The Impact of Vintage and Survival on Productivity:
Evidence from Cohorts of U.S. Manufacturing Plants

J. Bradford Jensen, Robert H. McGuckin, and Kevin J. Stiroh[*]

February 25, 2000

Abstract

This paper examines the evolution of productivity in U.S. manufacturing plants from 1963 to 1992. We define a "vintage effect" as the change in productivity of recent cohorts of new plants relative to earlier cohorts of new plants, and a "survival effect" as the change in productivity of a particular cohort of surviving plants as it ages. The data show that both factors contribute to industry productivity growth, but play offsetting roles in determining a cohort's relative position in the productivity distribution. Recent cohorts enter with significantly higher productivity than earlier entrants did, while surviving cohorts show significant increases in productivity as they age. These two effects roughly offset each other, however, so there is a rough convergence in productivity across cohorts in 1992 and 1987. (JEL Code: D24, L6)

I. Introduction

Economists have long argued that "best practice" technology is embodied in new capital so productivity of a manufacturing plant should be associated with its entry year or vintage. A plant's age, on the other hand, impacts productivity for different reasons. As plants age, managers accumulate experience, gain from learning-by-doing, undertake new investments, or achieve economies of scale, all of which can improve plant-level productivity.

This suggests two forces – a "vintage effect" from higher productivity of recent entrants and a "survival effect" from increasing productivity of surviving plants – interact to drive productivity growth for an industry and jointly push out the productivity frontier. Since many cohorts from different periods co-exist at any point in time, however, the competitive process should lead to a rough "convergence" in productivity across cohorts as the countervailing influences of vintage and survival balance.[1]

This paper explores and measures vintage and survival effects using cohort data from 19 U.S. manufacturing industries from 1963 to 1992. We compare the relative productivity of cohorts of different vintages (entry year) and different ages (number of years since entry) in order to answer several well-defined questions about productivity growth. First, do more recent cohorts enter with higher productivity than earlier cohorts? Second, do surviving plants of a particular vintage become more productive as they age? Third, how do these two effects trade off at a particular time? That is, do the improved capital, technologies, and practices behind vintage gains create an advantage for younger plants or do older, surviving plants show higher productivity from the experience, learning, and scale effects associated with age?

The paper is largely empirical and we structure our analysis around a carefully specified framework that systematically attempts to control for exogenous factors. Disembodied technical change, improved management techniques, and demand or supply shocks, for example, could affect the productivity of all plants in ways unrelated to the specific impact of vintage or age. Since failure to control for these factors can bias estimates of the vintage and survival effects, we use industry-wide variables like total industry output, average industry productivity, and change in total output to

---

[1]A robust finding from many studies is that plant productivity distributions are typically wide. This heterogeneity has been widely noted and the subject of considerable research. While it is somewhat of a puzzle that producers with very different productivity levels are able to survive under competitive conditions, there are many ways to reconcile the

control for the general, time-related factors in order to identify the vintage and survival effects. Our results suggest that both effects are large and important.

The 1992 cohort of new entrants was 51% more productive than the 1967 cohort in its entry year of 1967. Even after controlling for industry-wide factors and input differences, the vintage effect is over 50% from 1967 to 1992. Since this analysis compares successive cohorts of new entrants in their first period, we conclude that improved capital, technology, and operating practices embodied in new plants and equipment are an important source of productivity growth.

Analysis of a single cohort of surviving plants over time shows that age is also a good predictor of productivity. The 1967 cohort of surviving plants, for example, increased productivity by 57% between 1967 and 1992. While much of these gains were associated with industry-wide factors, productivity growth of 19% can be independently attributed to age-specific factors like experience and scale. These surviving plants showed steady gains and improved their relative standing in the productivity distribution.

Finally, the data suggest vintage and survival effects roughly offset each other in a given year. In 1992, all surviving cohorts that entered prior to 1992 showed average productivity within 6.5% of the industry average. Only the 1992 cohort was far from the industry average, 14.2% below, reflecting the important processes of competition and selection as competitive forces have not yet forced the lowest productivity plants from the market. Nonetheless, this sorting-out process appears quite rapid. The 1987 cohort, for example, entered with productivity 10% below the industry in 1987, but selection and survival gains allowed the surviving plants to improve to only 2.7% below the industry average by 1992. As low productivity plants failed and surviving plants improved with age, the cohort made large relative productivity gains.

These results highlight the complexity of the productivity dynamics that contribute to industry-level productivity growth. As new plants continually enter with higher productivity, competition weeds out the poor performers, and survivors improve with age, the productivity frontier is pushed out and average productivity rises. Only by looking at relative productivity across different cohorts at different points in time can we understand these important sources of productivity growth.

---

empirical regularities with competition. We do not dwell on this issue here and simply assume that the distribution within cohorts reflects idiosyncratic shocks or differences in the products produced within the industry.

II. Productivity and Micro-Data

Understanding productivity growth is obviously an important topic, but numerous measurement and conceptual problems make it difficult and controversial. One robust finding from micro-studies, however, is that distributions of plant-level productivity show enormous heterogeneity both within and across industries. Jensen and McGuckin (1997), for example, report productivity ranges of 4 or 5 to 1 as the norm within 4-digit manufacturing industries.

A second robust finding is the growth and survival of efficient producers, while inefficient producers decline and exit. Haltiwanger (1997), building on Baily, Hulten, and Campbell (1992) and Bartlesman and Dhrmyes (1998), attributes about 20% of U.S. manufacturing productivity growth to the entry of new plants and the exit of unsuccessful plants. New entrants exhibit average productivity levels well below those of incumbents, and a large fraction of these new entrants eventually fail. Haltiwanger (1997) estimates an additional 40% of manufacturing productivity growth is due to reallocations of output among surviving plants, with the remaining 40% reflecting industry-wide factors shared by all plants. Thus, approximately 60% of manufacturing productivity growth is associated with the growth of successful producers at the expense of the less successful rivals.

These facts imply evolution of the productivity distribution – which reflects a process of "creative destruction" – is a key factor in overall productivity growth. Plant vintage and plant age are two factors that contribute to this evolution. Recent empirical work, however, has paid little attention to distinguishing the two. With the notable exception of Bahk and Gort (1993), who focus directly on learning-by-doing and the separate impact of age and vintage, previous empirical work has typically treated these two characteristics as synonymous. Baily, Hulten, and Campbell (1992), for example, argue that the "the age of the plant is an obvious way to measure vintage (pg. 197)." For some purposes this is reasonable, but a deeper vintage and age story lies behind the evolution of plant-level productivity. This paper explores the interaction of these two key factors by isolating and measuring the impact of productivity changes for new entrants (vintage effect) and surviving incumbents (survival effect) in U.S. manufacturing plants from 1963 to 1992.

III. Definitions and Hypotheses

We define a plant's vintage, v, as the year a plant first produces, a plant's age, a, as the number of periods the plant has been in operation, and time, t, as the period in which a plant is observed. For each plant, therefore, $t \equiv v + a$. We begin with a traditional production function:

$$(1) \qquad Y_{t,i,a} = f_v(K_{t,i}, H_{t,i}, M_{t,i}, a_i, t)$$

where $f_v(.)$ is the production function for plants of vintage v, $Y_{t,i,a}$ is real gross output, $K_{t,i}$ is capital, $H_{t,i}$ is labor hours, $M_{t,i}$ is intermediate inputs, and a_i is age, all for plant i in period t.

Average labor productivity for plant i in period t of age a, $y_{t,i,a}$, is:

$$(2) \qquad y_{t,i,a} = \frac{Y_{t,i,a} - M_{t,i}}{H_{t,i}}$$

where $y_{t,i,a}$ is value-added per hour worked.[2]

To untangle the relationship between entrants, survivors, and productivity, we define a "vintage effect" as:

$$(3) \qquad V = \frac{d\,\hat{y}_{t,s,a}}{dv}\Big|_{a=0} = \frac{d\,\hat{y}_{t,s,0}}{dv}$$

and a "survival effect" as:

$$(4) \qquad S = \frac{d\,\hat{y}_{t,s,a}}{da}\Big|_v$$

where $\hat{y}_{t,s,a}$ is employment-weighted labor productivity of all plants in industry s at time t of age a.

The definitions appear similar, but they analyze different cuts of a larger data set. The vintage effect in Equation (3) compares productivity across cohorts of <u>new</u> plants ($a=0$) to measure how productivity evolves before age-related factors have an impact. The survival effect in Equation (4) compares productivity of the <u>same</u> set of surviving plants (a given vintage) as the plants age to compare how productivity varies with age.[3]

a) Plant Vintage

We treat the plant's production function as vintage-specific, e.g., $f_v(.)$, where different vintages utilize different production techniques. New plants fully embody the latest technology only if it is available at the time of entry. Models with this structure have a rich history, dating back at least to Solow (1960).[4]

---

[2] All empirical work was also done using gross output per employee, value-added per employee, and gross output per hour worked. The results were similar and we focus on value-added per hour worked to avoid the impact of variation in intermediate inputs and to capture variation in the workday.

[3] There are serious identification problems since each cohort is necessarily observed in different time periods. To identify the separate influence of vintage or age, therefore, one needs to control for general time effects. See below.

[4] Cooper, Haltiwanger, and Power (1995), Cooley, Greenwood, and Yorukoglu (1994), and Mortensen and Pissarides (1994) are more recent examples.

An important insight from this literature is that new plants enter with improved technology embodied in their plant and equipment and can outperform existing plants. Caballero and Hammour (1994) argue that new plants enter with higher productivity than older plants, earn rents, and then eventually exit when their vintage becomes outdated. This suggests $V>0$ in Equation (3) as <u>new</u> vintages are more productive than earlier ones. To measure these vintage differences, we compare the productivity of new plants ($a=0$) at different points in time.

One issue here is that existing plants may be able to "retool" by investing heavily and quickly in new types of capital, effectively making themselves indistinguishable from new plants. Given our definitions, however, retooling is part of the survival effect since vintage is defined as fixed parameter independent of capital age or investment. We do this for two reasons.

First, as a practical matter, available data do not provide any straightforward metric capable of distinguishing new and replacement investment. This is one reason why it is difficult to interpret Power's (1998) finding of little relationship between investment and productivity growth. Bahk and Gort (1993), on the other hand, develop a specific measure of vintage that takes account of post-entry investments, though we expect that their vintage measure is dominated by the plant's initial investments. Second, if investment is characterized by non-convexities in adjustment costs, as Doms and Dunne (1994) and Cooper, Haltiwanger, and Power (1995) suggest, then firms may not be able to quickly take advantage of technology embodied in new capital. Therefore, excluding changes in existing producers from the vintage effect and including them as part of the survival effect makes sense, particularly if one is interested in the pace and magnitude of technological change. While older survivors may improve by more or less than entrants, entrants provide a benchmark for how fast technology is changing.

b) Plant Age and Survival

Plants survive and become more productive with age for many reasons. Bartlesman and Doms (1997) and Bahk and Gort (1993) argue age-related productivity gains are due to changes in firm operations, e.g., productivity rises due to manager experience and learning-by-doing. Similarly, if older firms increase size, they may exploit economies of scale that improve productivity. Both of these plant specific, age-related factors can be captured by simply assuming $f_v(.)$ is a function of the plant's age. This assumption, while *ad hoc*, enables us to make the distinction between plant age and vintage. That is, a plant can get older, but it is still constrained to the same production function

determined by its vintage. If rapid investment allows an old plant to improve productivity, the age effect will capture the impact.

Alternatively, Baily et al. (1992), Jovanovic (1982), and Ericson and Pakes (1995) argue that selection and exit, which reflect plant-specific decisions and productivity shocks, are important determinants of the observed relationship between age and productivity. Baily et al. (1992) present a model where productivity growth is a random process and exit truncates the observed distribution from below. If plants receive a favorable draw then productivity grows as the plant ages, but if plants receive a negative draw, they exit and are lost from the sample. Jovanovic (1982) incorporates a fixed effect, which firms are able to identify only slowly so inefficient producers remain in business only until they realize they cannot be competitive. Ericson and Pakes (1995) provide a stochastic model of entry, exit, and competition where firms explore, via entry and investment decisions, potential profit opportunities. Relative to earlier models where firms passively receive a random shock or slowly learn about an immutable productivity parameter, firms take a more active role with success determined by endogenous investment decisions, the corresponding decisions of rivals, and random shocks, while exit is the result of an explicit, profit-maximizing choice.[5]

These different firm-level models suggest management, investment, and selection contribute to aggregate productivity growth, making it difficult to isolate particular age-related effects in empirical work. We use a balanced sample of survivors that combines the impact of selection and age-related factors like learning-by-doing, scale, investment, and experience. This implies $S>0$ in Equation (4) if the same set of surviving plants becomes more productive as they age.

IV. Data and Descriptive Statistics

Data are from the Longitudinal Research Database (LRD) at the Center for Economic Studies at the U.S. Census Bureau. This database contains outputs, inputs, and plant characteristics for individual manufacturing plants operating in the census years of 1963, 1967, 1972, 1977, 1982, 1987, and 1992.[6] There are 19 roughly two-digit manufacturing industries and we collected three sets of data for each year.

---

[5]We thank the referee for pointing out this comparison.

[6]See McGuckin and Pascoe (1988) for details on the LRD.

6

First, all plants in each industry for each period were aggregated to form the "Industry" measure, which serves as a benchmark. Second, all entrants in each industry for each period were aggregated to form a set of "New Plants," where each set of new plants is a distinct vintage. v_{63} enters prior to 1963; v_{67} enters between 1963 and 1967; v_{72} enters between 1968 and 1972; v_{77} enters between 1973 and 1977; v_{82} enters between 1978 and 1982; v_{87} enters between 1983 and 1987; and v_{92} enters between 1988 and 1992. Comparison of cohorts of New Plants allows us to measure the vintage effect. Note that the New Plants include all entrants regardless of whether the subsequently failed or survived. Third, "Surviving Cohorts" contain all plants for a given cohort that survived through 1992. Thus, each Surviving Cohort is a balanced panel, excluding plants that exited prior to 1992, which allows us to measure the survival effect.

The top panel in Table 1 shows the structure of the data. The first two rows include all plants (Industry) and all entrants (New Plants) in each year. For example, in 1967, there were 184,262 plants in operation, of which 49,984 were entrants. The next seven rows show Surviving Cohorts at different ages and time periods. Each diagonal contains the same number of plants, which reflects the construction of our balanced set of plants as it ages and moves through time. For example, 31,284 plants operated in 1963 and survived through 1992.

The number of plants in each Surviving Cohort increases over time relative to all New Plants, e.g., only 15% of the plants that entered between 1963 and 1967 survived to 1992, compared to 54% of the 1987 vintage. This shows the importance of selection and exit as most of entrants from early periods left the industry due to competitive pressures. In 1992, the numbers are identical since there has been no time to "weed out" entrants that will eventually exit.

Table 1 also presents mean output, mean employment, and average labor productivity, aggregated over 19 industries. Aggregation uses employment weights, which is appropriate for a representative firm analysis with constant returns to scale.[7]

The second panel shows New Plants are typically much smaller than the Industry, with their relative size remaining roughly constant over a 25 year period, e.g., 34% in 1967, 26% in 1977, and 35% in 1992. Plants in the Surviving Cohorts were larger in their entry year than those that did not survive (*a=0* row of Surviving Cohorts vs. New Plants), but not as large as the Industry (*a=0* row

---

[7]All output measures are in 1987 dollars, deflated with the 4-digit deflators in the NBER-CES Bartlesman-Becker-Gray database.

vs. Industry).[8] Finally, surviving plants steadily become larger with age. In 1992, for example, the average plant that entered between 1963 and 1967 produced nearly 5 times as much output as in the youngest cohort and twice the industry average.[9]

The third panel shows similar trends with respect to hours worked. New Plants are smaller than the Industry, Surviving Cohorts are larger than the plants that subsequently failed, and Surviving Cohorts grow in size as they age. It appears, however, that hours worked levels off after about 15-20 years and declines somewhat for the oldest plants.

The final panel in Table 1 shows average productivity for the Industry, New Plants, and each Surviving Cohort and suggests both the vintage effect and the survival effect are positive and quantitatively important. The vintage effect appears positive since the more recent cohorts of New Plants show higher productivity than earlier ones, e.g., the 1967 new plants showed productivity of $23.7 compared to $45.0 in 1992. Note the relative gap between the surviving cohorts and new plants widens as one moves back in time, which means that initial productivity is a good predictor of survival as reported by Baily et al. (1992). The survival effect appears positive since Surviving Cohorts generally becomes more productive as they age, e.g., the 1967 Surviving Cohort increased productivity from $29.8 in 1967 to $53.3 in 1992.

While both plant vintage and age appear to be good predictors of relative productivity, they will work in opposite directions at a given time. New plants gain from their more modern vintage of technology and capital, while older plants enjoy the benefits of experience, age, and scale that influence survival. While competition will make it hard for low productivity plants to survive, there is no fundamental reason to expect either the vintage or survival effect to dominate.

To explore this issue, we define a "net effect" as the productivity of each of the seven Surviving Cohorts relative to the industry in 1992. If newer plants are more productive, the vintage effect dominates; if older plants are more productive the age effect dominates. By looking down the last column of the final panel of Table 1, one can see an advantage for the oldest plants, approximate equality for the 1967 through 1987 vintages, and much lower relative productivity for the 1992 cohort. We interpret this as evidence of a productivity "convergence" where vintage and survival

---

[8]The pre-1963 group is not strictly comparable since all plants that entered prior to 1963 are labeled as "new" in 1963, the first observation period.

[9]As an artifact of the data collection, average output and employment per plant increase dramatically for the industry between 1963 and 1967. This is because the 1963 Census did not exclude very small plants. In subsequent years these plants were treated as administrative records with unreliable shipments data.

effects roughly offset. Only the 1992 surviving cohort shows a large productivity shortfall, which reflects the on-going selection process. Since this cohort only recently entered, it still contains many low productivity that are likely to fail.

These descriptive results show recent entrants are more productive than earlier entrants; survivors improve productivity with age; and vintage and survival gains roughly offset at a point in time. The next section augments these findings with a more rigorous econometric approach that controls for other confounding factors that might obscure the independent impact of vintage, age, and survival.

V. Econometric Issues

This section examines two difficulties associated with measuring the separate influences of vintage and survival. First, both effects must be separated from the general impact of time-related factors that affect the productivity of all plants. Second, one must control for variation in other inputs that affect productivity independent of vintage and survival.

a) Controlling for Time Effects

Consider again the age/time matrix in Table 1. A comparison of New Plants, say the 49,984 entrants in the 1967 cohort versus the 82,492 in the 1992, measures both the vintage effect (entry years are different) and a "time effect" since the observations are in different years. Likewise, a diagonal comparison of a Surviving Cohorts, say the 7,614 survivors of the 1967 cohort in 1967 and in 1992, measures both the survival effect (the cohort is observed at different ages) and a time effect (observations are in different years). Finally, a column comparison across Surviving Cohorts, say the 82,492 plants in the 1992 cohort versus the 7,614 survivors from the 1967 cohort, measures both the vintage effect (entry years are different) and the survival effect (cohorts are different ages). The comparisons from Table 1 show the combined impact of any two of these factors.

Time effects represent any factors that affect the productivity of <u>all</u> plants in a given period that are not uniquely attributable to vintage or survival. For example, if disembodied technical change raises the productivity of all plants, then all productivity gains along a diagonal cannot be interpreted as a survival effect. Similarly, changes in regulation, international competitiveness, or consumer tastes can impact the productivity of all plants and bias the measurement of vintage and survival effects.

Our goal is to remove these time effects so a comparison across the row of New Plants or along a diagonal for a Surviving Cohort in Table 1 will identify the vintage and survival effects, respectively. In the context of measuring price change, however, Hall (1971) shows that the three independent effects cannot be identified econometrically with an analysis of variance approach. Due to the exact collinearity of t, a, and v, it is impossible to include a complete set of time, age, and vintage dummy variables in a single productivity regression.

To avoid this problem, various techniques have been proposed.[10] We assume time effects are exogenous to the individual plant and control for them with industry-wide variables – average labor productivity, total output, and change in total output – that are not perfectly collinear with age and vintage, but still capture the time-related factors that affect plant-level productivity. If these industry variables control for industry-wide shocks, then OLS regressions with cuts of the data set described above will identify the vintage and survival effects.

This interpretation differs somewhat from Bahk and Gort (1993), who utilize a time trend to measure industry-wide learning-by-doing as a productivity residual. They find "industry-wide increases in the stock of knowledge affect output only insofar as they are uniquely related to embodied technical change of physical capital." That is, disembodied technological change is indistinguishable from embodied technical change. We interpret time effects more broadly to include the impact of all time-varying factors affecting productivity of all plants, e.g., technology, knowledge, regulation, business cycles, and demand shocks. As such, these time effects impact the entire industry, but are not perfectly correlated with measured inputs or particular cohorts.

b) **Controlling for Input Differences**

A second issue is different cohorts may face different input prices and choose different input mixes. Since data constraints force us to use labor productivity as the dependent variable, we must be careful that observed productivity differences across cohorts reflect the vintage and age effects, and not differences in other measurable inputs. For example, a higher quality workforce or a price-

---

[10]One possibility is to include non-linear functions of the age, time, and vintage variables in a simple regression. Another is to assume that some effect is fixed at an arbitrary level (possibly zero) and then examine a specific profile, e.g., the hedonic price literature typically examines the time-age profile and suppresses the impact of vintage. In a more recent study, Berndt, Griliches, and Rappaport (1995) estimate a "fully-saturated model," which includes the maximum number of time, age and vintage dummy variables without causing exact collinearity, and then infer the independent effects of time, age and vintage.

induced substitution towards more capital-intensive production raises labor productivity, but this does not represent the vintage and age effects in which we are interested.

To control for input variation, we use value-added rather than gross output in the numerator of our productivity measure. Subtracting the value of intermediate inputs controls for variation in the production methods across industries and across time. As a practical matter, we find little empirical difference between estimates based on value-added and those using gross output.

In addition, we employ LRD data on average wages and capital intensity to control for different input characteristics. In competitive markets, workers receive wages equal to their marginal revenue product, so average salary and wages per hour worked, $W_{t,s,a}$, is a natural proxy for the quality of a cohort's workforce. This is calculated on an employment-weighted basis for each cohort. Wages and salary income is gross earnings of all employees, production and non-production, deflated with a regional CPI index. We also include machinery per hour worked, $k_{t,s,a}$, as a measure of capital intensity, again calculated on an employment-weighted basis for each cohort. Machinery is the end of the period, book-value stock, deflated with the national CPI.

VI. Empirical Results

This section reports econometric estimates of the vintage, survival, and net effects discussed above. All estimates are based on cross-sectional regressions of 19 manufacturing industries with cohorts of New Plants or seven Surviving Cohorts from 1963 to 1992, as described in Table 1.

a) Measuring the Vintage Effect: New Plant Regressions

We estimate the vintage effect by comparing the productivity of New Plants across time. The identifying assumption is that industry-wide variables capture the influence of any factors affecting the productivity of all plants and that wage and capital variables control for input variation, so all remaining variation is due to the vintage effect. That is, conditional on the industry-wide and input controls, any systematic difference across these cohorts is attributable to plant vintage.[11]

The vintage effect compares the productivity of different cohorts of entrants at time t in industry s, $\hat{y}_{t,s,a=0}$, to the original vintage of new plants, $\hat{y}_{1963,s,a=0}$, controlling for time-related factors, labor quality, and capital intensity as:

(5) $\quad ln(\hat{y}_{t,s,a=0}) = a_0 + a_1 W_{t,s,a=0} + a_2\, k_{t,s,a=0} + b'_v\, I_v + \; d_1\, ln(\hat{y}_{1963,s,a=0}) + d_2\, ln(Y_{t,s}) + d_3\, DY_{t,s} + e$

---

[11]Note there are no age effects in this approach since all plants in this specification are new entrants.

where I_v includes 6 dummy variables for each vintage except the 1963 vintage, $Y_{t,s}$ is industry total output, and $DY_{t,s}$ is the change in total output over the preceding five year interval, for industry s at time t.

If the identifying assumption holds, the vintage effects are identified by b_v in Equation (5). Since the dummy variable for the 1963 vintage is dropped and the dependent variable is in logs, each b_v measures the proportional change in productivity of subsequent cohorts relative to the original 1963 vintage cohort, conditional on the other right-hand side variables.

Note that this represents a severe test for the vintage effect due to the construction of the dataset. The original 1963 cohort consists of all firms that entered in any year prior to 1963. Thus, the 1963 benchmark contains plants that were not entrants in the five-year period prior to 1963. Any evidence of a vintage effect, therefore, will indicate that each cohort is entering with substantial advances in productivity.

We first drop both the industry and input variables and estimate a simpler version of Equation (5). Column (1) in Table 2 shows large, positive, and significant estimates for the b_v coefficients. For example, productivity is 57% higher for the 1992 vintage relative to the 1963 vintage in their respective entry years.

As emphasized above, however, this does not control for time effects or input differences and thus does not isolate vintage effects. Column (2) includes industry controls and still shows a positive and significant vintage effect, although the coefficients are typically smaller.[12] The change in the b_v coefficients relative to Column 1 represents the productivity growth that is explained by industry-wide factors. Column (3) reports estimates after the inclusion of both industry controls and the wage and capital variables. Again, the magnitude of the vintage effect declines, but remains significant for all cohorts past the 1967 cohort.[13] Both the capital intensity and wage variables are positive and statistically significant.

These results show that more recent vintages of manufacturing plants entered with higher productivity than earlier entrants, even after controlling for industry-wide factors and input differences. This suggests that the vintage effect is an important source of industry productivity gains. It is important to point out, however, that even though recent cohorts of new plants are more

---

[12]An F-test indicates the two industry control variables are jointly significant at the 99% level.
[13]The industry controls remain jointly significant at the 98% level.

productive than earlier vintages, they are not as productive as the average plant in the industry when they enter. As Table 1 shows, this spread has remained relatively constant over time.

Finally, it is interesting to note that even entrants with the latest technology were not immune to the productivity slowdown in the mid-1970s. Middle cohorts show some productivity growth, but later cohorts show the largest relative gains. The negative coefficient for the 1967 cohort, which suggests negative productivity growth, undoubtedly reflects the somewhat unfair comparison to the 1963 cohort. This cohort is composed of all plants that entered prior to 1963 and survived to 1992 so it includes some very successful, very old survivors.[14]

b) **Measuring the Survival Effect: Surviving Cohort Regressions**

Empirical analysis of the survival effect examines productivity changes along a diagonal in Table 1. Our estimation framework is similar to the previous section, but requires a different cut of the data. We report results for the Surviving Cohort that entered in 1967 and analyze the evolution of productivity along a single diagonal in Table 1. We also examined later cohorts and the results are consistent with those reported here.[15]

To identify the survival effect, we include the input and industry-wide control variables in the following regressions:

(6) $ln(\hat{y}_{t,s,t-1967/92}) = a_0 + a_1 W_{t,s,t-1967} + a_2 k_{t,s,t-1967} + l'_a I_a + g'I_s + d_1 ln(\hat{y}_{t,s}) + d_2 ln(Y_{t,s}) + d_3 DY_{t,s} + e$

where I_a includes 5 dummy variables for each age except $a=0$ and the remaining variables are defined above and I_s is a vector of industry dummy variables.[16]

The identifying assumption is again that the industry and input control variables capture time effects and input differences, so any remaining variation reflects age and selection. If d_1, d_2, and d_3 capture the impact of all time-related factors, e.g., the general upward trend in plant-level productivity, then λ_a measures the independent survival effect associated with age. Since the dummy for $a=0$ is dropped, each λ_a measures the proportional change in productivity at each age relative to the productivity of the surviving 1967 cohort in 1967 when it entered.

---

[14]For the 1963 cohort, we simulated which plants were likely to have entered in 1963 and recalculated productivity for this group. The productivity for this more representative group of plants was significantly lower than the full 1963 cohort. Thus, if we could identify true 1963 entrants, the coefficient on the 1967 cohort would likely be positive.

[15]We use the 1967 cohort since the 1963 cohort consists of all plants that entered prior to 1963 and is thus made up of plants of very different ages. In addition, the capital data is not available for this period.

[16]Note there are no vintage effects in this approach since all plants entered in the same period.

This specification extends Table 4 in Bahk and Gort (1993), which reports production function estimates for a consistent set of plants in different years. They interpret shifts in the input elasticities over time as the impact of learning. In contrast, Equation (6) assumes a single production function with an age-varying intercept to capture the impact of age-related factors. This seems to be an improvement since the Bahk and Gort approach does not test a nested model for significant differences across year-to-year estimates. Our approach enables direct tests of the productivity impact of age, i.e., if average productivity increases conditional on capital intensity, labor quality, and time effects, then unobservable age-related factors are important.

We again first drop the industry and input controls and first estimate a simpler version of Equation (6). Results are reported in Column (1) of Table 3. The estimates of λ_a show a significant increase in productivity with age and imply that the 1967 cohort is 57% more productive in 1992 than it was in 1967.

This simple regression does not control for industry-wide factors that impact productivity and thus biases the estimate of the survival effect. Column (2) includes the industry-wide control variables and the magnitude of the survival effect drops substantially, although it remains positive and statistically significant in later years, e.g., 15% in 1992.[17] The decline in λ_a after the industry controls are included implies that a large part of the increased productivity of the 1967 cohort is not age-specific, but rather represents general productivity gains that affect all plants.

The inclusion of the wage and capital intensity variables, Column (3) in Table 5, does not change the results dramatically. Neither input variable is statistically significant, however, and the capital variable is surprisingly negative. After accounting for differences in labor quality, input intensity, and general time effects, the 1967 cohort shows a 19% productivity gain in 1992 relative to 1967.[18] These results show that age is an important factor that influences productivity growth.

This specification leads to a survival effect that must be interpreted broadly. That is, λ_a captures all factors that vary over the life of surviving plants such as scale effects, "retooling" through rapid investment, learning-by-doing, scale economies, or managerial efficiency gains. One area for future research involves decomposing the factors that contribute to this age effect. In principle, one could control for investment to isolate the retooling effect, but our data are from the

---

[17]An F-test indicates the three control variables are jointly significant at the 99% level.

[18]The industry-wide variables remain jointly significant at the 99% level.

full Census of Manufactures which is conducted at five-year intervals and we cannot obtain investment for intervening years.[19] In addition, there is the practical difficulty of distinguishing replacement investment from new investment as discussed earlier.

We can, however, include cohort average size to look for scale effects. If average size (mean value-added) is included in the regressions reported in Table 3, the age coefficients decline and are no longer statistically significant. While the causality is difficult to interpret – are more productive plants growing or does size increase productivity – this suggests size is closely linked with survival and productivity gains. Whether the size effect reflects learning or scale economies remains an unresolved issue. Future work with annual time series of plant-level data may be able to shed light on these interesting questions.

c) Measuring the Net Vintage and Survival Effects: Column Regressions

The previous section shows that plant vintage and plant survival/age are significant, but opposing determinants of plant-level productivity. That is, new cohorts gain from a later vintage, but suffer from their relative youth and inexperience. Older cohorts, on the hand, gain from experience, but suffer from a relatively old vintage. To examine this trade-off, we suppress the time dimension and compare productivity of different cohorts to measure the "net effect." In the context of Table 1, this compares productivity down Surviving Cohorts in a given year. Only 1992 and 1987 are analyzed since earlier years have too few active vintages.

The net effect regression is:

$$(7) \qquad ln(\hat{y}_{t,i,a}/_{92}) = a_0 + a_1 W_{t,i,a} + a_2 k_{t,i,a} + p'_a I_a + g'I_s + e$$

where $t=1992$ or $t=1987$ and p_a includes dummy variables for each age group, i.e., p_0 for the youngest cohort and p_6 for oldest cohort. The dummy for the 1963 cohort is dropped, so each p measures the proportional difference in productivity relative to that cohort of surviving plants in each year.[20]

---

[19]One possibility would be to use data from the Annual Survey of Manufactures (ASM). The ASM provides a measure of investment, but it is a survey with a rotating panel design that makes it difficult to consistently identify entrants and survivors over long periods of time.

[20]Only industry dummy variables, I_s, are included as controls since this analysis is in a single year. Any additional industry-specific variables would be constant for an industry and thus they would not provide additional information.

This specification is similar to Baily et al. (1992), but our interpretation is quite different. We view systematic differences across cohorts in a given year as the net effect of vintage and age, so p_a captures the net impact on productivity from these opposing influences.

Table 4 shows the estimates of Equation (7) for 1992 and 1987. We first estimate a simple regression without input controls in Columns (1) and (3). Results show negative and significant p coefficients for both years for recent cohorts and generally insignificant differences for earlier one. The 1992 cohort in 1992, for example, is 21% less productive than the 1963 cohort in 1992, while the 1987 cohort is 13% less productive than the 1963 cohort in 1987. This indicates that selection and age are dominant influences in a cohort's relative productivity.

When labor quality and capital intensity controls are included in Columns (2) and (4), however, there is a general increase in the coefficients. Much of the variation across cohorts appears to be associated with variation in labor quality and capital intensity. Once one controls for these differences, there is little systematic difference across cohorts as the effects of vintage and survival seem to offset.

This suggests the net effect of vintage and survival cancel. We interpret this as a "convergence" in productivity across the age dimension, after one accounts for selection phenomenon and input differences. Surviving plants gain with age and new plants gain from vintage, but the competitive process prevents either from dominating. Surviving cohorts, regardless of age or vintage, show similar productivity levels. Moreover, this weeding out process happens quite quickly. In 1987, new plants were 10% below the industry, but by 1992 the surviving plants from that cohort were less than 3% below the industry. In only 5 years, the combination of exiting plants with low productivity and surviving plants improving productivity brought the 1987 cohort to virtual parity with the industry

These results support the theoretical models of Jovanovic (1982) and Ericson and Pakes (1995), and support the empirical findings of Baily et al. (1992). Our data show entering cohorts contain a wide distribution of plants, some of which survive and flourish, while others fail and exit. This is consistent with both the unobserved productivity parameter model of Jovanovic and the active investment/exploration of Ericson and Pakes, although we don't rigorously distinguish between them. Casual observation, however, suggests that some important characteristics are

present at entry, e.g., surviving entrants show higher productivity than their peers that subsequently fail.[21]

VII. Summary and Conclusions

These results suggest five main conclusions about plant-level productivity growth. First, vintage matters: new plants embody better production technology and, even after controlling for labor quality and capital intensity, show higher productivity than earlier cohorts of entrants. There is some evidence that the process is not a smooth one as there appear to be "technological epochs" in the data that affected both entrants and incumbents. Productivity rapidly increases prior to 1972, remains flat from 1972-1982, and then increases until the early 1990's, but the aggregate nature of analysis precludes more than a tentative conclusion.

Second, our results are consistent with a number of other studies that find selection matters. Recent entrants show productivity levels below industry averages, but this is largely due to a large number of small, low productivity plants that subsequently fail. Rapid failure of these plants leaves behind larger, high productivity survivors.

Third, age has a positive effect on productivity. Surviving plants improve their relative standing in the productivity distribution as they age. While we can't identify the specific sources of these age gains in this analysis, it is likely that learning-by-doing, scale economies, and improvements in the quality of the workforce are principle factors. We cannot rule out, however, the possibility that these surviving plants undertake large investments that effectively allow them to retool and replicate the latest capital.

Fourth, labor quality and input intensity are important determinants of productivity differences across cohorts. These variables are typically significant and tend to weaken, but not eliminate, the separate impact of vintage and survival/age.

Fifth, convergence appears to be a feature of the productivity distribution, conditional on variation in inputs. The data suggest that, after about 5-10 years, all cohorts of surviving plants have similar productivity levels. Absent special circumstances such as monopoly power and/or regulation, plants remain viable only if productivity levels are comparable. This is sensible in a competitive world, but still must be reconciled with the pervasive heterogeneity observed in the productivity

---

[21]To formally distinguish these models requires tests of the ergodic properties of the equilibrium as discussed in Ericson and Pakes (1995), particularly pg. 69. These tests are beyond the scope of this paper.

distributions for individual plants in virtually every industry. Various explanations have been proposed for the observed distribution – measurement errors, differences in the products, selection, plant-level uncertainty, and alternative production processes that allow for differences in labor productivity – but none yet appears to be the whole story.

We close with the observation that these results hide substantial variation in evolutionary patterns within industries. This is to be expected since the pace of technological process varies significantly among industries. For example, technology, research and development, and the possibilities for exploiting information technology vary widely across industries. Exploring these differences is an important step that remains to be taken.

References

Bahk, Byong-Hyong and Michael Gort. (1993). "Decomposing Learning by Doing in New Plants." *Journal of Political Economy*. Vol. 101. No. 4. 561-583.

Baily, Martin N., Charles Hulten and David Campbell. (1992). "The Distribution of Productivity in Manufacturing Plants." *Brookings Papers: Microeconomics*. Washington, DC.

Bartelsman, Eric J. and Mark Doms. (1997). "Understanding Productivity: Lessons from Longitudinal Micro Datasets." manuscript.

Bartlesman, Eric J. and Phoebus J. Dhrymes. (1998). "Productivity Dynamics: US Manufacturing Plants, 1972-1986. *Journal of Productivity Analysis*. 9. 5-34.

Berndt, Ernst, Zvi Griliches and Neal Rappaport, (1995), "Econometric Estimates of Price Indexes for Personal Computers in the 1990s," *Journal of Econometrics*, Vol. 68. No. 1. July. 243-268.

Caballero, Ricardo and Mohamad Hammour. (1994). "The Cleansing Effects of Recessions." *American Economic Review*. 84. pp. 1356-1368.

Cooley, Thomas, Jeremy Greenwood, and Mehmet Yorukoglu. (1994). "The Replacement Problem." unpublished manuscript. University of Rochester.

Cooper, Russel, John Haltiwanger and Laura Power. (1995). "Machine Replacement and the Business Cycle: Lumps and Bumps." NBER Working Paper #5260.

Doms, Mark and Timothy Dunne. (1994). "Capital Adjustment Patterns in Manufacturing Plants." Center for Economic Studies. U.S. Bureau of the Census. Washington, DC.

Ericson, Richard and Ariel Pakes. (1995). "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*. Vol. 62. No. 1. 53-82.

Haltiwanger, John C. (1997). "Measuring and Analyzing Aggregate Fluctuations: The Importance of Building from Microeconomic Evidence." *Review*. Federal Reserve Bank of St. Louis. May/June. pp. 55-75.

Jensen, J. Bradford and Robert H. McGuckin. (1997). "Firm Performance and Evolution: Empirical Regularities in the U.S. Microdata." *Industrial and Corporate Change*. Vol. 6. No. 1. 25-47.

Jovanovic, Boyan. (1982). "Selection and the Evolution of Industry.*" Econometrica.* Vol. 50. No. 3. May. 649-679.

McGuckin, Robert H. and George A. Pascoe Jr. (1988). "The Longitudinal Research Database (LRD): Status and Research Possibilities." *Survey of Current Business*. 68(11). 30-37.

Mortensen, Date T., and Chistopher A. Pissarides. (1994). "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies*. Vol. 61. No. 3. 397-415.

Power, Laura (1998). "The Missing Link: Technology, Investment, and Productivity." *The Review of Economics and Statistics.* Vol. 80. No. 2. 300-313.

Solow, Robert M.   (1960). "Investment and Technical Progress." in eds. Kenneth Arrow, Samuel Karlin and Paul Suppes.   <u>Mathematical Methods in the Social Sciences</u>.   Stanford, CA.   Stanford University Press.

Table 1:
Descriptive Statistics for Cohorts of U.S. Manufacturing Plants, 1963-1992

		t = 1963	t = 1967	t = 1972	t = 1977	t = 1982	t = 1987	t = 1992
					Time (t)			

Number of Plants (000s)

		t = 1963	t = 1967	t = 1972	t = 1977	t = 1982	t = 1987	t = 1992
Industry		301,144	184,262	186,098	196,569	213,821	197,873	221,474
New Plants		301,149	49,984	67,090	71,873	77,196	63,495	82,492
				Surviving Cohorts				
	a = 0	31,284	7,614	12,140	16,662	23,225	34,296	82,492
	a = 1		31,284	7,614	12,140	16,662	23,225	34,296
	a = 2			31,284	7,614	12,140	16,662	23,225
Age (a)	a = 3				31,284	7,614	12,140	16,662
	a = 4					31,284	7,614	12,140
	a = 5						31,284	7,614
	a = 6							31,284

Value-Added per Plant (000s of 1987 Dollars)

		t = 1963	t = 1967	t = 1972	t = 1977	t = 1982	t = 1987	t = 1992
Industry		2,032.7	4,188.4	4,602.3	4,627.6	4,035.7	5,734.7	5,517.7
New Plants		2,034.6	1,409.6	1,480.8	1,209.9	1,099.9	2,128.7	1,955.3
				Surviving Cohorts				
	a = 0	9,498.1	3,437.2	3,359.8	2,345.4	1,798.7	2,896.0	1,955.3
	a = 1		12,434.	5,458.6	4,621.2	3,233.3	3,579.6	3,879.5
	a = 2			13,833.5	6,386.6	5,131.5	4,916.0	4,641.4
Age (a)	a = 3				14,449.5	6,478.2	7,719.2	5,536.1
	a = 4					13,046.	9,227.5	8,392.2
	a = 5						17,009.9	9,688.8
	a = 6							16,804.0

Mean Hours Worked per Plant (000s)

		t = 1963	t = 1967	t = 1972	t = 1977	t = 1982	t = 1987	t = 1992
Industry		80.7	148.0	138.9	131.6	105.5	117.4	102.3
New Plants		80.7	59.6	52.4	42.8	36.0	51.4	43.5
				Surviving Cohorts				
	a = 0	312.8	115.4	98.7	73.0	54.0	64.9	43.5
	a = 1		368.9	158.1	133.7	88.0	80.3	75.5
	a = 2			355.8	174.3	133.7	108.1	86.6
Age (a)	a = 3				352.8	164.6	159.4	112.3
	a = 4					294.8	183.9	159.3
	a = 5						306.0	181.9
	a = 6							276.2

Average Labor Productivity (Value-Added per Hour Worked)

		t = 1963	t = 1967	t = 1972	t = 1977	t = 1982	t = 1987	t = 1992
Industry		25.2	28.3	33.1	35.2	38.3	48.8	54.0
New Plants		25.2	23.7	28.3	28.3	30.5	41.4	45.0
				Surviving Cohorts				
	a = 0	30.4	29.8	34.0	32.1	33.3	44.6	45.0
	a = 1		33.7	34.5	34.6	36.7	44.6	51.4
	a = 2			38.9	36.6	38.4	45.5	53.6
Age (a)	a = 3				41.0	39.4	48.4	49.3
	a = 4					44.3	50.2	52.7
	a = 5						55.6	53.3
	a = 6							60.8

Table 2
The Impact of Vintage on Productivity – The Vintage Effect

Independent Variables	$\ln(\hat{y}_{t,s,a=0})$		
	(1)	(2)	(3)
Constant	0.433	0.251	0.991
	(0.106)	(0.329)	(0.321)
β_{67}	0.058	-0.066	-0.040
	(0.031)	(0.043)	(0.051)
β_{72}	0.237	0.168	0.119
	(0.040)	(0.043)	(0.043)
β_{77}	0.234	0.200	0.154
	(0.035)	(0.034)	(0.042)
β_{82}	0.253	0.297	0.250
	(0.051)	(0.042)	(0.044)
β_{87}	0.531	0.393	0.329
	(0.051)	(0.063)	(0.064)
β_{92}	0.571	0.538	0.468
	(0.061)	(0.056)	(0.052)
$\ln(\hat{y}_{1963,s,a=0})$	0.862	0.847	0.576
	(0.033)	(0.043)	(0.066)
$\ln(Y_{t,s})$		0.013	-0.008
		(0.024)	(0.025)
$\Delta Y_{t,s}$		0.517	0.345
		(0.133)	(0.116)
$k_{t,s,a=0}$			0.0020
			(0.0009)
$W_{t,s,a=0}$			0.0342
			(0.0049)
R^2	0.90	0.91	0.94
No. of Obs.	133	133	133
Prob > F		0.0004	0.0136

Heteroskedasticity-corrected standard errors in parentheses.
F-test for joint significance of $\ln(Y_{t,s})$ and $\Delta Y_{t,s}$.

Table 3:
The Impact of Age on Productivity – The Survival Effect

Independent Variables	$\ln(\hat{y}_{t,s,t\text{-}1967}|_{92})$		
	(1)	(2)	(3)
Constant	3.549	0.054	0.610
	(0.034)	(0.846)	(0.856)
λ_1	0.173	0.051	0.060
	(0.043)	(0.042)	(0.040)
λ_2	0.219	0.049	0.057
	(0.043)	(0.049)	(0.048)
λ_3	0.271	0.068	0.084
	(0.042)	(0.058)	(0.056)
λ_4	0.494	0.148	0.169
	(0.049)	(0.076)	(0.068)
λ_5	0.569	0.148	0.187
	(0.046)	(0.083)	(0.079)
$\ln(\hat{y}_{t,s})$		0.660	0.586
		(0.130)	(0.137)
$\ln(Y_{t,s})$		0.061	0.037
		(0.051)	(0.050)
$\Delta Y_{t,s}$		-0.089	-0.041
		(0.090)	(0.105)
$k_{t,s,t\text{-}1967}$			-0.0016
			(0.0014)
$W_{t,s,t\text{-}1967}$			0.014
			(0.012)
R^2	0.96	0.98	0.98
No. of Obs.	114	114	114
Prob > F		0.0000	0.0001

Industry dummy variables are not shown.
Heteroskedasticity-corrected standard errors are shown in parentheses.
F-test for joint significance of $\ln(Y_{t,s})$, $\Delta Y_{t,s}$, and $\ln(\hat{y}_{t,s})$.

The Net Impact of Vintage and Survival on Productivity in 1992 and 1987

Independent Variables	$\ln(\hat{y}_{t,s,a|92})$		$\ln(\hat{y}_{t,s,a|87})$	
	(1)	(2)	(3)	(4)
Constant	4.202	3.631	4.196	3.509
	(0.063)	(0.145)	(0.049)	(0.109)
π_0	-0.207	-0.056	-0.130	-0.001
	(0.036)	(0.040)	(0.041)	(0.034)
π_1	-0.091	0.016	-0.136	0.017
	(0.043)	(0.040)	(0.036)	(0.028)
π_2	-0.096	0.040	-0.090	0.055
	(0.040)	(0.042)	(0.035)	(0.027)
π_3	-0.120	0.004	-0.077	0.020
	(0.034)	(0.036)	(0.040)	(0.032)
π_4	-0.075	0.020	-0.054	0.022
	(0.045)	(0.035)	(0.039)	(0.031)
π_5	-0.054	0.012		
	(0.042)	(0.046)		
$k_{t,s,a}$		0.0011		0.0011
		(0.0006)		(0.0006)
$W_{t,s,a}$		0.0348		0.0396
		(0.0087)		(0.0065)
R^2	0.96	0.97	0.96	0.98
No. of Obs.	133	133	114	114

Industry dummy variables are not shown.
Heteroskedasticity-corrected standard errors in parentheses.