CAM

Centre for Applied Microeconometrics

Institute of Economics
University of Copenhagen

http://www.econ.ku.dk/CAM/

Productivity Growth and Worker Reallocation: Theory and Evidence

Rasmus Lentz
Dale T. Mortensen

2004-12

The activities of CAM are financed by a grant from The Danish National Research Foundation

Productivity Growth and Worker Reallocation: Theory and Evidence*

Rasmus Lentz
Boston University

Dale T. Mortensen
Northwestern University

July 19, 2004

Abstract

Dispersion in labor and factor productivity across firms is large and persistent, large flows of workers move across firms, and worker reallocation is an important source of productivity growth. The purpose of the paper is to provide a formal explanation for these observations that clarifies the role of worker reallocation as a source of productivity growth. Specifically, we study a modified version of the Schumpeterian model of growth induced by product innovation developed by Klette and Kortum (2002). More productive firms are those that supply higher quality products in the model. We show that more productive firms grow faster and the reallocation of workers across continuing firms contributes to aggregate productivity growth if and only if current productivity predicts future productivity. We provide evidence in support of the hypothesis that more productive firms become larger in Danish data. In addition, we provide estimates of the distribution of productivity at entry and the parameters of the cost of investment in innovation function and other structural parameters that all firms are assumed to face by fitting the model to observations on value added, employment, and wages drawn from a panel of Danish firms for the years 1992-1997.

*We would like to thank Victor Aguirregabiria for helpful comments. Centre for Applied Microeconometrics, Center for Corporate Performance, the Danish Social Science Research Council, and Bent Jesper Christensen generously provided data access.

1 Introduction

In their review article on firm productivity, Bartelsman and Doms (2000) draw three lessons from empirical studies based on longitudinal plant and firm data: First, the extent of dispersion in relative productivity across production units, firms or establishments, is large. Second, productivity rank of any unit in the distribution is highly persistent. Third, a large fraction of aggregate productivity growth is the consequence of worker reallocation. In their recent study of wage and productivity dispersion trends in U.S. Manufacturing, Dunne, Foster, Haltiwanger, and Troske (2002) find that wage differences in wages across plants is an important and growing component of total wage dispersion, most of the between plant increase in wage differences is within industries, and wage and productivity dispersion between plants has grown substantially in the recent past. Although the explanations for productive heterogeneity across firms are not fully understood, economic principles suggest that wage and productivity dispersion should induce worker reallocation from less to more productive firms as well as from exiting to entering firms. Indeed, workers should move voluntarily to capture wage gains while more productive employer have an incentive to expand production.

There is ample evidence that workers do flows from one firm to another frequently. As Davis, Haltiwanger, and Schuh (1996) and others document, job and worker flows are large, persistent, and essentially idiosyncratic in the U.S. Recently, Fallick and Fleischman (2001) and Stewart (2002) find that job to job flows without a spell of unemployment in the U.S. represent at least half of the separations and is growing. In their analysis of Danish matched employer-employee IDA data, Frederiksen and Westergaard-Nielsen (2002) report that the average establishment separation rate over the 1980-95 period was 26%. About two thirds of the outflow represents the movement of workers from one firm to another. Using firm level data based on the same source, Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005) document considerable cross firm dispersion in the average wage paid. Furthermore, they show that separation rates decline steeply with a firm's relative wage suggesting that workers do move from lower to higher paying jobs.

Baily, Hulton, and Campbell (1992) find a strong positive correlation between productivity and wages paid across plants in U.S. manufacturing and Bartelsman and Doms (2000) report that the finding is present in similar studies. Mortensen (2003) argues that dispersion in wages paid for observably equivalent workers is hard to explain unless they reflect differences in firm productivity. To the extent that wage dispersion reflects differences in firm specific labor productivity, direct voluntary flows of workers from lower to higher paying firms as well as indirect flows through unemployment from less to employment with more productive firms improve the overall allocation labor in the economy. As noted earlier, the studies cited by Bartelsman and Doms (2000) document that labor reallocation of this form is a major contributor to aggregate productivity growth.

The purpose of this paper is to clarify the role of worker reallocation in the growth process. The model developed by Klette and Kortum (2002), which itself builds on the endogenous growth model of Grossman and Helpman (1991), is adapted for this purpose. Their version of the model is designed to be consistent with stylized facts about product innovation and its relationship to the dynamics of firm size evolution and the distribution of firm size. In the model, firms are monopoly suppliers of differentiated products viewed as inputs in the production of a final consumption good. Better quality products are introduced from time to time as the outcome of R&D investment by both existing firms and new entrants.

As a theoretical result, we show that more productive firms, those that have developed higher quality products in the past, tend to grow larger by developing more product lines in the future only if a firm's future product quality is positively correlated with it past innovation success. If product quality were iid across innovations, then investment in R&D would be independent of a firm's current productivity. Interestingly, the qualitative relationship between employment size and labor productivity is ambiguous in the first case and is negative in the second because innovations are labor saving in the sense that fewer workers are required to produce higher quality products.

If more productive firms do grow faster, then aggregate productivity growth reflects the fact that workers flow from less to more productive employers as well as from exiting to entering firms. The model developed in the paper provides a useful framework for interpreting empirical growth decomposition exercises such as those reviewed in Foster, Haltiwanger, and Krizan (2001). When output weights are used as required by our model, they find that about 34% of productivity growth in U.S. Manufacturing in the 1977-1987 time period can be attributed to entry while 24% is due to worker reallocation across continuing establishments. Our model implies that the latter figure is zero when firms don't differ with respect to the expected productivity of future innovations.

We find support for the hypothesis that more productive firms grow faster in Danish firm data in the sense that value added is positively associated with value added per worker across firms but employment size is not. By fitting the moments implied by the model to those derived from panel observations of value added, employment, and wages for Danish firms during the period 1992-1997, we also obtain meaningful estimates of the initial distribution of productivity across firms at entry as well as the parameters of the model. These include the overall rate of creative destruction as well as the parameters of the cost of innovation function that all firms are assumed to face.

The remainder of the paper is composed of five sections. In section 2, an adaptation of the Klette-Kortum model of product creation and destruction is introduced. The implication of productive heterogeneity for differences in average firm size and the composition of aggregate productivity growth are developed in section 3. A full general equilibrium model with a competitive labor market is sketched in section 4. Existence of at least one equilibrium solution to the model for the aggregate rate of creative destruction and the wage rate is demonstrated for the case of heterogeneous firms. The empirical evidence and estimation results based on Danish firm data are presented in section 5. The paper concludes with a brief review of the paper's contributions.

2 A Model of Creative Destruction

As is well known, firm employment growth is roughly independent of labor force size; Gebrat's law holds at least as an approximation. Klette and Kortum (2002) construct a stochastic market equilibrium model of firm innovation and growth that consistent with this and other stylized facts regarding firm growth and the size distribution of firms. Although they allow for productive heterogeneity across firms, firm productivity and growth are unrelated because costs and benefits of growth are both proportional to firm productivity in the model. Although we do not make this assumption in our version of the model, the independent of current firm productivity and expected future firm growth is a special case of a more general formulation in which future and current productivity may or may not be correlated.

2.1 Household Preferences

Following Grossman and Helpman (1991), households consume a continuum of different goods indexed by $j \in [0,1]$. Households are identical and live forever. Intertemporal utility of the representative household at time t is

$$U_t = \int_t^\infty \ln C_s e^{-\rho(s-t)} ds \tag{1}$$

where ρ represents the discount rate and

$$\ln C_t = \int_0^1 ln[x_t(j) z_t(j)] dj. \tag{2}$$

is instantaneous unity of consumption where $x_t(j)$ is the service flow of good j at time t, $z_t(j)$ represents the quality of good j at date t. For each good type, the quality level develops through a series of product innovations such that

$$z_t(j) = \Pi_{i=1}^{J_t(j)} q_i(j) \tag{3}$$

where $J_t(j)$ is the number of innovations up to date t and $q_i(j) > 1$ denotes the quantitative improvement in the quality of innovation i over the previous version of good j.

Households can borrow and lend at nominal interest rate r. The household's intertemporal budget constraint is implicit in the following law of motion for interest bearing assets:

$$\frac{da}{dt} = ra + \int_0^1 \pi_t(j)dj + w_t \ell - \int_0^1 p_t(j)x_t(j)dj.$$

In this equation, a represents the net asset position of the household, $\pi_t(j)$ is the profit earned by supplying the j^{th} good, $p_t(j)$ is its price, and w_t is the wage earned by employed participants at time t, and ℓ is the fixed labor endowment.

A household's demands for goods are time paths that maximize intertemporal utility subject to the intertemporal budget constraint and the constraint on the available supply of labor. As households are identical, the only interest rate consistent with equilibrium in the asset market and the necessary transversality condition for intertemporal optimality is the discount rate. Of course, total expenditure by each household is constant when $r = \rho$. Given the form of the utility function, the household spreads it expenditure evenly over the continuum of market good types. Following Grossman-Helpman, total aggregate expenditure is set equal to unity by an appropriate choice of the numeraire. This normalization implies that the marginal utility of income is also unity. Hence, the expenditure flow on each commodity is unity.

2.2 The Value of a Firm

Each individual firm is the monopoly supplier of the products created in the past that have survived to the present. The price it can charge for each is limited by the ability of suppliers of previous version to provide a substitute. In Nash-Bertrand equilibrium, any innovator takes over the market for its good type by setting the price just below that at which consumers are indifferent between the higher quality product supplied by the innovator and an alternative supplied by the previous supplier of the product type. The price charged is the product of the relative quality improvement and the previous producer's marginal cost of production. Given the symmetry of demands for the different good types and the assumption that future quality improvements are independent of the type of good, one can drop the good subscript without confusion.

Labor service is the only factor of production and output per worker is normalized at unity for every product type. Hence, $p = qw$ is the price of the good in terms of the numeraire as well as the value of labor productivity where w represent the marginal cost of production of the previous supplier and $q > 1$ is the step up in quality of the innovation. As total expenditure is normalized at unity and there is a unit measure of product types, it follows that total revenue per product type is also unity, i.e., $px = 1$. Hence, product output and employment are both equal to

$$x = \frac{1}{p} = \frac{1}{wq}. \tag{4}$$

and the gross profit associated with supplying the good is

$$1 > \pi = px - wx = 1 - \frac{1}{q} > 0. \tag{5}$$

Following Klette and Kortum (2002), the discrete number of products supplied by a firm, denoted as k, is defined on the integers. Its value evolves over time as a birth-death process reflecting product creation and destruction. In their interpretation, k reflects the firm's past successes in the product innovation process as well as current firm size. New products are generated by R&D investment. The firm's R&D investment flow generates new product arrivals at frequency γk. The total R&D investment is $wc(\gamma)k$ where $c(\gamma)k$ represents the labor input required in research and development process. The function $c(\gamma)$ is assumed to be strictly increasing and convex. According to the authors, the implied assumption that the total cost of R&D investment is linearly homogenous is the new product arrival rate and the number of existing product, "captures the idea that a firm's knowledge capital facilitates innovation." In any case, this cost structure is needed to obtain firm growth rates that are independent of size as typically observed in the data.

The market for any current product supplied by the firm is destroyed by the creation of a new version by some other firm, which occurs at the rate δ. Below we refer to γ as the firm's creation rate and to δ as the common destruction rate faced by all firms.[1] As product gross profit and product quality are one-to-one, the profits earned on each products reflect a firm's current labor productivity. The firm chooses the creation rate γ to maximize the expected present value of its future net profit flow conditional on information that is relevant for predicting the product profits of future innovations.

Let the parameter θ summarize past profit realizations. We assume that this indicator is a sufficient statistic for prediction the distribution of the next innovation's profit rate. For example,

[1] These are in fact the continuous time job creation and job destruction rates respectively as defined in Davis, Haltiwanger, and Schuh (1996).

the product quality sequence might be a first order Markov process, in which case θ is the profit on the last product innovation. Alternatively, we might think of the problem as one in which a firm's product profitability is initially unknown but can be learned over time by observing the past realization. In Jovanovic's original normal-normal case the sufficient statistic is pair which include both the current estimate of the mean and its precision. In general, θ will be updated in response to the realized profitability of any new product.

Let $\Pi^k = (\pi_1, \pi_2, .., \pi_k)$ denote the firm's vector of profits for the products currently supplied, let $\Pi^{k+1} = (\Pi^k, \pi')$ represent the profits of the $k+1$ products where $\pi_{k+1} = \pi'$, and let $\Pi^k_{\langle i \rangle}$ denote Π^k excluding element $i \in \{1, ..., k\}$. In terms of this notation, the current value of the firm is a function of its state characterized by Π^k and θ. It solves the Bellman equation

$$\begin{aligned} rV_k(\Pi^k, \theta) &= \max_{\gamma \geq 0} \left\{ \sum_{i=1}^{k} \pi_i - wc(\gamma)k + \gamma k \left\{ E\left[V_{k+1}((\Pi^k, \pi'), \theta')|\theta \right] - V_k(\Pi^k, \theta)] \right\} \right. \\ &\quad \left. +\delta \left[\sum_{i=1}^{k} V_{k-1}(\Pi^k_{\langle i \rangle}, \theta) - V_k(\Pi^k, \theta) \right] \right\}. \end{aligned} \tag{6}$$

where $E\{\cdot|\theta\}$ is the expectation operator conditional on information about the quality of the firm's future products and and θ' is the updated value of θ given the realized profit of the next innovation, denoted π'. Notice that no information about future profitability is gained or lost when a product line is destroyed although the firm's scale as reflected in the number of product supplied fall by one unit. The first term on the right side is current gross profit flow accruing to the firms product portfolio less current expenditure of R&D. The second term is the expected capital gain associated with the arrival of a new product line. Finally, because product destruction risk is equally likely across the firm's current portfolio, the last term represents the expected capital loss associated with the possibility that one among the existing product lines will be destroyed.

Consider the conjecture that the solution takes the following additively separable form

$$V_k(\Pi^k, \theta) = \sum_{i=1}^{k} \frac{\pi_i}{r+\delta} + R_k(\theta). \tag{7}$$

That is, we suppose that the value of the firm is the sum of the expected present value of the rents accruing to the firm's current products plus the value of R&D activities. The latter depends only on expectations about the profitability of future innovations and the current number of product lines. Since $V_{k+1}((\Pi^k, \pi'), \theta') = \sum_{i=1}^{k} \frac{\pi_i}{r+\delta} + \frac{\pi'}{r+\delta} + R_{k+1}(\theta')$ under the conjecture, equation (6) can be rewritten as

$$
\begin{aligned}
rV_k(\Pi^k,\theta) &= r\sum_{i=1}^{k}\frac{\pi_i}{r+\delta}+rR_k(\theta) \\
&= \sum_{i=1}^{k}\pi_i + k\max_{\gamma}\left\{\gamma E\left\{\frac{\pi'}{r+\delta}+R_{k+1}(\theta')-R_k(\theta)|\theta\right\}-wc(\gamma)\right\} \\
&\quad -\delta\sum_{i=1}^{k}\frac{\pi_i}{r+\delta}+\delta k\left[R_{k-1}(\theta)-R_k(\theta)\right]
\end{aligned}
$$

Because the term on the left cancels with the two terms on the right that involve the profits of the products currently supplied, the conjecture holds for any sequence of functions $R_k(\theta)$, $k = 1, 2, ...$ that satisfies the difference equation

$$
\begin{aligned}
rR_k(\theta) &= k\max_{\gamma}\left\{\gamma E\left\{\frac{\pi'}{r+\delta}+R_{k+1}(\theta')-R_k(\theta)|\theta\right\}-wc(\gamma)\right\} \\
&\quad +\delta k\left[R_{k-1}(\theta)-R_k(\theta)\right].
\end{aligned} \tag{8}
$$

In words, the return on the value of the R&D department is the expected gain in future profit associated with the next innovation plus the expected capital gains and losses to the R&D operation associated with the possibility of product creation and destruction. In general, these terms are non-zero because a new innovation changes expectations about the profitability of future innovation and because a change in scale affects future returns to and costs of R&D.

Note that equation (8) can be rewritten as

$$
R_k(\theta) = k\max_{\gamma}\left\{\frac{\gamma E\left\{\frac{\pi'}{r+\delta}+R_{k+1}(\theta')|\theta\right\}-wc(\gamma)+\delta R_{k-1}(\theta)}{r+(\delta+\gamma)k}\right\}.
$$

Because the right hand side satisfies Blackwell's sufficient conditions for a contraction that maps the set of non-negative functions defined on the product of the non-negative reals and non-negative integers into itself, a unique solution exists. If the uncertain profit of the next innovation, π', is stochastically increasing in expected profitability as summarized by θ, the unique solution is an increasing function of θ for every value of k by the same argument. Similarly, the fact that the right hand side is strictly increasing in k, $R_{k+1}(\theta')$ and $R_{k-1}(\theta)$ also implies that the contraction maps the functions increasing in k into itself. In sum, the solution has the properties $\theta' > \theta \Rightarrow R_k(\theta') \geq R_k(\theta)$ and $R_{k+1}(\pi) > R_k(\pi)$.

As an implication of (8), a firm's optimal product creation rate maximizes the expected net return to R&D activity:

$$\begin{aligned}\gamma(\theta) &= \arg\max_{\gamma}\left\{\gamma E\{V_{k+1}((\Pi^k,\pi'),\theta')|\theta\} - V_k(\Pi^k,\theta) - wc(\gamma)\right\} \\ &= \arg\max_{\gamma}\left\{\gamma E\left\{\frac{\pi'}{r+\delta} + R_{k+1}(\theta') - R_k(\theta)|\theta\right\} - wc(\gamma)\right\}.\end{aligned} \tag{9}$$

By implication, the expected growth rate, the difference between the chosen creation rate γ and the market determined destruction rate δ, is independent of the firm's current productivity and size if the profitability of the next innovation is independent of past realization of product quality. *When past successes have no consequence for future prospects, there is no incentive for firms that are currently more profitable to grow faster and to become larger.*

3 Deterministic Productive Heterogeneity

In this section, we explore the implications of the case of deterministic productivity dispersion. These are compared with the alternative hypothesis that all firms are exante identical in the sense that a firm's product qualities are iid across innovations.

3.1 Product Creation

We restrict the analysis to the case of deterministic heterogeneity in product quality indexed by $\pi = 1 - q^{-1}$. Namely, assume that the profitability of the every innovation is π with probability one. Since

$$\begin{aligned}rR_k(\pi) &= k\max_{\gamma}\left\{\gamma\left(\frac{\pi}{r+\delta} + R_{k+1}(\pi) - R_k(\pi)\right) - wc(\gamma)\right\} \\ &\quad +\delta k\left[R_{k-1}(\pi) - R_k(\pi)\right]\end{aligned}$$

from (8) in this case, it follows that the solution for $R_k(\pi)$ is proportional to k. Namely, $R_k(\pi) = k\Delta R(\pi)$ where by substitution

$$\Delta R(\pi) = \max_{\gamma\geq 0}\left\{\frac{\gamma\frac{\pi}{r+\delta} - wc(\gamma)}{r+\delta-\gamma}\right\} \tag{10}$$

is the value of R&D per product line for a firm of type π.

From equation (9), an interior solution for the firm's creation rate choice, denoted $\gamma(\pi)$, satisfies the following first order condition:

$$wc'(\gamma) = \frac{\pi}{r+\delta} + \Delta R(\pi) = \max_{\gamma\geq 0}\frac{\pi - wc(\gamma)}{r+\delta-\gamma} \tag{11}$$

Obviously, the optimal creation rate is a strictly increasing function of the firm's profit rate. We conjecture that the latter conclusion also holds when expected profitability is positively correlated with past realization as in the case of learning but we don't have a formal proof.

3.2 The Distribution of Firm Size

As the set of firms with k products at a point in time must either have had k products already and neither lost nor gained another, have had $k-1$ and innovated, or have had $k+1$ and lost one to destruction over any sufficiently short time period, the equality of the flows into and out of the set of firms of type π with $k>1$ product requires

$$\gamma(\pi)(k-1)M_{k-1}(\pi)+\delta(k+1)M_{k+1}(\pi)=(\gamma+\delta)kM_k(\pi)$$

for every π where $M_k(\pi)$ is the steady state mass of firm of type π that supply k products.[2] Because an incumbent dies when it looses its last product but entrants flow into the set of firms with a single product at rate η,

$$\phi(\pi)\eta+2\delta M_2(\pi)=(\gamma(\pi)+\delta)M_1(\pi)$$

where as defined above $\phi(\pi)$ is the fraction of the new entrant flow that realize profit π. Birth must equal deaths in steady state and only firms with one product that looses it die. Therefore, $\phi(\pi)\eta=\delta M_1(\pi)$ and

$$M_k(\pi)=\frac{k-1}{k}\gamma(\pi)M_{k-1}=\frac{\phi(\pi)\eta}{\delta k}\left(\frac{\gamma(\pi)}{\delta}\right)^{k-1} \tag{12}$$

by induction.

The size distribution of firms conditional on type can be derived using equation (12). Specifically, the total mass firms of type π is

$$\begin{aligned} M(\pi) &= \sum_{k=1}^{\infty}M_k(\pi)=\frac{\phi(\pi)\eta}{\delta}\sum_{k=1}^{\infty}\frac{1}{k}\left(\frac{\gamma(\pi)}{\delta}\right)^{k-1} \\ &= \frac{\eta}{\delta}\ln\left(\frac{\delta}{\delta-\gamma(\pi)}\right)\frac{\delta\phi(\pi)}{\gamma(\pi)}. \end{aligned}$$

if finite.Hence, the fraction of type π firm with k product is

$$\frac{M_k(\pi)}{M(\pi)}=\frac{\frac{1}{k}\left(\frac{\gamma(\pi)}{\delta}\right)^{k}}{\ln\left(\frac{\delta}{\delta-\gamma(\pi)}\right)}. \tag{13}$$

This is the logarithmic distribution. Note that the firm size distribution is well defined if and only if the creation rate $\gamma(\pi)$ is less than the overall destruction rate δ. Later we show that this condition must hold in any meaningful market equilibrium.

[2] This equation does not hold in the general case in which an individual firm's type is transitory. In that case, one must also account for type identity switches that occur as new innovations arrive.

Consistent with the observations on firm size distributions, that implied by the model is highly skewed to the right. Furthermore, the mean of distribution conditional on firm profitability,

$$E\{k|\pi\} = \sum_{k=1}^{\infty} \frac{kM_k(\pi)}{M(\pi)} = \frac{\frac{\gamma(\pi)}{\delta-\gamma(\pi)}}{\ln\left(\frac{\delta}{\delta-\gamma(\pi)}\right)},$$

is increasing in $\gamma(\pi)$. Formally, because $(1+a)\ln(1+a) > a > 0$, the expected number of products is increasing in firm profitability,

$$\frac{\partial E\{k|\pi\}}{\partial \pi} = \left(\frac{(1+a(\pi))ln(1+a(\pi)) - a(\pi)}{(1+a(\pi))ln^2(1+a(\pi))}\right)\frac{\delta\gamma'(\pi)}{(\delta-\gamma(\pi))^2} > 0 \tag{14}$$

where $a(\pi) = \frac{\gamma(\pi)}{\delta-\gamma(\pi)}$, if and only if $\gamma'(\pi) > 0$.

3.3 Selection and Worker Reallocation

When permanent differences in product quality exist across firms, workers move from less to more profitable surviving firms as well as from exiting to entering firms. This selection effect can be demonstrated by noting that more profitable firms are over represented relative to their fraction at entry among those that produce more than one product and that this "selection bias" increases with the number of products produced. Namely, the relative fraction of the more profitable firms in the surviving population, given by

$$\frac{M_k(\pi')}{M_k(\pi)} - \frac{\phi(\pi')}{\phi(\pi)} = \frac{\phi(\pi')}{\phi(\pi)}\left[\left(\frac{\gamma(\pi')}{\gamma(\pi)}\right)^{k-1} - 1\right], \tag{15}$$

is positive and increasing in k where $\pi' > \pi$.

The selection effect induced by differential firm rates of product creation has important implications for empirical growth decomposition exercises such as those reviewed by Foster, Haltiwanger, and Krizan (2001). Since every employed worker produces one unit of product per period, the labor productivity improvement attributable to an innovation of quality q relative to the version of the product replaced is $q - 1 = \frac{\pi}{1-\pi}$. In turn, the aggregate rate of labor productivity growth is the product of the innovation rate and the average relative productivity improvement of entrants and surviving firms. Formally, in either special case

$$\frac{\dot{P}}{P} = \eta\sum_{\pi}\left(\frac{\pi}{1-\pi}\right)\phi(\pi) + \sum_{\pi}\gamma(\pi)\left(\frac{\pi}{1-\pi}\right)\sum_{k=1}^{\infty}kM_k(\pi)$$

where $\phi(\pi)$ is the fraction of new firms that enter with a product of quality $q = \frac{1}{1-\pi}$ and $M_k(\pi)$ is the steady state fraction of incumbent firms of type π that supply k products as defined above.

Let

$$\overline{\gamma} = \sum_{\pi} \gamma(\pi) \sum_{k=1}^{\infty} k M_k(\pi) \tag{16}$$

represent the average product creation rate of incumbent firms. Then, the equation

$$\begin{aligned} \frac{\dot{P}}{P} &= \eta \sum_{\pi} \left(\frac{\pi}{1-\pi}\right) \phi(\pi) + \overline{\gamma} \sum_{\pi} \left(\frac{\pi}{1-\pi}\right) \sum_{k=1}^{\infty} k M_k(\pi) \\ &\quad + \sum_{\pi} [\gamma(\pi) - \overline{\gamma}] \left(\frac{\pi}{1-\pi}\right) \sum_{k=1}^{\infty} k M_k(\pi) \end{aligned} \tag{17}$$

represents a decomposition of the aggregate rate of productivity growth into three parts attributable to the entry, average within firm growth in productivity, and a between firm component respectively. The between firm components captures the growth in productivity attributable to the movement of workers from less to more productive incumbent firms. Specifically, the last term is zero if there is only one firm type or if product quality is iid across innovations. Because $\pi/(1-\pi)$ is increasing in π and the productivity contingent size distribution of firms is stochastically increasing in π, the between firm share of productivity growth is strictly positive if $\gamma'(\pi) > 0$.

The empirical literature on the sources of aggregate productivity growth, recently reviewed by Foster, Haltiwanger, and Krizan (2001), suggests that entry and within establishment productivity growth are both important. However, the evidence for reallocation across surviving firms as a source of aggregate growth is mixed. This literature starts by defining the level of aggregate productivity at time t as measured by an index

$$P_t = \sum_{i \in I_t} s_{it} p_{it}$$

where p_{it} is represents a measure of (labor or total factor) productivity of the i^{th} firm or establishment in period t and s_{it} is the (employment or output) share in period t. Hence, aggregate growth in the index can be be represented as

$$\begin{aligned} \frac{\Delta P_t}{P_t} &= \sum_{i \in E_t} s_{it+1} \left(\frac{p_{it+1}}{P_t} - 1\right) - \sum_{i \in X_t} s_{it} \left(\frac{p_{it}}{P_t} - 1\right) \\ &\quad + \sum_{i \in C_t} s_{it} \frac{\Delta p_{it}}{P_t} + \sum_{i \in C_t} \left(\frac{p_{it} + \Delta p_{it}}{P_t} - 1\right) \Delta s_{it} \end{aligned} \tag{18}$$

where C_t represent the set of continuing units, E_t is the set of those entering and X_t denotes the set of those exiting the market in period t. In order, the terms represent the contributions to overall productivity growth of entering, exiting, and continuing units where the latter is further decomposed into within and between components. Our theoretical decomposition equation (17) corresponds exactly to equation (18) in steady state.

Table 1: U.S. Manufacturing Productivity Growth Decomposition, 1977-1987. Growth Rates and Source Shares.

Measure	Weight	Growth Rate	Net Entry	Within	Between
TFP	Gross Output	10.24	26%	48%	26%
Labor	Gross Output	25.56	31%	45%	24%
Labor	Manhours	21.32	29%	77%	-6%
Labor	Employment	23.02	29%	74%	-3%

Source: Foster, Haltiwanger, and Krizan (2001), Table 4.

To establish the claim of equivalence, note that the steady state mass of firms of type π that supply a single product is $M_1(\pi) = \frac{\eta}{\delta}\phi(\pi)$. Because the relative frequency distribution of types in this set is the same as that of entrants, on the one hand, and because only firms with one product are subject to exit risk in any sufficient short time interval on the other, every new successful entrant of each type can be exactly matched with an exiting firm of the same type. Similarly, because the number of firms of each type with k products that gain another and the number that loose one per period are equal, the within and between components of (18) correspond respectively to the last two term of (17) in steady state. As a measurement issue, it is also important to note that all innovations are equally weighted in equation (17). Since the value of sales are the same for both the innovator and the current supplier, namely $px = 1$ for every product, the implicit labor productivity index, P, can be considered as one in which a firm's labor productivity is weighted by the firm's share of value of output rather than by its share of employment.

Table 1 summarizes the shares of the three components of productivity growth represented on the right side of (18) reported by Foster, Haltiwanger, and Krizan (2001) for U.S. Manufacturing establishment data over the period 1977-1987. The results very across the rows due to differences in the productivity concept and the weights used to construct the productivity index. The two measures of productivity considered are multifactor productivity (TFP) and output per worker or hour. Obviously, these results support the hypothesis that more profitable firm grow at a more rapid rate ($\gamma'(\pi) > 0$) when output weights are used in the calculation of shares as is implicit in the theoretical equation (17). It is not clear what to infer from the fact that the between term is essentially zero when labor input weights are used. The interpretation problem is even cloudier given the fact that more profitable firms employ more R&D labor in the model given that measures of labor input used in these calculations do not distinguish between production and other types of workers in the firm.

4 Market Equilibrium

In this section, we complete the specification of the market model and establish existence of an equilibrium solution in the case of deterministic productive heterogeneity.

4.1 Firm Entry and Labor Market Clearing

The entry of a new firm requires an innovation. The cost of entry is the expected cost of the R&D effort required of a potential entrant to discover and develop a new successful product. Hence, if a potential entrant obtains ideas for new products at frequency h per period, the expected opportunity cost of her effort per innovation is w/h, the expected earnings forgone during the required period of R&D activity. As no entrant knows the profitability of its product a priori but all know its distribution, new firms enter if and only if the expected value of a new product given no entry exceeds the cost. Assuming that the condition holds, the endogenous equilibrium product destruction rate, δ, adjusts though entry to equate the expected cost and return. Given that product quality at entry is uncertain but that its distribution is common knowledge, the equality of the expected return and cost of entry require that

$$\sum_{\pi} V_1(\pi,\theta)\phi(\pi) = \sum_{\pi} \max_{\gamma\geq 0}\left\{\frac{\pi - wc(\gamma)}{r+\delta-\gamma}\right\}\phi(\pi) = \frac{w}{h} \tag{19}$$

from equations (7) and (11) where $\phi(\pi)$ is fraction of entrants with product quality $q=(1-\pi)^{-1}$.[3]

Because the new product arrival rate of a firm of type π with k products is $\gamma(\pi)k$ and the measure of such firms is $M_k(\pi)$, the aggregate rate of destruction is the sum of the entry rate and the creation rates of all the incumbents given that the mass of products is fixed. That is

$$\begin{aligned}
\delta &= \eta + \sum_{\pi}\sum_{k=1}^{\infty}\gamma(\pi)kM_k(\pi) = \eta + \sum_{\pi}\sum_{k=1}^{\infty}\gamma(\pi)\frac{\phi(\pi)\eta}{\delta}\left(\frac{\gamma(\pi)}{\delta}\right)^{k-1} \\
&= \eta\left(1+\sum_{\pi}\frac{\phi(\pi)\gamma(\pi)}{\delta}\sum_{k=1}^{\infty}\left(\frac{\gamma(\pi)}{\delta}\right)^{k-1}\right) = \eta\left(\sum_{\pi}\frac{\delta\phi(\pi)}{\delta-\gamma(\pi)}\right).
\end{aligned}$$

where the second equality follows from (12) and the last equality requires that the aggregate rate of creative destruction exceeds the creation rate for every firm type. Using the assumption that the measure of firms is unity, a direct derivation of the same relationship follows:

$$1 = \sum_{\pi}\sum_{k=1}^{\infty}kM_k(\pi) = \sum_{\pi}\frac{\eta\phi(\pi)}{\delta}\sum_{k=1}^{\infty}\left(\frac{\gamma(\pi)}{\delta}\right)^{k-1} = \eta\sum_{\pi}\frac{\phi(\pi)}{\delta-\gamma(\pi)}. \tag{20}$$

[3]For simplicity, we assume that the number of different product qualities is finite.

Note that $\delta - \gamma(\pi) > 0$ for all π in the support of the entry distribution from (20) if and only if entry is the entry rate η is positive. Below, we will seek a market solution that satisfies this property. In general, restrictions on fundamental parameters are required to insure that the condition holds.

There is a fixed measure of available workers, denoted by ℓ, seeking employment at any positive wage. In equilibrium, these are allocated across production and R&D activities, those performed by both incumbent firms and potential entrants. Since the number of workers employed for production purposes per product of quality q is $x = 1/wq = (1-\pi)/w$ from equations (4) and (5), the total number demanded for production activity by firms of type π with k products is $\ell_x(k,\pi) = k(1-\pi)/w > 0$. The number of R&D workers employed by incumbent firms of type π with k products is $\ell_R(k,\pi) = kc(\gamma(\pi))$. Because a potential entrant innovates at frequency h, the total number so engaged in R&D is $\ell_E = \eta/h$ given entry rate η. Hence, the equilibrium wage satisfies the labor market clearing condition

$$\begin{aligned}
\ell &= \sum_{\pi}\sum_{k=1}^{\infty}\left[\ell_x(k,\pi)+\ell_R(k,\pi)\right]M_k(\pi)+\ell_E \qquad (21)\\
&= \sum_{\pi}\sum_{k=1}^{\infty}\left(\frac{1-\pi}{w}+c(\gamma(\pi))\right)kM_k(\pi)+\frac{\eta}{h}\\
. &= \sum_{\pi}\left(\frac{1-\pi}{w}+c(\gamma(\pi))\right)\frac{\phi(\pi)\eta}{\delta}\sum_{k=1}^{\infty}\left(\frac{\gamma(\pi)}{\delta}\right)^{k-1}+\frac{\eta}{h}\\
&= \eta\left(\frac{1}{h}+\sum_{\pi}\left(\frac{1-\pi}{w}+c(\gamma(\pi))\right)\frac{\phi(\pi)}{\delta-\gamma(\pi)}\right)
\end{aligned}$$

where again the last equality is implied by equation (12) and the requirement that $\delta > \gamma(\pi)$ for all π.

4.2 Existence

Definition A steady state *market equilibrium*with positive entry is a triple composed of a labor market clearing wage w, a positive entry rate η, and positive creative destruction rate δ that satisfy equation (19), (20), and (21).

From (19), the free entry condition is

$$\sum_{\pi}\max_{\gamma\geq 0}\left\{\frac{\pi - wc(\gamma)}{r+\delta-\gamma}\right\}\phi(\pi) = \frac{w}{h} \qquad (22)$$

By using equation (20) to eliminate the positive entry rate η and equation (22) to eliminate w/h

in equation (21), one can write the result as

$$\begin{aligned} w\ell\left(\sum_{\pi}\frac{\delta}{\delta-\gamma(\pi)}\phi(\pi)\right) &= \delta\left(\sum_{\pi}\left(\frac{1-\pi+wc(\gamma(\pi))}{\delta-\gamma(\pi)}+\max_{\gamma\geq 0}\frac{\pi-wc(\gamma)}{r+\delta-\gamma}\right)\phi(\pi)\right) \\ &= \sum_{\pi}\frac{\delta}{\delta-\gamma(\pi)}\left(1-r\max_{\gamma\geq 0}\frac{\pi-wc(\gamma)}{r+\delta-\gamma}\right)\phi(\pi) \end{aligned}$$

where the first equality is implied by the fact that $\sum_{\pi}\phi(\pi)=1$ and the second is a consequence of the fact that $\gamma(\pi)$ is the optimal choice of the creation rate for a type π firm. Hence,

$$1 = w\ell + \frac{r\sum_{\pi}\left(\max_{\gamma\geq 0}\frac{\pi-wc(\gamma)}{r+\delta-\gamma}\right)\frac{\phi(\pi)}{\delta-\gamma(\pi)}}{\sum_{\pi}\frac{\phi(\pi)}{\delta-\gamma(\pi)}}. \tag{23}$$

Since total value added is unity by choice of the numeraire, this expression is the income identity. Namely, the total wage bill plus the return on the values of all the operating firms in the economy is equal to value added.

In order to focus on the case in which incumbents invest in R&D, we assume their cost, $c(\gamma)$, is strictly convex and that $c(0)=c'(0)=0$. Under these restrictions, the optimal creation rate for each type conditional on the market wage and rate of creative destruction is uniquely determined by the following first order condition state as equation (11). Since the optimal creation rate is strictly increasing in productivity and strictly decreasing in the market wage, a necessary and sufficient condition for the optimal choice to be less than the rate of creative destruction, $\gamma(\pi)<\delta$, at any point (w,δ) is

$$wc'(\delta) > \frac{\pi-wc(\delta)}{r} \Leftrightarrow w > \frac{\pi}{rc'(\delta)+c(\delta)}\forall\pi. \tag{24}$$

Of course, the entry rate, η, is positive only in the union of these regions from equation (20).

The boundary of the admissible set, that defined by (24) with $\pi=\overline{\pi}$, the upper support of the given distribution of profit at entry, is labeled BB in Figure 1. All pairs above BB satisfy the requirement that $\gamma(\pi)<\delta$ because $\gamma(\pi)\leq\gamma(\overline{\pi})$ for all $\pi\leq\overline{\pi}$. As illustrated, the wage on the boundary is positive, tends to infinity as δ tend to zero, is strictly decreasing in δ, and tends to zero as δ tends to infinity given the assumed properties of the R&D cost function.

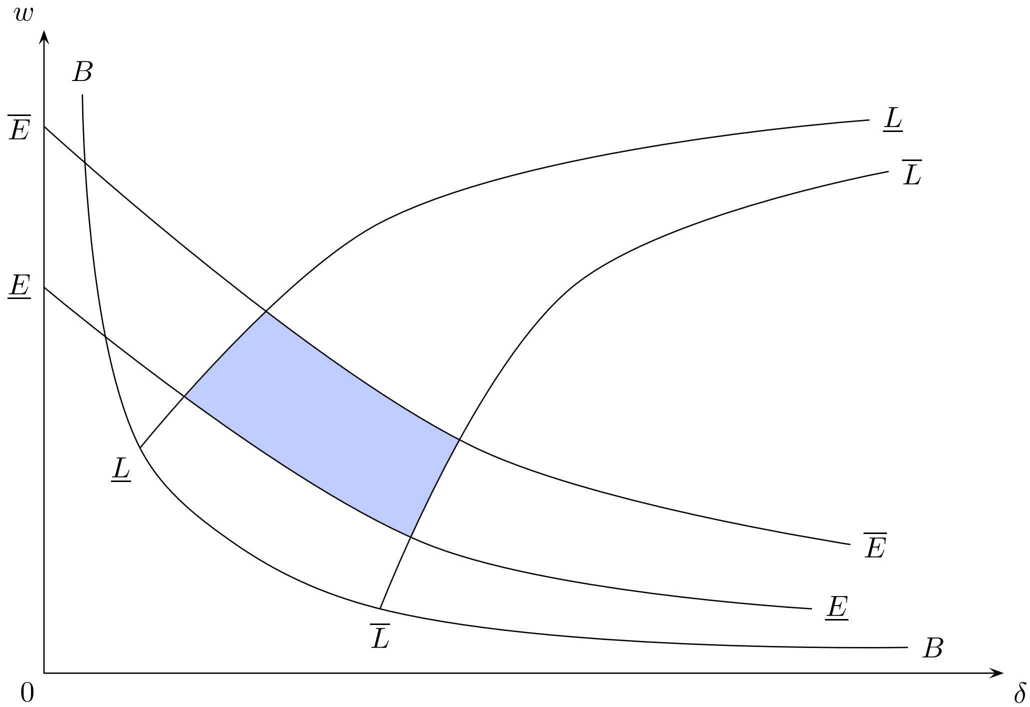
An equilibrium is any (w,δ) pair satisfying equation (22) and (23) provided that $\delta\geq\gamma(\pi)$ on the support of the distribution of profits at entry. Let $w=E_{\pi}(\delta)$ represent the locus of points defined by

$$\max_{\gamma\geq 0}\frac{\pi-wc(\gamma)}{r+\delta-\gamma}=\frac{w}{h} \tag{25}$$

and let $w=L_{\pi}(\delta)$ represent solution to

$$1 = w\ell + r\left(\max_{\gamma\geq 0}\frac{\pi-wc(\gamma)}{r+\delta-\gamma}\right) \tag{26}$$

Figure 1: Equilibrium Wage and Creative-Destruction Rates.



in the region defined by (24). Since $E'_\pi(\delta) < 0$ and $L'_\pi(\delta) > 0$, at most one solution exists to both given π. Furthermore, because the solution to (25) for w is monotone increase in h and because both (24) and (26) are independent of h, a solution exists in the required region for all values of h above some critical value.

As equations (22) and (23) collapse to (25) and (26) respectively when there is a single firm type, we have established sufficient conditions for both existence and uniqueness in this case. In the case of firm heterogeneity, the same argument implies existence. Specifically, because $\sum_\pi \phi(\pi) = 1$ where $\phi(\pi)$ is the fraction of entrants of type π and because $E_\pi(\delta)$ is increasing in π from (25), the locus of point that satisfy the entry condition (22) is bounded above by $E_{\overline{\pi}}(\delta)$ and below by $E_{\underline{\pi}}(\delta)$ where $\underline{\pi}$ is the lower and $\overline{\pi}$ is the upper support of the type distribution at entry. Similarly, because $L_\pi(\delta)$ is decreasing in π from (26), the solution to the labor market clearing condition (23) is bounded above by $L_{\underline{\pi}}(\delta)$ and below by $L_{\overline{\pi}}(\delta)$.

In Figure 1, the curves $\overline{LL}$ and $\underline{LL}$ represent $w = L_{\overline{\pi}}(\delta)$ and $w = L_{\underline{\pi}}(\delta)$ respectively. Similarly, $w = E_{\overline{\pi}}(\delta)$ and $w = E_{\underline{\pi}}(\delta)$ are represented as $\overline{EE}$ and $\underline{EE}$. It follows that any joint solution to the entry and labor market clearing conditions must lie in the shaded area. Given continuity of the relationship, at least one common solution exists to (22) and (23) in that region. Finally, the shaded area lies above BB in the figure for all sufficiently large values of h because $E_\pi(\delta)$ is monotone increasing in h and both (24) and (26) are independent of h. Indeed, the critical value is that for which the intersection of $w = L_{\underline{\pi}}(\delta)$ and $w = E_{\overline{\pi}}(\delta)$ lies on the boundary. Since $\widehat{w} = h/(r + h\ell)$ at any joint solution to equations (25) and (26), the critical value of h, denoted $\widehat{h}$, and the associated rate of creative destruction at the intersection, $\widehat{\delta}$, are the unique solutions to

$$\frac{1-\underline{\pi}}{\ell - c(\widehat{\delta})} = \widehat{w} = \frac{\widehat{h}}{r + \widehat{h}\ell} = \frac{\overline{\pi}}{rc'(\widehat{\delta}) + c(\widehat{\delta})}.$$

A unique triple $(\widehat{w}, \widehat{\delta}, \widehat{h})$ exists under the hypothesis to the following result.

Proposition 1 *If the cost of R&D function, $c(\gamma)$, is strictly convex and $c'(0) = c(0) = 0$, then a steady state market equilibrium exists for all $h > \widehat{h}$. In the case of a single firm type, there is only one.*

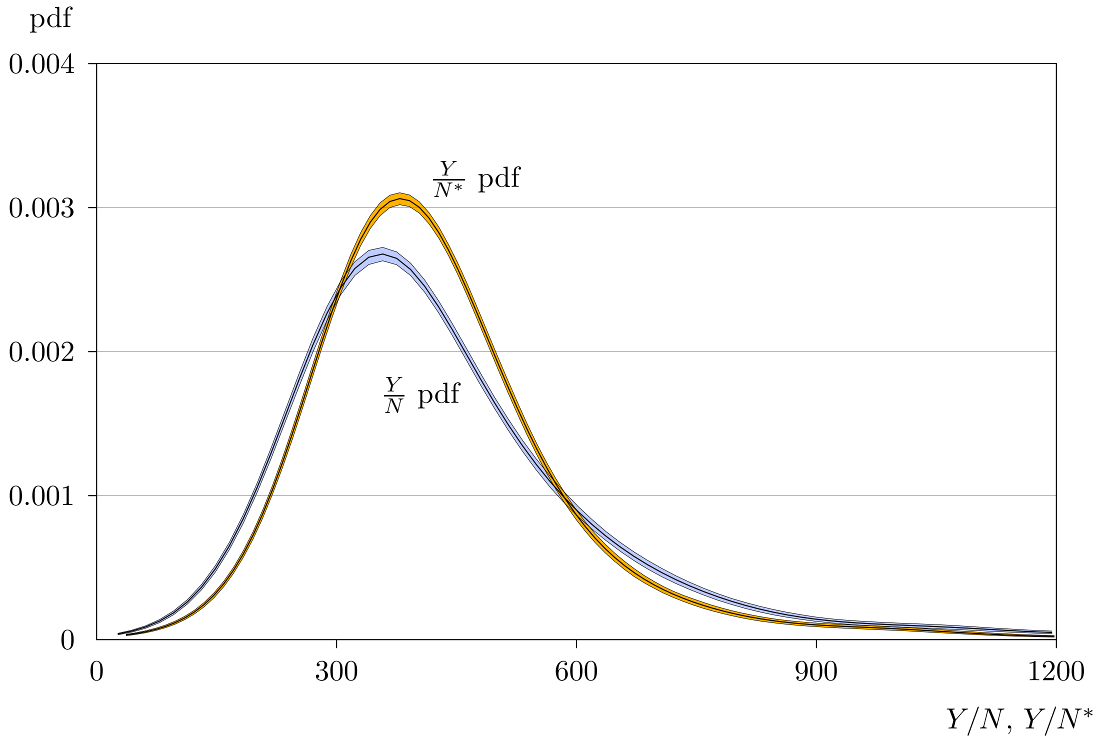
5 Evidence and Estimation

If product quality is a permanent firm characteristic, then differences in firm profitability are associated with differences in the product creation rates chosen by firms. Specifically, more profitable firms grow faster, are more likely to survive in the future, and supply a larger number of produces on average. Hence, a positive cross firm correlation between current gross profit per product and sales volume should exist. Furthermore, worker reallocation from slow growing firms that supply products of lesser quality to more profitable fast growing firms will be an important sources of aggregate productivity growth. On the other hand, if product quality is iid across innovations and firms, all firms grow at the same rate even though persistent differences in profitability exist as a consequence of different realizations of product quality histories. In the section, we demonstrate that firm specific differences in profitability are required to explain Danish the interfirm relationships between value added, employment, and wages paid. In the process of fitting the model to the data, we also obtain estimates of the investment cost of innovation function that all firms face as well as the sampling distribution of firm productivity at entry.

5.1 Danish Firm Data

If more productive firm's grow faster in the sense that $\gamma'(\pi) > 0$, then (14) implies that more productive firms also supply more products and sell more on average. However, because production employment per product decreases with productivity, total expected employment, nEk where $n = (1-\pi)/w + c(\gamma(\pi))$, need not increase with π in general and decreases with π when growth is

Figure 2: Value Added per Worker and per Standardized Worker pdf's.



Note: The shaded areas represent 90% confidence intervals based on bootstrapping.

independent of a firm's past product quality realizations. These implications of the theory can be tested directly.

Danish firm data provide information on the relationships among productivity, employment, and value added. The available data set is an annual panel of privately owned firms for the years 1992-1997 drawn from the Danish Business Statistics Register. The sample of approximately 6,700 firms is restricted to those with 20 or more employees. The variables observed in each year include value added (Y), full-time equivalent employment (N), and the total wage bill (W). The model is estimated on an unbalanced panel of 5,254 firms drawn from the firm panel. The panel is constructed by selecting all existing firms in 1992 and following them through time, while all firms that enter the sample in the subsequent years are excluded. Furthermore, the top and bottom 1% of the firms in the value added distribution for 1992 are censored from the panel to ease numerical challenges in the estimation and to avoid extreme observation bias. The censoring means a loss of roughly 110 firms.

Figure 2 presents non-parametric estimates of the distributions of two alternative measures labor productivity. The first measure is value added per worker (Y/N) while the second is valued added per unit of quality adjusted employment (Y/N^*). The first measure misrepresents cross firm productivity differences to the extent that labor force quality differs across firms. However, if more productive workers are compensated with higher pay as would be true in a competitive labor market, one can use a wage weighted index of employment to correct for this source of cross firm differences in productive efficiency. Formally, the constructed quality adjusted employment of firm j is defined as $N_j^* = \frac{W_j}{w}$where $w = \sum_j W_j / \left(\sum_j N_j\right)$is the average wage paid per worker in the market. Although correcting for wage differences across firms in this manner does reduce the spread and skew of the implied productivity distribution somewhat, both distributions have high variance and skew and are essentially the same shape.

Figure 3 illustrates non-parametric regressions of value added and employment size on the two productivity measures. The top and bottom curves in the figures represent a 90% confidence interval for the relationship. Hence, these results strongly reject the hypothesis that firm growth is independent of the firm's profitability in favor of the alternative that the sales of more productive firms grow larger.

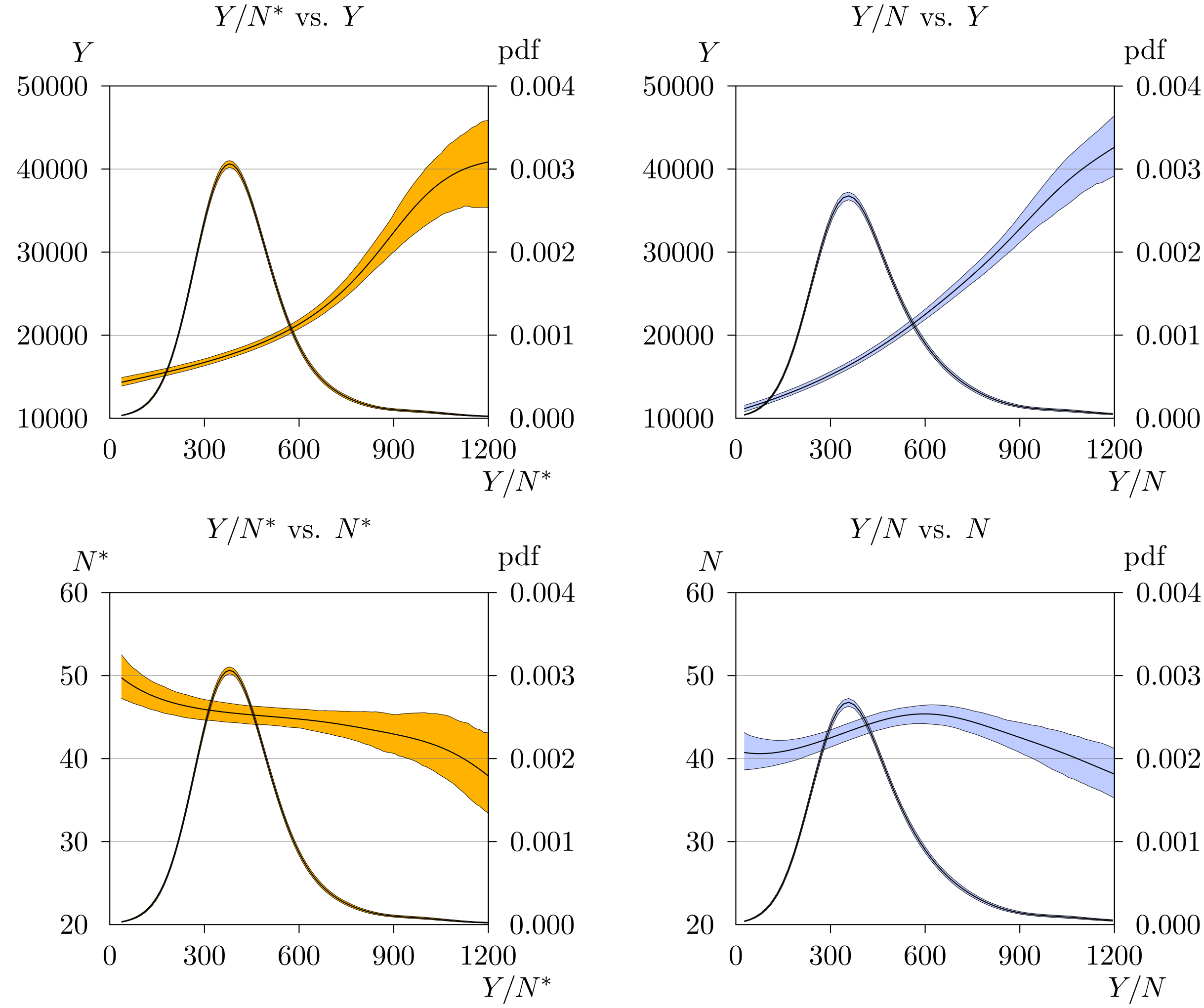
5.2 Model Estimation

The following identifies the deterministic permanent firm types case of the model. An observation in the panel is given by $\psi_{jt} = \left(Y_{jt}, W_{jt}, N_{jt}^*\right)$, where Y_{jt} is real value added, W_{jt} the real wage sum, and N_{jt}^* quality adjusted labor force size of firm j in year t. Let ψ_j be defined by, $\psi_j = \left(\psi_{j1,...,}\psi_{jT}\right)$.

The model is estimated by use of a simulated minimum distance estimator as described in for example Gourieroux, Monfort, and Renault (1993), Hall and Rust (2003), and Alvarez, Browning, and Ejrnæs (2001). First, define a set of sample auxiliary parameters, $\Gamma(\psi_1, \ldots, \psi_J)$, which in this case takes a cross-section form for each time period. Specifically, 10 data moments are generated for each year: Number of surviving firms, $E[Y]$, $Std[Y]$, $E[W]$, $Std[W]$, $Corr[Y, W]$, $Corr[Y/N^*, Y|Y > 0]$, $Corr[Y/N^*, N^*|Y > 0]$, $Median[Y|Y > 0]$, $Median[W|W > 0]$. Thus, $\Gamma(\cdot)$ consists of 60 moments.

Second, $(\psi_1^s(\omega), \ldots, \psi_J^s(\omega))$ is simulated from the model for a given set of model parameters ω. The model simulation is initialized by assuming that the economy is in steady state in the first year and consequently that firm observations are distributed according to the ω-implied steady state distribution. Alternatively, one can initialize the simulation according to the observed data in the first year, $(\psi_{11}, \ldots, \psi_{1J})$. The assumption that the economy is initially in steady state provides

Figure 3: Regressions of Value Added per Worker against Firm Size (1992).



Note: The shaded areas represent 90% confidence intervals based on bootstrapping.

aditional identification in that $(\psi_{11}, \ldots, \psi_{1J})$ can be compared to the model-implied steady state distribution $(\psi^s_{11}(\omega), \ldots, \psi^s_{1J}(\omega))$. The simulated auxiliary parameters are then given by,

$$\hat{\Gamma}(\omega) = \frac{1}{S} \sum_{s=1}^{S} \Gamma(\psi^s_1(\omega), \ldots, \psi^s_J(\omega)),$$

where S is the number of simulations.

The estimator is then the choice of parameters that minimizes the weighted distance between the sample auxiliary parameters and the simulated auxiliary parameters,

$$\hat{\omega} = \arg\min_{\omega \in \Omega} \big(\hat{\Gamma}(\omega) - \Gamma(\psi_1, \ldots, \psi_J)\big)' A^{-1} \big(\hat{\Gamma}(\omega) - \Gamma(\psi_1, \ldots, \psi_J)\big),$$

where A is some positive definite matrix. If A is the identity matrix, $\hat{\omega}$ is the equally weighted minimum distance estimator (EWMD). If A is the covariance matrix of the auxiliary parameter vector $\Gamma(\cdot)$, $\hat{\omega}$ is the optimal minimum distance estimator (OMD). The OMD estimator is asymptotically more efficient than the EWMD estimator. However, Altonji and Segal (1996) show that the estimate of A as the second moment matrix of $\Gamma(\cdot)$ may suffer from serious small sample bias. Horowitz (1998) suggest an alternative estimator of A based on bootstrap methods. The analysis will adopt Horowitz's (1998) estimator of the covariance matrix A.

5.3 Model Simulation

To fit the data, the model simulation produces time paths for value added (Y), the wage sum (W), and labor force size (N) for J firms. Rather than normalize the total consumer expenditure for each product at unity, the expenditure for each product is set at Z. Hence, the demand for each good is $x_j = Z/p_j$. Denote by k_{jt} the number of products of firm j at time t. Let the type of firm j be represented by its quality improvement q_j.

To properly capture the labor share in the data, a capital cost $\kappa \equiv K/Z$ is added to the model. K is the capital associated with the production of a given product. κ is the capital cost relative to product expenditure. This modifies the pricing of the intermediary goods. Now, providing an intermediary good at price p yields operational profits, $Z(1 - w/p) - K$. Thus, the price of the intermediary goods for which firm j is the quality leader is, $p_j = q_j w/(1-\kappa)$. Firm j's total profits at time t is given by,

$$\begin{aligned}
\Pi_{jt} &= k_{jt}\left[p_j x_j - w x_j - K - wc\big(\gamma(\pi_j)\big)\right] \\
&= k_{jt}\left[Z - \frac{Z}{q_j}(1-\kappa) - K - wc\big(\gamma(\pi_j)\big)\right] \\
&= k_{jt} Z\left[\pi_j - \kappa - w\tilde{c}\big(\gamma(\pi_j)\big)\right],
\end{aligned}$$

where $\pi_j \equiv 1 - (1-\kappa)/q_j$.

The value added of firm i at time t (Y_{jt}) is given by,

$$\log Y_{jt} = \log k_{jt} + \log Z + \varepsilon_y, \tag{27}$$

where ε_Y is a noise term which can be interpreted as measurement error and/or demand side shocks. ε_y is assumed iid with $E[\varepsilon_y] = 0$ and $Var[\varepsilon_y] = \sigma^2_{\varepsilon_y}$.The wage bill of firm j at time t (W_{jt}) is given by,

$$\begin{aligned}
\hat{W}_{jt} &= k_{jt}\left(w\frac{Z}{wq_j}(1-\kappa) + wc\big(\gamma(\pi_j)\big)\right) \\
&= k_{jt} Z\left[1 - \pi_j + w\tilde{c}\big(\gamma(\pi_j)\big)\right],
\end{aligned}$$

where $\tilde{c}(\gamma) = c(\gamma)/Z$. Define the labor share of firm j by,

$$\alpha_j \equiv 1 - \pi_j + w\tilde{c}\big(\gamma(\pi_j)\big).$$

Firm j's wage bill at time t is then given by,

$$\log W_{jt} = \log k_{jt} + \log Z + \log \alpha_j + \varepsilon_W, \tag{28}$$

where ε_W is another iid noise term with $E[\varepsilon_W] = 0$ and $Var[\varepsilon_W] = \sigma^2_{\varepsilon_W}$.

By (11), firm j's choice of creation rate solves,

$$\gamma(\pi_j) = \arg\min_{\gamma} \frac{\pi_j - \kappa - w\tilde{c}(\gamma)}{r + \delta - \gamma}. \tag{29}$$

Specify the cost function $\tilde{c}(\gamma) = c_0\gamma^{1+c_1}$. Then the first order condition for the optimal creation rate choice is,

$$w(1+c_1)c_0\gamma^{c_1}(r+\delta-\gamma) = \pi_j - \kappa - wc_0\gamma^{1+c_1}.$$

Substituting the first order condition into the definition of the labor share yields,

$$\alpha_j = 1 - \kappa - \big(r + \delta - \gamma(\pi_j)\big)w(1+c_1)c_0\gamma^{c_1}. \tag{30}$$

(27) and (28) provide the foundation for the model simulation. It then remains to simulate product paths for all firms. The simulation is initialized by the assumption of steady state. Let $G(\pi)$ be the unknown steady state distribution of firm types. To simplify matters discretize the support of the type distribution to $(\pi_1, \ldots, \pi_M)$. By (13), the steady state product size distribution conditional on survival is given by,

$$\Pr(k^* = k|\pi) = \frac{\frac{1}{k}\left(\frac{\gamma(\pi)}{\delta}\right)^k}{\ln\left(\frac{\delta}{\delta - \gamma(\pi)}\right)}. \tag{31}$$

First, firm j's type, π_j, is determined according to $G(\cdot)$. Then, the initial product size of a firm j (k_{j1}) is determined according to (31).

With a given initial product size k_{j1}, simulation of the subsequent time path requires knowledge of the transition probability function $\Pr(k_{j2} = k|k_{j1}, \pi_j)$. Denote by $p_{\pi,n}(t)$ the probability of a type π firm having product size n at time t. As shown in Klette and Kortum (2002), $p_{\pi,n}(t)$ evolves according to the ordinary differential equation system,

$$\begin{aligned} \dot{p}_{\pi,n}(t) &= (n-1)\gamma(\pi)p_{\pi,n-1}(t) + (n+1)\delta p_{\pi,n+1}(t) - (\delta + \gamma(\pi))p_{\pi,n}(t), \ \forall n \geq 1 \\ \dot{p}_{\pi,0}(t) &= \delta p_{\pi,1}(t). \end{aligned} \tag{32}$$

Hence, with the initial condition,

$$p_{\pi,n}(0) = \begin{cases} 1 \text{ if } n = k_{j1} \\ 0 \text{ otherwise.} \end{cases} \tag{33}$$

one can determine $\Pr(k_{j2} = k | k_{j1}, \pi_j)$ by solving the differential equation system in (32) for $p_{\pi_j,k}(1)$. Solving for $p_{\pi_j,k}(1)$ involves setting an upper reflective barrier to bound the differential equation system. It has been set sufficiently high so as to avoid biasing the transition probabilities. Based on the transition probabilities $\Pr(k_{jt+1} = k | k_{jt}, \pi_j)$ one can then iteratively simulate product size paths for each firm.

5.4 Identification

The set of model parameters to be identified (ω) is given by,

$$\omega = \{c_0, c_1, \delta, \kappa, Z, (\pi_1, \ldots, \pi_M), (p_1, \ldots, p_M)\} \in \Omega,$$

where $p_m = \Pr(\pi = \pi_m)$. and Ω is the feasible set of model parameters choices. The interest rate will be set at $r = .05$ and the noise processes governing ε_Y and ε_W will be taken as given. The wage w is immediately identified as the average worker wage in the sample $w = 221.73$. Since $\sum_{m=1}^{M} p_m = 1$, this implies that the estimation will be identifying $2M + 4$ parameters. Notice that to simulate product size paths and generate ψ_j^s according to (27) and (28), it is necessary and sufficient to know

$$\psi = \{\delta, Z, (\gamma_1, \ldots, \gamma_M), (\alpha_1, \ldots, \alpha_M), (p_1, \ldots, p_M)\},$$

which is $3M + 1$ parameters. The choice of ω maps into ψ according to (29) and (30). Denote the mapping by, $\psi = \Psi(\omega)$. The dimension of ω is strictly greater than ψ if $M \leq 2$. Thus, in the case where there are less than 3 distinct productivity types, there may be multiple ω choices that map into the same ψ which suggests a fundamental identification problem in these cases. Suppose $M \leq 2$ and there exists a $\omega' \in \Omega$ different from $\omega'' \in \Omega$ such that $\psi = \Psi(\omega') = \Psi(\omega'')$. In this case, the simulated data is the same for ω' and ω'', that is $(\psi_1^s(\omega'), \ldots, \psi_J^s(\omega')) = (\psi_1^s(\omega''), \ldots, \psi_J^s(\omega''))$, and the distance criterion for the SMD estimator will be the same for ω' and ω''. The example suggests a potential for failure of identification for $M \leq 2$.[4]

When $M \geq 3$, the dimension of ω is greater or equal to the dimension of ψ. While a choice of $M \geq 3$ resolves the identification problem associated with the mapping between ω and ψ, it remains necessary that there is enough identifying variation in the data to identify the $2M + 4$

[4]Indeed, experimentation with estimation of the model with $M = 2$ resulted in serious identification problems. The estimation pointed to a region of parameter values but failed to identify an actual point estimate.

Table 2: Model Parameter Estimates

	Point Estimate	Lower Bound	Upper Bound
c_0	21.569	–	–
c_1	4.832	–	–
κ	0.360	–	–
Z	$10,086.943$	–	–
δ	0.177	–	–
π_1	0.374	–	–
π_2	0.583	–	–
π_3	0.589	–	–
$\Pr(\pi_1)$	0.465	–	–
$\Pr(\pi_2)$	0.337	–	–
$\Pr(\pi_3)$	0.197	–	–
γ_1	0.072	–	–
γ_2	0.139	–	–
γ_3	0.140	–	–
α_1	0.628	–	–
α_2	0.465	–	–
α_3	0.461	–	–
w	221.734	–	–
r	0.05	–	–

model parameters. This is the standard identification problem and increasing M will all else equal strain identification on this dimension. The model is estimated for $M = 3$ and turns out to be identified under this choice.

5.5 Estimation Results

The model parameter estimates are given in table 2. The creation rates γ_m and labor shares α_m for each type are derived from the model parameter estimates. The interest rate has been set at $r = .05$ and the wage level is identified as the average worker wage in the data for 1992. The lower and upper bounds of the double sided 90% confidence interval are generated by naive bootstrapping.

Table 3 produce a comparison of the data moments and the simulated moments associated with the model parameter estimates.

First of all, it is seen that the model is quite successful in capturing the overall characteristics of the data. However, the model tends to over estimate the mean and median of the Y and W distributions somewhat. The revenue per product parameter (Z) is central in this respect. It can be shown that irrespective of firm type, the mode of the product size distribution is equal to 1 in the model and consequently the mode of the Y distribution is equal to Z. As is seen in figure 4, the model estimation fails to perfectly match the mode in the observed Y distribution. The model

Table 3: Data Moments and Model Fit

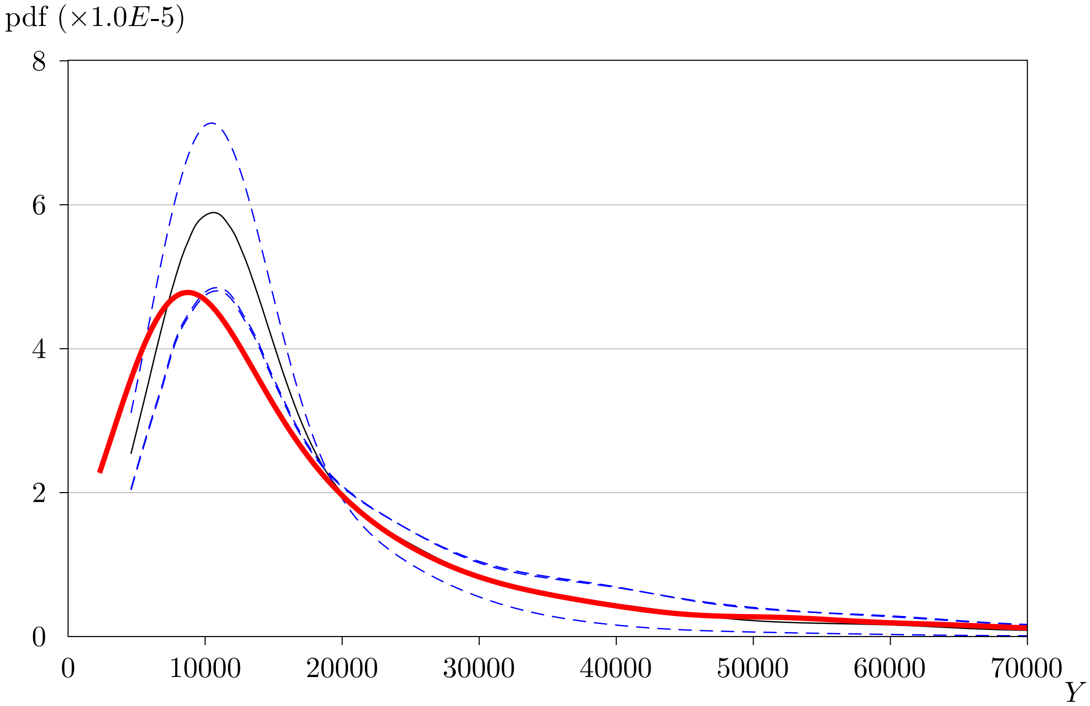
		1992	1993	1994	1995	1996	1997
Survivors	Data	$5,254$	$4,861$	$4,405$	$3,963$	$3,599$	$3,167$
	Sim	$5,254$	$4,714$	$4,253.97$	$3,859.39$	$3,512$	$3,209$
$E[Y]$	Data	$19,395.84$	$17,955.18$	$17,720.47$	$16,797.64$	$15,531.19$	$14,444.30$
	Sim	$19,569.88$	$18,397.22$	$17,270.21$	$16,477.96$	$15,585.87$	$14,740.87$
$Med[Y]_{Y>0}$	Data	$12,092.39$	$12,109.85$	$12,847.01$	$13,606.83$	$14,048.41$	$14,731.58$
	Sim	$11,873.31$	$12,277.76$	$12,741.21$	$13,413.68$	$14,103.26$	$14,845.71$
$E[W]$	Data	$9,962.15$	$9,156.43$	$8,615.13$	$8,351.38$	$7,804.45$	$7,149.29$
	Sim	$9,905.29$	$9,303.62$	$8,744.49$	$8,231.65$	$7,741.54$	$7,282.16$
$Med[W]_{W>0}$	Data	$6,546.79$	$6,539.92$	$6,717.87$	$7,240.28$	$7,517.31$	$7,827.89$
	Sim	$6,550.73$	$6,694.84$	$6,858.94$	$7,063.21$	$7,337.77$	$7,782.63$
$Std[Y]$	Data	$20,319.85$	$21,617.13$	$24,998.44$	$26,341.06$	$24,820.29$	$25,210.02$
	Sim	$20,065.89$	$20,693.82$	$21,215.06$	$21,800.13$	$22,107.39$	$22,307.12$
$Std[W]$	Data	$10,328.97$	$10,389.14$	$10,332.56$	$11,420.99$	$11,472.95$	$11,190.37$
	Sim	$8,861.19$	$9,290.46$	$9,631.06$	$9,857.90$	$10,022.83$	$10,124.32$
$Cor[Y,W]$	Data	0.89	0.88	0.82	0.83	0.91	0.91
	Sim	0.94	0.94	0.95	0.95	0.95	0.96
$Cor\left[\frac{Y}{N^*}, Y\right]$	Data	0.24	0.25	0.44	0.27	0.24	0.22
	Sim	0.30	0.29	0.29	0.28	0.28	0.27
$Cor\left[\frac{Y}{N^*}, N^*\right]$	Data	-0.06	-0.03	-0.02	-0.02	-0.02	-0.02
	Sim	0.05	0.05	0.06	0.06	0.06	0.07

estimated distribution is shifted somewhat to the right. The reason for the right shift is found in the under estimation of the standard deviation of Y and W. Again, Z is an important determinant of this moment. The higher the value of Z, the greater the variance in Y and W. While Z is not the only determinant, it seems that the estimation has sacrificed some of the first moment and median fits to improve the fit to the second moments.

The firm type distribution also affects the second moment fit, though, and one might suspect that allowing more types in the distribution support could introduce more variance and consequently allow for a lower Z estimate to bring the model estimates a bit more in line with the observed first moments and medians. Thus, the current Z estimate is probably an upward biased estimate. Figure 4 compares the observed and estimated distribution of value added. The dashed lines depict the value added distribution associated one of the three possible firm types. The estimated distribution of value added is a mixture of the three single-type distributions. The higher the profit of a type, the greater the variance in the distribution.

Figure 5 shows the change in the distribution of value added from 1992 to 1997 in the sample. It is seen that the observed distribution shifts to the right and is more spread out. The model captures this change and explains it as a change in firm type composition over time. The low profit types create new products at a lower rate than high profit types and consequently will tend to reduce

Figure 4: Observed and Estimated Value Added pdf's.

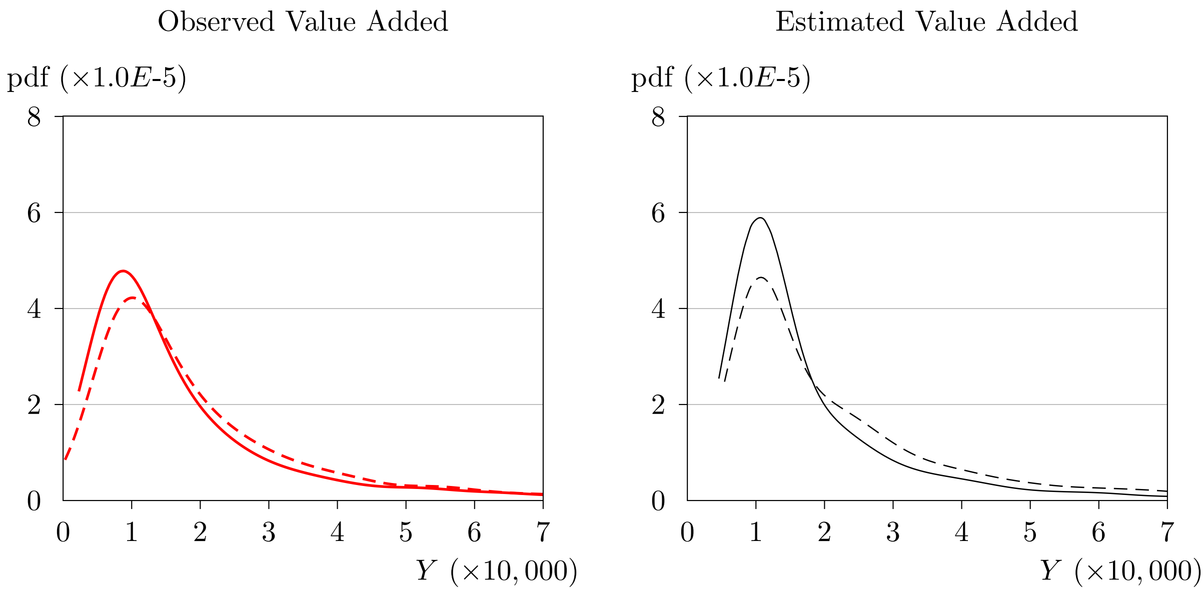


Note: Observed value added distribution drawn in bold pen. Estimated value added distribution drawn in solid thin pen. Single-type hypothetical value added distributions drawn in dashed pen.

in size at a greater rate than the high profit types. Therefore, the composition of firm types will switch towards high profit types over time and generate the increased spread in the distribution of value added. This particular source of variation in the data turns out to be an important identifier of δ. Experimentation with estimation of the model for fixed, lower values of δ results in less change in the survival conditional mean and median values of Y and W over time, thus forcing the model to over estimate the means and medians of Y and W early on in the sample and under estimate them in the later years. The lower value of δ implies a smaller difference in survival probabilities across firm types and consequently a slower rate of change in the firm type composition over time. Therefore, a lower value of δ results in a slower rate of change in the estimated conditional means and medians of Y and W over time.

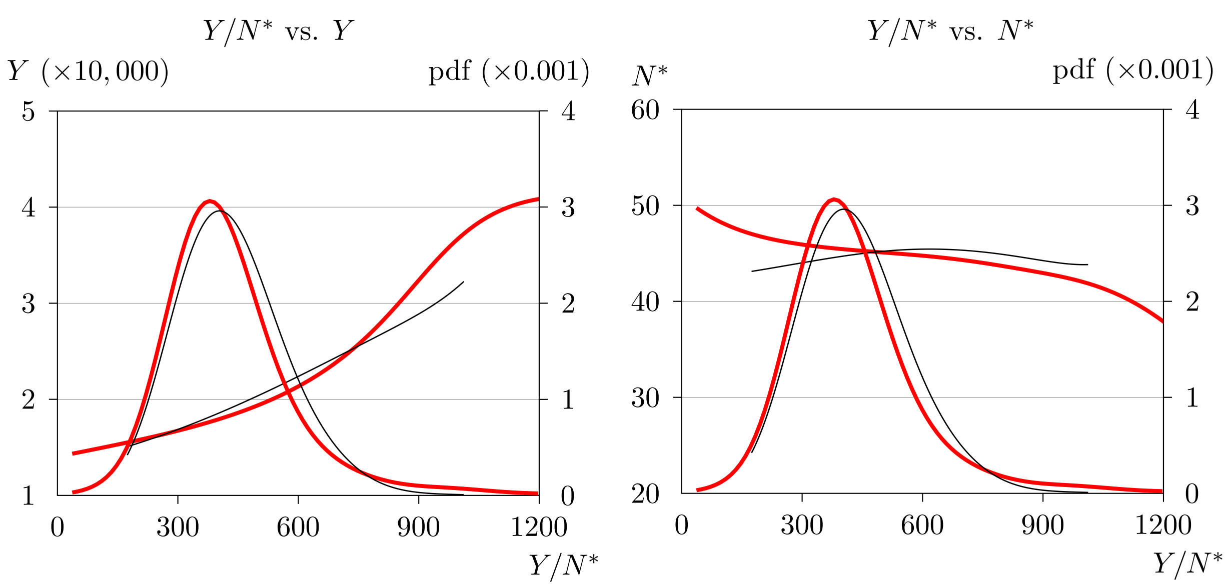
Figure 6 shows the model fit relative to the correlation between worker productivity and firm size (as measured by either value added or labor force size). It is seen that the model captures the relationships quite successfully. There is not quite enough noise in the model to generate as much spread in the worker productivity distribution support as is seen in data. Furthermore, the model

Figure 5: Change in Observed and Estimated Value Added Distribution from 1992 to 1997.



Note: Pdf in 1992 drawn in solid pen. Pdf of survivors in 1997 drawn in dashed pen.

Figure 6: Value Added per Standardized Worker versus Value Added and Labor Force Size.



Note: Observed relationships drawn in bold pen and estimated relationships drawn in thin pen.

estimates a somewhat flatter relationship between Y/N^* and Y than the observed relationship. The model also over estimates the wage share slightly resulting in a small over estimate of the labor force size.

6 Concluding Remarks

Large and persistent differences in firm productivity and size exist. Evidence suggests that the reallocation of workers across firms and establishments is an important source of economic growth. In the paper, we explore the Schumpeterian model of aggregate growth and firm evolution developed by Grossman and Helpman (1991) and by Klette and Kortum (2002).

We find that firms with higher measurable labor productivity will grow larger in the future and that worker reallocation from the less to more productive will contribute to growth only if current productivity predict future productivity in the model. Specifically, there is no relationship between current productivity and expected future firm sales and there will be no contribution to growth of worker reallocation across existing firms if profits are independently distributed over the sequence of new product innovations. Furthermore, one should find a negative relationship between employment size and current productivity measures in this case. However, if some firms consistently develop better products, then profit maximization will imply that the more productive will innovate more frequently and can expect to enjoy larger future sales as a consequence.

Existing studies that provide an empirical decomposition of aggregate productivity growth provide strong evidence for the importance of worker reallocation from exiting to entering firms and establishments. The evidence for the importance of reallocation across continuing firms is less clear. However, If gross output weights are used in constructing the productivity index as our model would require, then the two sources of growth are equally important and together explain over half of the productivity growth in the U.S. Manufacturing sector during the 19977-1987 period according to Foster, Haltiwanger, and Krizan (2001).

Our own evidence from Danish firm level data supports the conclusion that more productive firms grow faster. Specifically, the hypothesis that there is no relationship between size as measured by value added and labor productivity is clearly rejected in favor of a positive association between the two. Furthermore, a structural version of our model in which there are three types of firms that vary with respect to the quality of their products does an excellent job of explaining the moments of panel data observations on value added, employment size, and firm survival rates.

References

Altonji, J. G., and L. M. Segal (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 14, 353–366.

Alvarez, J., M. Browning, and M. Ejrnæs (2001): "Modelling Income Processes with Lots of Heterogeneity," *Working Paper*.

Baily, M., C. Hulton, and D. Campbell (1992): "Productivity Dynamics in Manufacturing Plants," *Brookings Papers on Economic Activity, Microeconomics*, pp. 187–249.

Bartelsman, E., and M. Doms (2000): "Understanding Productivity: Lessons for Longitudinal Microdata," *Journal of Economic Literature*, 38, 569–594.

Christensen, B. J., R. Lentz, D. T. Mortensen, G. Neumann, and A. Werwatz (2005): "On the Job Search and the Wage Distribution," *forthcoming in Journal of Labor Economics*.

Davis, S. J., J. C. Haltiwanger, and S. Schuh (1996): *Job Creation and Destruction*. MIT Press, Cambridge and London.

Dunne, T., L. Foster, J. Haltiwanger, and K. Troske (2002): "Wage and Productivity Dispersion in U.S. Manufacturing: The Role of Computer Investment," *Journal of Labor Economics*.

Fallick, B., and C. Fleischman (2001): "The Importance of Employer-to-Employer Flows in the U.S. Labor Market," *Federal Reserve Board Finance and Economics Discussion Paper*, 2001-18.

Foster, L., J. Haltiwanger, and C. J. Krizan (2001): "Aggregate Productivity Growth: Lessons from Microeconomic Evidence," in *New Developments in Productivity Analysis*, ed. by C. R. Hulten, E. R. Dean, and M. J. Harper. University of Chicago Press, Chicago.

Frederiksen, A., and N. Westergaard-Nielsen (2002): "Where Did They Go?," *Århus School of Business Working Paper*.

Gourieroux, C., A. Monfort, and E. Renault (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8(0), S85–S118.

Grossman, G. M., and E. Helpman (1991): *Innovation and Growth in the Global Economy*. MIT Press, Cambridge, Massachusetts and London, England.

Hall, G., and J. Rust (2003): "Simulated Minimum Distance Estimation of a Model of Optimal Commodity Price Speculation with Endogenously Sampled Prices," *Working Paper*.

Horowitz, J. L. (1998): "Bootstrap Methods for Covariance Structures," *Journal of Human Resources*, 33, 39–61.

Klette, T. J., and S. Kortum (2002): "Innovating Firms and Aggregate Innovation," *NBER Working Paper 8819*.

Mortensen, D. T. (2003): *Wage Dispersion: Why Are Similar Workers Paid Differently?* MIT Press, Cambridge, Massachusetts and London, England.

Stewart (2002): "Recent Trends in Job Stability and Job Security," *U.S. BLS Working Paper*, 356.