83.16 percent were always single-unit plants; 16.03 percent were always multiunit plants; and only 0.85 percent (6,891 plants) changed ownership status. Virtually all of the changes that do occur are from single- to multiunit status.

17. Plants in each of the three entry cohorts can be classified into 200 possible cells in the first time period in which they exist (20 industries, 1 age, 2 ownership

period t which survive until period $t + 1$ and N_{ct} is the total number of plants in period t.

Let g_{it}^c represent the discrete employment growth rate of establishment i in cell c between census years t and $t + 1$. The sample mean of the growth rates for all establishments in cell c between t and $t + 1$ is calculated as

$$\bar{g}_{ct} \equiv \left(\frac{1}{N_{ct}}\right) \sum_{i=1}^{N_{ct}} g_{it}^c,$$

where N_{ct} is the number of plants in the cell. If the plant fails before year $t + 1$, then $g_{it}^c = -1$. Similarly, the sample mean growth rate for all surviving plants can be constructed using only the growth rates for the M_{ct} surviving plants. The sample variance of the growth rates for all plants in cell c between t and $t + 1$ is

$$S_{ct}^2 \equiv \left(\frac{1}{N_{ct} - 1}\right) \sum_{i=1}^{N_{ct}} (g_{it}^c - \bar{g}_{ct})^2,$$

where failing plants again have $g_{it}^c = -1$. The variance of growth rates for nonfailing plants is similarly constructed from the growth rates of surviving plants.

2. Empirical Model

The five summary statistics provide consistent estimates of the corresponding population parameters for the distributions $f(g|x)$ and $h(g|x)$ within each data cell. A regression model is now developed to examine the across-cell pattern for each summary statistic Y_{ct} $(= \bar{g}_{ct}, S_{ct}^2, \text{or } F_{ct})$. The regression model is written as

$$(3) \qquad Y_{ct} = \sum_{i=1}^{60} \alpha_i D_i + \sum_{k=1}^{18} \beta_k D_k^c + \epsilon_{ct},$$

where D_i and D_k^c are dummy variables that distinguish cell characteristics or interactions among characteristics and ϵ_{ct} is a random disturbance with zero mean and variance σ_{ct}^2.

---

types, 5 current-size categories) and 520 possible cells in each subsequent time period (20 industries, 1 age, 2 ownership types, 13 current-size–initial-size categories). Of these 2,160 cells, 163 never contained any plants. Fifty cells have a failure rate equal to one, and forty-seven have no failure. In order to calculate the moments of the growth rate distributions described below, it is necessary to have at least four surviving plants and one failing plant per cell. There are a total of 1,627 cells: 703 for single-unit plants and 924 for multiunit plants, with sufficient observations. Together these cells cover 99.6 percent of the plant-year observations in the data set. This will be the group of data cells used in the empirical analysis developed below. On average, these cells for single-unit plants contain 326 plants, and these cells for multiunit plants contain 104 plants.

The dummy variables D_i $(i = 1 \ldots 60)$ represent each of the twenty industries in each of the three time periods. Industries undergo expansions and contractions at different points in time, and the complete set of industry-time dummies is needed to control for this variation. The dummy variables D_k^c $(k = 1 \ldots 18)$ include 14 dummy variables to represent the 15 current-size–age combinations. The theoretical model predicts nonlinear effects of size and age on growth as well as interactions between size and age. Modeling these interactions requires fourteen dummy variables rather than simply four dummies for the current-size categories and two for the age categories. The base category is a new plant in the smallest size class. The remaining four dummy variables are used to distinguish whether the current-size class of age group 2 and 3 plants (6–10 and 11–15 years) is greater than or less than their initial-size class. The theoretical model assumes that there are no adjustment costs so that all information about a plant's history is summarized by its current size. In practice, size adjustment may be slow due to the fixity of some inputs or financial constraints. The base group in each age category is the plants whose current and initial-size classes are equal.

The model is applied separately to the single-unit and multiunit plants. This recognizes that the importance of the learning process at the plant level, and thus its effect on plant growth and failure rates, may depend upon whether the firm which owns the plant has experience from operating other plants. Regression equation (3) is estimated using weighted least squares for each of the five summary measures of the growth rate distributions. The $\mathrm{var}(\epsilon_{ct})$ is not constant across cells but can be estimated using the multiple observations on plants within each cell.

Failure Proportion. When applied to failure proportions, the empirical model given by equation (3) is used with $Y_{ct} = F_{ct}$ and $\mathrm{var}(\epsilon_{ct}) = F_{ct}(1 - F_{ct})/N_{ct}$. Weighted least squares regression in this case is equivalent to the minimum chi-square estimation method [Maddala, 1983, p. 28].

Mean Growth Rates. Equation (3) is estimated with $Y_{ct} = \bar{g}_{ct}$. The within-cell variation in plant growth rates and number of plants is used to estimate $\mathrm{var}(\epsilon_{ct}) = S_{ct}^2/N_{ct}$. When examining the growth rates of successful plants, the denominator of $\mathrm{var}(\epsilon_{ct})$ is M_{ct} and the numerator is the estimated within-cell variance for nonfailing plants. The heteroskedasticity in this regression equation occurs because both σ_{ct}^2 and the number of observations in a cell vary across cells.

Variance of Growth Rates. Equation (3) is estimated with

$Y_{ct} = S_{ct}^2$. The var(ϵ_{ct}) is estimated as $(\mu_4^c - ((N_{ct} - 3)/(N_{ct} - 1))$ $(S_{ct}^2)^2)$, where μ_4^c is the estimated fourth central moment of g_{it}^c. This is equal to $(1/N_{ct})(\Sigma g_{it}^4 - 4\bar{g}_{ct}(\Sigma g_{it}^3) + 6(\bar{g}_{ct}^2)(\Sigma g_{it}^2) - 4(\bar{g}_{ct}^3)(\Sigma g_{it}) + N_{ct}\bar{g}_{ct}^4)$, where all summations are over N_{ct} [Wilks, 1962, pp. 76 and 199]. Similarly, summations are over M_{ct} when examining nonfailing plants.

3. Comparison with Alternative Estimation Methods

The methodology employed here groups plants into data cells and uses the multiple observations within each cell to consistently estimate parameters of the distributions of realized growth rates for all plants $f(g|x)$ and for nonfailing plants $h(g|x)$ within the cell. A regression model is then used to summarize the across-cell variation in these estimates. The basic assumption is that all plants within a cell are homogeneous up to a random disturbance with a zero mean and constant, cell-specific variance. No assumption is made about the specific form of the distribution of potential growth rates $j(g'|x)$.

An alternative methodology for examining the mean of $f(g|x)$ and $h(g|x)$ is available which is based on Tobit estimation [Tobin, 1958] of the distribution of potential growth rates $j(g'|x)$. As developed by Maddala [1983, Ch. 6], Heckman [1979], and Amemiya [1984], $j(g'|x)$ could be treated as a latent-variable distribution. When combined with the failure rule, this generates a censored sample of plant growth rates. By making specific assumptions about the form of the latent variable distribution, the conditional mean of $j(g'|x)$ can be estimated. This is the procedure used by Evans [1987a, 1987b] and Hall [1987] to study firm growth rates. The effect of variations in x on the mean of $f(g|x)$ and $h(g|x)$ can then be examined using the information gained from estimation of the parameters of the latent variable distribution [Maddala, 1983, p. 158]. This procedure requires a distributional assumption for the latent variable. The usually invoked assumption of normality is inappropriate for this study and would result in inconsistent parameter estimates because the population distribution of discrete growth rates is highly skewed and, as a result of the truncation of plant sizes at zero, is truncated at -1. While the Tobit model can be used with latent variable distributions other than the Normal, the details are less well developed [Lee, 1982; Maddala, 1983, p. 187].

The grouped data techniques used here provide several advantages over the standard Tobit model for examining the distributions of realized growth rates. Distributional assumptions are avoided. A great degree of (piecewise) nonlinearity is allowed in the

failure rates and the mean and variance of the observed values. This allows us to avoid the difficult issue—which arises in the Tobit framework—of separating sample selection, heteroskedasticity, and the nonlinear effects of the explanatory variables on the conditional mean of the latent variable distribution. There are two costs to using this procedure. First, as with any grouping procedure, there is a loss of information. However, in this case all of the explanatory variables except the current and initial size of the plant are qualitative variables, and there is no information lost on these qualitative explanatory variables by grouping. Second, the procedure does not allow identification of the parameters of the latent variable distribution $j(g'|x)$ separately from the selection rule. However, given our focus on the plant growth patterns that underlie the observed employment flows, knowledge of the latent variable distribution is not necessary.

VI. EMPIRICAL RESULTS

This section reports results of the estimation of equation (3). Table II reports the regression coefficients for plant failure rates. All coefficients represent deviations from an establishment appearing in its first census with a size between 5 and 20 employees. Separate coefficients are reported for regressions on plants owned by single-plant and multiplant firms. For plants owned by multi-plant firms, examined in column two of the table, two patterns are present. First, within any age category the failure rate declines with increases in the plant's current size. This can be seen by noticing that the age–current-size coefficients decrease with current size while holding age fixed. The second pattern is that the failure rate declines with a plant's age, holding current size fixed. In every size class the regression coefficients decline as the plant ages, indicating less failure. Examining the results for single-unit plants, the negative effect of plant age on the failure rate is again present. However, the effect of plant size on the failure rate is less apparent for these plants. In general, there is a downward trend in the failure rate with increasing plant size, particularly for age groups 2 and 3, but it is not monotonic.[18]

---

18. Inferences concerning the largest size class (≥ 250 employees) for single-unit plants are less precise because there are relatively few data cells with this combination of characteristics. Of the 703 data cells for these plants, only 19 pertain to the largest size class. This compares with between 130 and 215 cells for each of the other four size classes. Higher standard errors will be seen on the coefficients for large, single-unit plants throughout the remainder of this paper.

TABLE II
REGRESSION COEFFICIENTS FOR PLANT FAILURE RATES: 1967, 1972, 1977 ENTRANTS
(STANDARD ERRORS IN PARENTHESES)

		Single-plant	Multiplant
Intercept[a]		0.426	0.471
Age/initial-size (IS) versus current-size (CS) category			
2	IS > CS	0.084 (0.012)*	0.104 (0.010)*
2	IS < CS	0.008 (0.010)	0.001 (0.007)
3	IS > CS	0.043 (0.020)	0.065 (0.021)
3	IS < CS	0.002 (0.016)	−0.013 (0.012)
Age/current-size category			
1	2	−0.035 (0.005)*	−0.064 (0.006)*
1	3	−0.002 (0.009)	−0.116 (0.006)*
1	4	0.055 (0.014)	−0.206 (0.007)*
1	5	−0.039 (0.056)	−0.306 (0.008)*
2	1	−0.061 (0.005)*	−0.070 (0.009)*
2	2	−0.166 (0.008)*	−0.158 (0.008)*
2	3	−0.151 (0.012)*	−0.206 (0.008)*
2	4	−0.095 (0.019)*	−0.271 (0.008)*
2	5	−0.174 (0.097)	−0.356 (0.008)*
3	1	−0.087 (0.008)*	−0.150 (0.018)*
3	2	−0.199 (0.013)*	−0.238 (0.015)*
3	3	−0.180 (0.021)*	−0.250 (0.015)*
3	4	−0.099 (0.037)	−0.315 (0.014)*
3	5	−0.193 (0.179)	−0.386 (0.014)*

Age categories: 1 = 1–5 years; 2 = 6–10 years; 3 = 11–15 years.
Size categories: 1 = 5–19 employees; 2 = 20–49 employees; 3 = 50–99 employees; 4 = 100–249 employees; 5 = ≥250 employees.
a. The reported intercept is the average value of the 60 estimated industry-time intercepts.
*Significant at the 0.05 level using Leamer's]1978] correction for sample size.

The pattern of age and size effects on failure rates can be examined more closely by testing several restricted versions of the model.[19] The first hypothesis is that a plant's current size has no effect; the second is that plant age has no effect; the final hypothesis is that there are no age-size interactions. The results of the tests are reported in the first column (single-unit plants) and sixth column (multiunit plants) of Table III. They indicate that current size and

19. All hypothesis tests employ the standard F-statistic based on the proportional difference in the restricted and unrestricted sum of squared residuals. Because of the large sample sizes, Leamer's [1978, p. 114] correction to the critical value of the F-statistic is adopted. Given the sample sizes used here, 703 observations for single-unit plants and 924 for multiunit plants, the correction results in an increase in the conventional 0.05 critical value by a factor of six or seven. This correction is also made when reporting the significance of individual regression coefficients in Tables II, IV, and V.

TABLE III
HYPOTHESIS TEST STATISTICS

| | | Single-unit plants | | | | | Multiunit plants | | | | | |
| | | Successful plants | | | All plants | | | Successful plants | | | All plants | |
Hypothesis	Number of restrictions	F_{ct}	$\bar{g}_t^c$	S_{ct}^2	$\bar{g}_t^c$	S_{ct}^2	F_{ct}	$\bar{g}_t^c$	S_{ct}^2	$\bar{g}_t^c$	S_{ct}^2
No current-size effect	12	26.06*	84.36*	89.19*	89.02*	110.44*	234.35*	66.38*	26.72*	11.47*	53.00*
No age effect	10	59.09*	47.54*	47.65*	8.90	58.97*	45.55*	40.67*	31.02*	8.01	41.50*
No current-size–age interaction	8	11.24	7.91	26.74*	9.61	30.86*	4.67	24.85*	17.93*	8.63	23.08*
No initial-size effect	4	13.20	6.74	39.52*	10.45	16.51*	30.37*	1.47	15.01	11.26	2.49
No industry effects	57	35.51*	16.87*	8.04	27.34*	7.54	18.07*	13.56*	9.58	15.94*	5.67
No year effects	40	10.62*	11.14*	4.31	17.68*	6.65	13.64*	8.67	2.48	11.62*	2.74

*Hypothesis is rejected at the 0.05 significance level using Leamer's [1978] correction for sample size.

age both have significant effects on the failure rate but there are no significant interaction effects. In summary, failure rates are lower for older plants, regardless of ownership type, and for larger plants, particularly those owned by multiplant firms. The effects of size and age are consistent with the theoretical selection model.

In addition to age and size effects, variables to distinguish the initial size and current size of the plant as well as industry-time effects are included as control variables. The parameter estimates reported in Table II indicate that plants in age group 2 (6 to 10 years old) which are in a smaller size class than their entering size class have failure rates that are between eight and ten percentage points higher than plants of comparable age that remain in their initial size class. There is no evidence of lower failure rates for plants that have expanded beyond their initial size class.

The significant effect of larger initial size indicates that there may be some role for plant history in explaining current failure rates. For example, a large initial size or employment level may signal that the plant has installed a relatively large capital stock and hence has relatively high fixed costs which cannot be quickly altered. A plant with relatively high fixed costs is more likely to fail. This also provides some evidence to support Jovanovic's [1982] suggestion that the fixity of capital as well as the selection mechanism may affect the growth and failure process.

The final inference on failure rates concerns the importance of both industry and year-specific factors. The test statistics reported in Table III indicate that any attempt to restrict the full set of 60 industry-year intercepts for either single- or multiunit plants is rejected. In particular, industry differences are found to be important. Many previous studies do not control for these factors.

Table IV reports the regression coefficients for the mean and variance of plant growth rates using only the information on nonfailing plants. The first two columns report coefficients for the mean growth rate of nonfailing plants; while the third and fourth columns report results for the variance of plant growth rates. The pattern found in most previous studies—mean growth rates declining with size—is strongly evident for both the single-unit and multiunit plants. Within each age category, growth declines monotonically across virtually every size class as predicted by the selection model. In addition, mean growth rates decline monotonically with increases in plant age for virtually every size class. The only exception is for the largest plants. They show small, insignificant increases in growth rates as they move from age group 2 to 3.

TABLE IV

REGRESSION COEFFICIENTS FOR SUCCESSFUL PLANTS: 1967, 1972, 1977 ENTRANTS (STANDARD ERRORS IN PARENTHESES)

	Mean growth rate		Variance of growth rate	
	Single-plant	Multiplant	Single-plant	Multiplant
Intercept[a]	0.492	0.709	1.650	2.446
Age–initial-size (IS) versus current-size (CS) category				
2 IS > CS	−0.038 (0.025)	0.010 (0.018)	−0.103 (0.014)*	−0.052 (0.010)*
2 IS < CS	0.061 (0.014)	0.026 (0.011)	0.080 (0.012)*	0.015 (0.008)
3 IS > CS	−0.023 (0.037)	0.009 (0.029)	−0.078 (0.020)	−0.056 (0.012)
3 IS < CS	0.045 (0.022)	0.004 (0.017)	0.057 (0.015)	0.004 (0.007)
Age–current-size category				
1 2	−0.292 (0.013)*	−0.360 (0.034)*	−1.166 (0.080)*	−1.812 (0.287)*
1 3	−0.448 (0.016)*	−0.504 (0.032)*	−1.383 (0.078)*	−1.939 (0.284)*
1 4	−0.551 (0.023)*	−0.598 (0.032)*	−1.488 (0.078)*	−2.072 (0.284)*
1 5	−0.664 (0.054)*	−0.683 (0.033)*	−1.651 (0.079)*	−2.167 (0.284)*
2 1	−0.229 (0.014)*	−0.432 (0.039)*	−1.110 (0.083)*	−2.037 (0.285)*
2 2	−0.450 (0.015)*	−0.575 (0.032)*	−1.465 (0.078)*	−2.117 (0.284)*
2 3	−0.550 (0.018)*	−0.636 (0.032)*	−1.545 (0.077)*	−2.148 (0.284)*
2 4	−0.579 (0.024)*	−0.680 (0.032)*	−1.640 (0.077)*	−2.174 (0.283)*
2 5	−0.587 (0.074)*	−0.717 (0.032)*	−1.715 (0.082)*	−2.202 (0.283)*
3 1	−0.310 (0.021)*	−0.517 (0.054)*	−1.257 (0.086)*	−2.130 (0.285)*
3 2	−0.529 (0.020)*	−0.664 (0.036)*	−1.547 (0.078)*	−2.208 (0.284)*
3 3	−0.628 (0.027)*	−0.663 (0.036)*	−1.602 (0.078)*	−2.205 (0.284)*
3 4	−0.681 (0.045)*	−0.664 (0.035)*	−1.616 (0.080)*	−2.220 (0.284)*
3 5	−0.506 (0.269)	−0.683 (0.036)*	−1.472 (0.234)*	−2.248 (0.284)*

Age categories: 1 − 1–5 years; 2 − 6–10 years; 3 − 11–15 years.
Size categories: 1 − 5–19 employees; 2 − 20–49 employees; 3 − 50–99 employees; 4 − 100–249 employees; 5 − ≥250 employees.
a. The reported intercept is the average value of the 60 estimated industry-time intercepts.
*Significant at the 0.05 level using Leamer's [1978] correction for sample size.

The hypothesis test statistics reported in Table III indicate that age, size, and industry have significant effects on growth patterns. Age-size interactions are also important for multiunit plants. Initial plant size has no significant effect in these regressions.

While the pattern of size and age effects is virtually identical for plants owned by single-plant and multiplant firms, the levels differ. This can be seen by comparing the intercept and age-size coefficients within each ownership group. For single-unit plants the average growth rate of nonfailing plants is negative for the larger size classes and older plants. For single-unit plants in the oldest age category, only the smallest size class (5–19 employees) will have a positive average growth rate for successful plants. In contrast, the average growth rate of successful multiunit plants is positive for all but the age = 2/current size = 5 category. On average, the patterns of employment expansion or contraction for continuing plants will thus differ with ownership type.

The regression results for the variance of the growth rates are reported in the final two columns of Table IV. For both ownership types the variance declines with both size and age. For the multiunit plants differences in the variance across size classes are virtually eliminated as the plants age, while small size effects persist as single-unit plants age. The theoretical selection model predicts the negative effect of age and gives ambiguous sign predictions for the effect of size. The test statistics in columns three and eight of Table III indicate that size, age, and size-age interactions are significant factors; industry and year effects are not.

These results can be summarized simply. The failure rate, mean growth rate of nonfailing plants, and variance of the growth rate of nonfailing plants decline with size and age for both single-unit and multiunit plants. The mean growth rate for nonfailing, single-unit plants in the larger size classes is negative which contrasts with the positive mean growth rates of nonfailing multiunit plants across virtually all size and age classes.

The regression coefficients for the mean and variance of growth rates over all plants are reported in Table V. Focusing first on single-unit plants, mean growth declines with size for all but the largest size class in age groups 2 and 3. The mean growth rate increases between size classes 4 and 5 for age groups 2 and 3, although the difference is not statistically significant. In general, the mean growth rate of a randomly chosen single-unit plant will thus be negative and decline with increases in the plant's size. This negative net effect appears because the decline in the mean growth

TABLE V
REGRESSION COEFFICIENTS FOR ALL PLANTS: 1967, 1972, 1977 ENTRANTS (STANDARD ERRORS IN PARENTHESES)

		Mean growth rate		Variance of growth rate	
		Single-plant	Multiplant	Single-plant	Multiplant
Intercept[a]		−0.126	−0.051	1.503	2.09
Age–initial-size (IS) versus current-size (CS) category					
2	IS > CS	−0.106 (0.022)*	−0.104 (0.017)*	−0.073 (0.024)	−0.001 (0.016)
2	IS < CS	0.043 (0.014)	0.012 (0.011)	0.094 (0.015)	0.033 (0.011)*
3	IS > CS	−0.054 (0.036)	−0.045 (0.034)	−0.016 (0.039)	−0.007 (0.026)
3	IS < CS	0.037 (0.023)	0.006 (0.019)	0.061 (0.020)	0.002 (0.016)
Age–current-size category					
1	2	−0.143 (0.010)*	−0.132 (0.020)*	−0.877 (0.048)*	−1.197 (0.155)*
1	3	−0.278 (0.013)*	−0.168 (0.019)*	−1.093 (0.047)*	−1.452 (0.151)*
1	4	−0.415 (0.017)*	−0.143 (0.019)*	−1.235 (0.047)*	−1.606 (0.150)*
1	5	−0.458 (0.056)*	−0.116 (0.021)*	−1.400 (0.054)*	−1.730 (0.150)*
2	1	−0.072 (0.011)*	−0.188 (0.025)*	−0.801 (0.053)*	−1.550 (0.153)*
2	2	−0.118 (0.013)*	−0.173 (0.020)*	−1.184 (0.047)*	−1.595 (0.150)*
2	3	−0.212 (0.017)*	−0.170 (0.020)*	−1.274 (0.047)*	−1.681 (0.150)*
2	4	−0.316 (0.025)*	−0.144 (0.020)*	−1.332 (0.049)*	−1.734 (0.150)*
2	5	−0.234 (0.100)*	−0.097 (0.021)	−1.401 (0.073)*	−1.812 (0.150)*
3	1	−0.090 (0.018)*	−0.152 (0.041)	−0.968 (0.058)*	−1.663 (0.157)*
3	2	−0.149 (0.020)*	−0.159 (0.027)*	−1.279 (0.049)*	−1.704 (0.151)*
3	3	−0.248 (0.028)*	−0.144 (0.027)*	−1.350 (0.049)*	−1.732 (0.151)*
3	4	−0.360 (0.045)*	−0.085 (0.026)	−1.353 (0.055)*	−1.774 (0.151)*
3	5	−0.156 (0.241)	−0.058 (0.027)	−1.205 (0.216)*	−1.828 (0.151)*

Age categories: 1 = 1–5 years; 2 = 6–10 years; 3 = 11–15 years.
Size categories: 1 = 5–19 employees; 2 = 20–49 employees; 3 = 50–99 employees; 4 = 100–249 employees; 5 = ≥250 employees.
[a] The reported intercept is the average value of the 60 estimated industry-time intercepts.
*Significant at the 0.05 level using Leamer's [1978] correction for sample size.

rate of nonfailing plants with increased size overwhelms the reduction in the failure rate with increased size.

The size pattern for multiunit plants is quite different. For young plants the pattern is U-shaped with mean growth rates declining with size for plants less than 100 employees and increasing with size for plants over 100 employees. For plants over five years old mean growth increases with plant size. This latter pattern occurs because the reduction in the failure rate as we move to larger plants outweighs the reduction in the growth rate of successful plants. On average, the expected growth of a multiunit plant will increase with size.

The net effect of age on plant growth is less uniform than the effects of size. For single-unit plants mean growth declines with age for the smallest group of plants but increases and then declines with age for larger plants. For multiunit plants with less than 100 employees, mean growth initially declines with age but then increases. Multiunit plants with over 100 employees show an increase in mean growth rates with age. However, based on the test statistics in Table III, these age effects are not statistically significant for either group of plants.

In summary, the most striking characteristic of the results for all plants is the effect of ownership status. Plants owned by single-plant firms have mean growth rates that decline with size and, to a lesser extent, with age. This occurs because the reduction in failure rates with size and age is too small to overcome the reduction in the growth rate of the nonfailing plants. The opposite pattern characterizes plants owned by multiplant firms. Mean growth rates for these plants tend to increase with size and age because a substantial reduction in their failure rate overwhelms a more modest reduction in the growth of nonfailing plants.

The final set of inferences relates to the variance of growth rates over all plants. Here the pattern of size and age effects is relatively simple. The variance for both single and multiunit plants tends to decline with size and age. The test statistics reported in Table III indicate that size, age, and size-age interactions are significant factors.

VII. Stylized Plant Histories

The patterns of plant failure and growth rates of nonfailing plants can be used to derive an expected employment history for an average manufacturing plant in each ownership and initial-size

category. After choosing an initial employment size for the plant, the growth rates in Table IV can be used to construct the expected size at ages 5, 10, and 15 years of a plant that does not fail. In addition, the cumulative survival rate for these plants as they age can be constructed using the information in Table II.

Table VI reports these average histories for manufacturing plants with five different initial sizes. The initial plant sizes were chosen to be approximately in the middle of each size class used in the regression models. Within the group of single-unit plants, two patterns are of interest. First, the differences in the 15-year survival rates are fairly minor across plants with different initial sizes. Between approximately one quarter and one third of the plants in each size class survive 15 years. Second, the net employment change after 15 years, for the plants that survive, decreases substantially with initial plant size. On average, surviving small single-unit plants expand while large ones decline over the period.

The pattern of postentry growth and survival is substantially different for plants owned by multiplant firms. First, the 15-year survival rate increases substantially with plant size. It varies from a low of 0.278 for a plant that begins with 15 employees to a high of 0.676 for a 400-employee plant. Second, while the growth rates for successful multiunit plants vary substantially with initial size, the resulting net change in plant employees does not. The net increase

TABLE VI
AVERAGE PLANT HISTORIES

Initial plant employment	Employment of surviving plants at age (cumulative survival rate)			Net employment change of surviving plants over 15 years
	5 years	10 years	15 years	
I. Single-unit plants				
15	22 (0.574)	23 (0.421)	23 (0.325)	8
35	42 (0.609)	44 (0.447)	42 (0.346)	7
75	78 (0.576)	74 (0.418)	64 (0.315)	−11
175	165 (0.519)	150 (0.347)	122 (0.234)	−53
400	331 (0.613)	300 (0.459)	296 (0.352)	−104
II. Multiunit plants				
15	26 (0.529)	29 (0.363)	30 (0.278)	15
35	47 (0.593)	54 (0.436)	56 (0.339)	21
75	90 (0.645)	97 (0.474)	101 (0.369)	26
175	194 (0.735)	200 (0.588)	209 (0.496)	34
400	410 (0.835)	407 (0.739)	418 (0.676)	18

in the number of employees after 15 years varies from 15 to 34 across multiunit plants with different initial sizes. This reflects the fact that small growth rates for surviving large plants can have as substantial an impact on the level of employment as large growth rates for surviving small plants.

A final interesting pattern present in Table VI is based on a comparison of plants with different ownership types but similar initial sizes. Small multiunit plants have lower survival rates but higher growth rates if they survive than single-unit plants have. Large multiunit plants, however, have both substantially higher survival rates and growth rates if they survive than comparable single-unit plants have. The postentry employment experience of large plants thus varies substantially with ownership type.

VIII. CONCLUSION

The fluctuations in labor supply arising from the movement of workers between employment and unemployment as well as the movement in and out of the labor force have been extensively studied. The analogous fluctuations in labor demand arising from the entry, growth, and exit of individual employers have not. In this paper we examine the patterns of employment growth and exit for plants in the U. S. manufacturing sector over the 1967–1982 period.

We find that the patterns of plant growth and failure are systematically related to the plant's age, size, and ownership type. In general, both the failure rate and the growth rate of nonfailing plants decline with age. There is, however, variation in this pattern with plant size and ownership type. Relative to small plants, large plants have lower failure rates and lower growth rates if they survive. Large multiunit plants have both lower failure rates and higher growth rates if successful than large single-unit plants have. The latter plants have negative average growth rates even when they survive.

These empirical results have several implications for the way we view labor market turnover. First, the U. S. manufacturing sector is characterized by the continual turnover of large numbers of employers. This turnover occurs even after controlling for industry and time-specific factors, such as fluctuations in the demand for the industry's output. Second, the turnover is not randomly distributed across all producers but rather systematically related to observable producer characteristics including size, age, and owner-

ship type. The turnover patterns are generally consistent with those predicted by the theoretical model of producer heterogeneity, learning, and market selection developed by Jovanovic [1982].

This underlying producer turnover results in the continual loss of existing employment opportunities as well as the creation of new positions, even with little or no change in the total demand for labor. This is evidenced by the fact that the gross employment flows resulting from producer turnover are substantially larger than the resulting net change in employment. This suggests that accounting for the turnover of employment positions may be valuable in explaining both employment fluctuations and the duration of employment spells.[20]

To see the implication for the duration of employment spells, suppose, for example, that there are two types of employers, one of which has a higher probability of surviving than the other has. This heterogeneity can give rise to a labor market in which significant employment turnover coexists with a large number of long-duration employment opportunities. This occurs because the plant turnover takes the form of relatively high turnover rates that are concentrated in a small subset of employers. The continual entry and exit of the high failure probability employers generates employment turnover, while the high survival rate of the low failure probability employers ensures that many employment opportunities are of long duration. As a result, information on turnover rates alone will not provide an accurate indication of the stability of employment opportunities.[21]

The systematic variation in the expected duration of employment opportunities with observable employer characteristics has implications for the worker-employer matching process. If worker quits are costly, then employers who have a high probability of surviving have incentives to reduce their turnover costs by searching for workers who are less likely to quit and by offering wage premiums to secure these workers. As a result, firms that provide relatively long-lived employment opportunities may offer relatively high wages. Given the finding of this paper that exit is more highly concentrated in smaller plants, this may help explain the observed

---

20. Leonard [1985, 1987] makes a similar argument and provides some estimates of the effect of employer turnover on unemployment patterns.
21. Dunne and Roberts [1987] develop these implications in greater detail. They estimate that, despite high rates of employment turnover, 54.7 percent of the manufacturing positions present in 1982 will have a completed length of at least 20 years.

positive correlation between plant size and wage rates [Brown and Medoff, 1988]. Similarly, if workers differ in the value they place on having a long-term stable employment opportunity, because, for example, their opportunity cost of being unemployed varies, then systematic sorting of workers and employers based on the probability the employer survives could occur.

Finally, the results are also relevant to the debate on the importance of new small businesses as the source of most new job creation. Previous studies have observed that small plants grow faster than large ones, and this has led to arguments that policies designed to achieve employment growth should target small plants and, more specifically, the formation of new small plants. The results reported here reveal that small plants grow faster than large ones, if they survive, but that they are much less likely to survive. As shown in Table Ib, the average failure rate for plants with 5–19 employees is 12.7, 34.4, and 104.7 percent higher than for plants with 20–99, 100–249, and more than 250 employees, respectively. In addition, the results reported here indicate that small plants owned by multiplant firms generally have higher rates of employment growth than similar plants owned by single-plant firms. The relevant question is thus not whether successful small entrants have faster rates of employment growth than larger, older plants but whether the growth rates of the small, often single-plant entrants, are large enough to compensate for their higher attrition rates. Finally, if the selection model captures the important aspects of the growth and failure process, as the results suggest, then policies designed to encourage the establishment of new plants may simply elevate the level of small plant failure if the policy-induced entrants are candidates who are less likely to succeed than their older competitors.

CENTER FOR ECONOMIC STUDIES, BUREAU OF THE CENSUS
PENNSYLVANIA STATE UNIVERSITY
PENNSYLVANIA STATE UNIVERSITY

REFERENCES

Amemiya, Takeshi, "Tobit Models: A Survey," *Journal of Econometrics,* XXIV (1984), 3–61.
Brown, Charles, and James Medoff, "The Employer Size Wage Effect," mimeo, Harvard University, 1988.
Clark, Kim B., and Lawrence H. Summers, "Labor Market Dynamics and Unemployment: A Reconsideration," *Brookings Papers on Economic Activity* (1979), 13–60.

Dunne, Timothy, and Mark J. Roberts, "The Duration of Employment Opportunities in U. S. Manufacturing," Department of Economics, Pennsylvania State University, 1987.
——, ——, and Larry Samuelson, "Patterns of Firm Entry and Exit in U. S. Manufacturing Industries," *Rand Journal of Economics,* XIX (1988), 495–515.
——, ——, and ——, "Plant Turnover and Gross Employment Flows in the U. S. Manufacturing Sector," *Journal of Labor Economics,* VII (1989), 48–71.
Evans, David S., "The Relationship Between Firm Growth, Size and Age: Estimates for 100 Manufacturing Industries," *Journal of Industrial Economics,* XXXV (1987a), 567–82.
——, "Tests of Alternative Theories of Firm Growth," *Journal of Political Economy,* XCV (1987b), 657–74.
Feller, William, *An Introduction to Probability Theory and Its Applications,* Vol. 2 (New York: John Wiley and Sons Inc., 1966).
Hall, Bronwyn, "The Relationship Between Firm Size and Firm Growth in the U. S. Manufacturing Sector," *Journal of Industrial Economics,* XXXV (1987), 583–606.
Hall, Robert E., "Turnover in the Labor Force," *Brookings Papers on Economic Activity* (1972), 709–56.
Hart, P. E., and S. J. Prais, "The Analysis of Business Concentration: A Statistical Approach," *Journal of the Royal Statistical Society,* CXIX (1956), 150–91.
Heckman, James J., "Sample Selection Bias as a Specification Error," *Econometrica,* XLVII (1979), 153–61.
Hymer, Stephen, and Peter Pashigian, "Firm Size and the Rate of Growth," *Journal of Political Economy,* LXX (1962), 556–69.
Ijiri, Yuji, and Herbert A. Simon, *Skew Distributions and the Sizes of Business Firms* (Amsterdam: North-Holland Publishing Company, 1977).
Jovanovic, Boyan, "Selection and Evolution of Industry," *Econometrica,* (1982), 649–670.
Killingsworth, Mark, and James J. Heckman, "Female Labor Supply: A Survey," in *Handbook of Labor Economics,* Orley Ashenfelter and Richard Layard, eds. (Amsterdam: North-Holland Publishing Company, 1986).
Kmenta, Jan, *Elements of Econometrics,* second edition (New York: Macmillan Publishing Company, 1986).
Leamer, Edward E., *Specification Searches* (New York: John Wiley and Sons, 1978).
Lee, Lung Fei, "Some Approaches to the Correction for Selectivity Bias," *Review of Economic Studies,* XLIX (1982), 355–72.
Leonard, Jonathan S., "On the Size Distribution of Employment and Establishments," mimeo, School of Business Administration, University of California, Berkeley, 1985.
——, "In the Wrong Place at the Wrong Time: The Extent of Frictional and Structural Unemployment," in *Unemployment and the Structure of Labor Markets,* Kevin Lang and Jonathan S. Leonard, eds. (New York: Basil Blackwell, 1987).
Maddala, G. S., *Limited-Dependent and Qualitative Variables in Econometrics* (Cambridge, England: Cambridge University Press, 1983).
Mansfield, Edwin, "Entry, Gibrat's Law, Innovation and the Growth of Firms," *American Economic Review,* LII (1962), 1031–51.
Pakes, Ariel, and Richard Ericson, "Empirical Implications of Alternative Models of Firm Dynamics," Social Systems Research Institute Working Paper 8803, University of Wisconsin, 1988.
Pencavel, John, "Labor Supply of Men: A Survey," in *Handbook of Labor Economics,* Orley Ashenfelter and Richard Layard, eds. (Amsterdam: North-Holland Publishing Company, 1986.).
Singh, Ajit, and Geoffrey Whittington, "The Size and Growth of Firms," *Review of Economic Studies,* XLII (1975), 15–26.
Tobin, James, "Estimation of Relationships for Limited Dependent Variables," *Econometrica,* XXVI (1958), 24–36.
Wilks, Samuel S., *Mathematical Statistics* (New York: John Wiley and Sons 1962).