An Empirical Model of Growth Through Product Innovation *

Rasmus Lentz
University of Wisconsin-Madison
and CAM

Dale T. Mortensen
Northwestern University
NBER and IZA

July 2005
Revised: March 9, 2007

Abstract

Productivity differences across firms are large and persistent but the evidence for worker reallocation as an important source of aggregate productivity growth is mixed. The purpose of the paper is to estimate the structure of an equilibrium model of growth through innovation designed to identify and quantify the role of resource reallocation in the growth process. The model is a version of the Schumpeterian theory of firm evolution and growth developed by Klette and Kortum (2004) extended to allow for firm heterogeneity. The data set is a panel of Danish firms that includes information on value added, employment, and wages. The model's fit is good. We show that the empirical growth decomposition identity popularized by Baily, Hulten, and Campbell (1992) does not identify the contribution of resource reallocation to growth in a steady state stochastic growth model such as ours. However, the theory implies that more productive firms in each cohort grow faster and consequently crowd out less productive firms in steady state. This selection effect accounts for 55% of aggregate growth in the estimated version of the model.

Keywords: Labor productivity growth, worker reallocation, firm dynamics, firm panel data estimation.

JEL Classification Numbers: E22, E24, J23, J24, L11, L25, O3, O4.

*This research was supported by a collaborative grant to the authors from the NSF. Funding for data access was provided by the Danish Social Science Research Council through a grant to Bent Jesper Christensen as part of other collaborative research. Rasmus Lentz acknowledges research support from Centre for Applied Microeconometrics at University of Copenhagen. The authors wish to thank Victor Aguirregabiria, Joseph Altonji, Jonathan Eaton, Robert Hall, John Kennan, Samuel Kortum, Giuseppe Moscarini, Jean-Marc Robin, Rob Shimer, and three anonymous referees for useful comments and suggestions. All remaining errors are ours.

1 Introduction

In their review article of empirical productivity studies based on longitudinal plant and firm data, Bartelsman and Doms (2000) conclude that the extent of dispersion in productivity across production units, firms or establishments, is large. Furthermore, the productivity rank of any unit in the distribution is highly persistent. Although the explanations for firm heterogeneity in productivity are not fully understood, economic principles dictate that its presence will induce the reallocation of resources from less to more profitable firms.

To quantify the effect of worker reallocation on growth, decompositions of productivity growth into terms associated with productivity growth within firms and between firms have been proposed and implemented.[1] Studies based on these decomposition identities provide mixed evidence of the importance of reallocation as a source of aggregate productivity growth. Based on data on U.S. manufacturing firms, Bartelsman and Doms (2000) find that roughly one quarter of growth can be attributed to gross reallocation, another quarter to net entry, and roughly half of all growth to be a result of within firm growth. However, based on the same data Foster, Haltiwanger, and Krizan (2001) find that "...much of the increase in labor productivity would have occurred even if labor share had been held constant at their initial levels."[2] In a study of a number of different OECD countries, Scarpetta, Hemmings, Tressel, and Woo (2002) also find that the majority of growth can be attributed to within firm growth. In this paper we argue that the between firm component does not capture the role of reallocation in the growth process. Indeed, in a broad class of models including that studied in this paper, the term is zero in the absence of transitory shocks and measurement error.

Our model is an extension on that proposed by Klette and Kortum (2004), which itself builds on the endogenous growth model of Grossman and Helpman (1991). It is designed to capture the implications for growth through reallocation induced by the creative destruction process. In the model, final consumption output is produced by a competitive sector using a continuum of

[1] The literature on the connection between aggregate and micro productivity growth include: Foster, Haltiwanger, and Krizan (2001), Baily, Hulten, and Campbell (1992), Baily, Bartelsman, and Haltiwanger (1996), Bartelsman and Dhrymes (1994), Griliches and Regev (1995), and Olley and Pakes (1996),Tybout (1996), Aw, Chen, and Roberts (1997), and Liu and Tybout (1996).

[2] The discrepancy can be traced to variation in the particular choices of productivity and weighting measures.

differentiated intermediate products as inputs. More productive or higher quality versions of each intermediate product type are introduced from time to time as the outcome of R&D investment by both existing firms and new entrants. The supplier of the current version has monopoly power based on frontier knowledge and uses it to set price above the marginal cost of production. As new products and services displace old, the process of creative destruction induces the need to reallocate workers across activities. In the version of the model estimated here, firms differ with respect to the expected productivity of the intermediate goods and services that they create. The model has two principal empirical implications. First, a firm that is of a more innovative type in the sense that the quality improvement embodied in its products is higher, can charge a higher price, is more profitable, and as a consequence invests more in innovation and grows relatively faster after entry. Second, the expected firm growth conditional on firm type is independent of size.

In an earlier paper, Lentz and Mortensen (2005), we establish the existence of a general equilibrium solution to a simplified version of the model applied in this paper. In this paper, we use the equilibrium relationships and information on value added, employment, and wage payments drawn from a Danish panel of firms over the period 1992-1997 to estimate the model's parameters by the method of indirect inference. Providing a good fit to data, the model is estimated on a number of cross section and dynamic moments including size, productivity, and firm growth distribution moments. The model is also estimated to fit the growth decomposition pioneered by Foster, Haltiwanger, and Krizan (2001) found in our data.

In spite of the fact that that all growth arises because resources are reallocated from less to more rapidly growing firms in the model, the term typically interpreted as the contribution of gross reallocation is close to zero in our data. We show that this result is to be expected in a stochastic equilibrium model such as ours. Although in the model more profitable firms in each cohort grow faster on average as a consequence of more frequent innovation, the aggregate share of products supplied and inputs required by each firm type are constant in the model's ergodic steady state by definition. As a consequence, the "between" and "cross" terms in the Foster, Haltiwanger, and Krizan (2001) decomposition should be zero in the absence of measurement error and transitory shocks.

In our model, the aggregate growth rate in final good consumption is equal to the sum of the expected percentage increase in the productivity of the intermediate inputs weighted by their contributions to final consumption output. This term can be decomposed by type of firm into the net contribution of entrants and incumbents. As in the empirical decomposition literature, the net contribution of entry is the average increase in productivity of the entrants relative to those that exit the market within each period. The second term, that associated with continuing firms, can be decomposed into two parts designed to reveal the consequences of the selection process associated with differences in firm growth and survival rates. The first is the contribution of incumbents if the share of value added supplied by each firm type were to remain equal to that at entry and the second is the contribution of the difference between the steady state share and the share at entry. Because a more productive firm type grows faster, its share in steady state exceeds that at entry which implies that selection contributes positively to growth. Indeed, our estimated model implies that net entry accounts for 20% of the aggregate growth rate while 55% can be attributed to the selection effect.

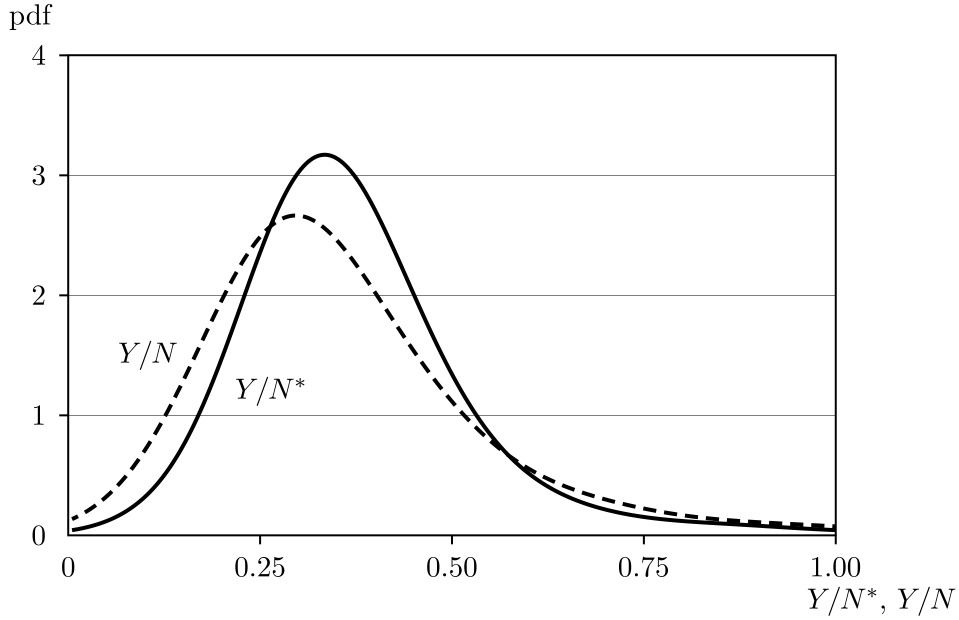
2 Danish Firm Data

Danish firm data provide information on productivity dispersion and the relationships among productivity, employment, and sales. 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 4,900 firms is restricted to those with 20 or more employees. The sample does not include entrants.[3] The variables observed in each year include value added (Y), the total wage bill (W), and full-time equivalent employment (N). In this paper we use these relationships to motivate the theoretical model studied. Both Y and W are measured in Danish Kroner (DKK) while N is a body count.

Non-parametric estimates of the distributions of two alternative empirical measures of a firm's labor productivity are illustrated in Figure 1. The first empirical measure of firm productivity is value added per worker (Y/N) while the second is valued added per unit of quality adjusted

[3]The full panel of roughly 6,700 firms contains some entry, but due to the sampling procedure, the entrant population suffers from significant selection bias. Rather than attempt to correct for the bias, we have chosen not to rely on the entrant population for identification of the model.

Figure 1: Observed firm productivity distribution, 1992



Note: Value added (Y) measured in 1 million DKK. N is the raw labor force size measure. N^* is the quality adjusted labor force size.

employment (Y/N^*). Standard labor productivity misrepresents cross firm productivity differences to the extent that labor 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 $N_j^* = W_j/w$ where

$$w = \frac{\sum_j W_j}{\sum_j N_j} \tag{1}$$

is the average wage paid per worker in the market.[4] 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 general shape.

Both distributions are consistent with those found in other data sets. For example, productivity distributions are significantly dispersed and skewed to the right. In the case of the adjusted measure

[4]In the case, where a firm is observed over several periods, the implicit identification of the firm's labor force quality is taken as an average over the time dimension to address issues of measurement error. The alternative approach of identifying a quality measure for each year has no significant impact on the moments of the data set.

Table 1: Productivity – Size Correlations

	Employment (N)	Adjusted Employment (N^*)	Value Added (Y)
Y/N	0.0017	0.0911	0.3138
Y/N^*	-0.0095	-0.0176	0.1981

of productivity, the 5^{th} percentile is roughly half the mode while the 95^{th} percentile is approximately twice as large are the mode. The range between the two represents a four fold difference in value added per worker across firms. These facts are similar to those reported by Bartelsman and Doms (2000) for the U.S.

There are many potential explanations for cross firm productivity differentials. A comparison of the two distributions represented in Figure 1 suggests that differences in the quality of labor inputs does not seem to be the essential one. The process of technology diffusion is a well documented. Total factor productivity differences across firms can be expected as a consequence of slow diffusion of new techniques. If technical improvements are either factor neutral or capital augmenting, then one would expect that more productive firms would acquire more labor and capital. The implied consequence would seem to be a positive relationship between labor force size and labor productivity. Interestingly, there is no correlation between the two in Danish data.

The correlations between the two measures of labor productivity with the two employment measures and sales as reflected in value added are reported in Table 1. As documented in the table, the correlation between labor force size and productivity using either the raw employment measure or the adjusted one is zero. However, note the strong positive associate between value added and both measures of labor productivity.

The theory developed in this paper is in part motivated by these observations. Specifically, it is a theory that postulates labor saving technical progress of a specific form. Hence, the apparent fact that more productive firms produce more with roughly the same labor input per unit of value added is consistent with the model.

3 An Equilibrium Model of Creative Destruction

As is well known, firms come is an amazing range of shapes and sizes. This fact cannot be ignored in any analysis of the relationship between firm size and productivity. Furthermore, an adequate theory must account for entry, exit and firm evolution in order to explain the size distributions observed. Klette and Kortum (2004) construct a stochastic model of firm product innovation and growth that is consistent with stylized facts regarding the firm size evolution and distribution. The model also has the property that technical progress is labor saving. For these reasons, we pursue their approach in this paper.

Although Klette and Kortum (2004) allow for productive heterogeneity, firm productivity and growth are unrelated because costs and benefits of growth are both proportional to firm productivity in their model. Allowing for a positive relationship between firm growth and productivity is necessary for consistency with the relationships found in the Danish firm data studied in this paper.

3.1 Preferences and Technology

The model is set in continuous time. Intertemporal utility of the representative household at time t is given by

$$U_t = \int_t^{\infty} \ln C_s e^{-r(s-t)} ds \tag{2}$$

where $\ln C_t$ denotes the instantaneous utility of the single consumption good at date t and r represents the pure rate of time discount. Each household is free to borrow or lend at interest rate r_t. Nominal household expenditure at date t is $E_t = P_t C_t$. Optimal consumption expenditure must solve the differential equation $\dot{E}/E = r_t - r$. Following Grossman and Helpman (1991), we choose the numeraire so that $E_t = z$ for all t without loss of generality, which implies $r_t = r$ for all t. Note that this choice of the numeraire also implies that the price of the consumption good, P_t, falls over time at a rate equal to the rate of growth in consumption.

The consumption good is supplied by many competitive providers and the aggregate quantity produced is determined by the quantity and productivity of the economy's intermediate inputs. Specifically, there is a measure 1 continuum of different inputs and consumption is determined by the CES production function

$$C_t = \left[\int_{j=0}^{1} Z(j) \big(A_t(j) x_t(j) \big)^{\frac{\sigma}{\sigma-1}} dj \right]^{\frac{\sigma-1}{\sigma}} , \ \ \sigma \geq 0 \tag{3}$$

where $x_t(j)$ is the quantity of input j at time t and $A_t(j)$ is the productivity of input j at time t. $Z(j)$ reflects that expenditure shares vary across the intermediary inputs. The level of productivity of each input is determined by the number of technical improvements made in the past. Specifically,

$$A_t(j) = \prod_{i=1}^{J_t(j)} q_i(j), \tag{4}$$

where $J_t(j)$ is the number of innovations made in input j up to date t and $q_i(j) > 1$ denotes the quantitative improvement (step size) in the input's productivity attributable to the i^{th} innovation in product j. Denote by $q(j)$ the latest quality improvement of good j. Innovations arrive at rate δ which is endogenous but the same for all intermediate products under the assumption that innovation is equally likely across the set of intermediate goods.

If the latest innovation in good j happened at time t, the profit maximizing demand for intermediate good j at time $t+a$ can be expressed as,

$$x_{t+a}(j) = \frac{z_{t+a}(j)}{p_{t+a}(j)}, \tag{5}$$

where

$$z_{t+a}(j) = z_t(j) e^{g(1-\sigma)a}, \text{ and } z_t(j) = z Z_t(j)^{\sigma} \left(\frac{P_t(j)}{P_t} \right)^{1-\sigma} . \tag{6}$$

$P_t(j) = p_t(j)/A_t(j)$ is the per quality unit price of good j, and the price of the final consumption good is given by,

$$P_t = \left(\int_{j=0}^{1} P_t(j)^{1-\sigma} Z(j)^{\sigma} dj \right)^{\frac{1}{1-\sigma}} .$$

In between quality improvement events, the demand for good j is time dependent because its quality remains fixed while the quality of the other intermediate goods is on average growing at rate g. For a given price, if the intermediate goods are gross substitutes ($\sigma > 1$), consumers will decrease demand for good j as the quality of the alternatives is increasing. If goods are gross compliments the demand will increase over time. Only in the unit elastic case is demand stationary.

3.2 The Value of a Firm

Each individual firm is the monopoly supplier of the products it has created in the past that have survived to the present. The price charged for each is the minimum of the monopoly price and the limit price resulting from competition with suppliers of previous versions of the good. In Nash-Bertrand equilibrium, the limit price exactly prices out all suppliers of previous version of the good. Consumers are exactly indifferent between the higher quality intermediate good supplied by the quality leader at the limit price and the highest quality alternative priced at marginal cost. The limit price is the product of the magnitude of the latest quality improvement and the marginal cost of production. If the monopoly price is below the limit price, the firm will simply charge the monopoly price which also prices out the previous suppliers.

The output of any intermediate good requires labor and capital input in fixed proportions. Total factor productivity is the same across all goods and is set equal to unity without loss of generality. Denote by w the wage of a unit of labor, and by κ the cost of capital. The price charged for good j can be expressed as,

$$p(j) = m\big(q(j), \sigma\big)(w + \kappa), \text{ where } m\left(q, \sigma\right) = \begin{cases} \frac{\sigma}{\sigma-1} & \text{if } q > \frac{\sigma}{\sigma-1} \text{ and } \sigma > 1 \\ q & \text{otherwise.} \end{cases} \tag{7}$$

The gross profits associated with supplying good j at time $t + a$ is,

$$\begin{aligned} \Pi_{t+a}(j) &= \big(p(j) - w - \kappa\big)\frac{z_{t+a}(j)}{p(j)} \\ &= \pi\big(q(j), \sigma\big) z_{t+a}(j), \end{aligned} \tag{8}$$

where $\pi(q, \sigma) = 1 - m(q, \sigma)^{-1}$.

Following Klette and Kortum (2004), 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. A firm enters with one product and a firm exit when it no longer has leading edge products. In Klette and Kortum's 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 cost is $wc(\gamma)k$ where $c(\gamma)k$ represents the labor input required in the 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 in 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, the cost structure implies that Gibrat's law holds in the sense that innovation rates are size independent contingent on type.

The market for any current product supplied by a 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. The firm chooses the creation rate γ to maximize the expected present value of its future net profit flow.

At entry the firm instantly learns its type, τ, which is a realization of the random variable, $\tilde{\tau} \sim \phi(\cdot)$. When an innovation occurs, the productivity improvement realization is drawn from a type conditional distribution. Specifically, a τ-type's improvement realizations are represented by the random variable, $\tilde{q}_\tau$, that is distributed according to the cumulative distribution function, $F_\tau(\cdot)$. It is assumed that a higher firm type draws realizations from a distribution that stochastically dominates that of lower firm types, that is if $\tau' > \tau$ then $F_{\tau'}(\tilde{q}) \leq F_\tau(\tilde{q})$ for all $\tilde{q} \geq 1$.[5] Assume that the lower bound of the support of $\tilde{q}_\tau$ is 1 for all τ.

By assumption firms cannot direct their innovation activity toward a particular market. Furthermore, their ability to create new products is not specific to any one or subset of product types.[6] Since product demand $Z(j)$ and quality levels vary across products, firms face demand uncertainty for a new innovation resolved only when the product type of an innovation is realized. By equation (6) the initial product line demand realization $z_t(j) = zZ_t(j)^\sigma \left(\frac{P_t(j)}{P_t}\right)^{1-\sigma}$ is a result of a random draw over the product space $j \in [0,1]$. Denote by $G(\cdot)$ the steady state cumulative distribution function of $z_t(j)$ across products. The initial demand for an innovation is then determined as a realization of the random variable $\tilde{z} \sim G(\cdot)$. By definition of $z_t(j)$ it follows that $z = E[\tilde{z}]$. $\tilde{z}$ and

[5]The "noise" in the realization of quality step size suggests the need for a new entrant to learn about its type in response to the actual realizations of q. We abstract from this form of learning. Simulation experiments using the parameter estimates obtained under this assumption suggest that learning ones type is not an important feature of the model's equilibrium solution.

[6]On its face, this feature of the model is not realistic in the sense that most firms innovate in a limited number of industries. However, if there are a large number of product variants supplied by each industry, then it is less objectionable. In the appendix we show that similar results are obtained when estimating the model within broadly defined industries.

$\tilde{q}_\tau$ are independent.

A firm's state is characterized by the number of products it currently markets, k, and the particular productivity improvement and demand realization for each products as represented by the vectors, $\tilde{q}^k = \{\tilde{q}_1, \ldots, \tilde{q}_k\}$ and $\tilde{z}^k = \{\tilde{z}_1, \ldots, \tilde{z}_k\}$. Because the demand for any surviving product changes deterministically, the current demand level is a sufficient statistic. Given such a state, the value of a type τ firm is accordingly given by,

$$\begin{aligned} rV_\tau\left(\tilde{q}^k, \tilde{z}^k, k\right) &= \max_{\gamma \geq 0} \Bigg\{ \sum_{i=1}^{k} \tilde{z}_i \pi\left(\tilde{q}_i\right) - kwc\left(\gamma\right) + k\gamma \left[E_\tau \left[V_\tau\left(\tilde{q}^{k+1}, \tilde{z}^{k+1}, k+1\right)\right] - V_\tau\left(\tilde{q}^k, \tilde{z}^k, k\right)\right] \\ &\quad + k\delta \left[\frac{1}{k} \sum_{i=1}^{k} V_\tau\left(\tilde{q}^{k-1}_{\langle i \rangle}, \tilde{z}^{k-1}_{\langle i \rangle}, k-1\right) - V_\tau\left(\tilde{q}^k, \tilde{z}^k, k\right)\right] + \dot{V}_\tau\left(\tilde{q}^k, \tilde{z}^k, k\right) \Bigg\}, \end{aligned} \tag{9}$$

where $\left(\tilde{q}^{k-1}_{\langle i \rangle}, \tilde{z}^{k-1}_{\langle i \rangle}\right)$ refers to $\left(\tilde{q}^k, \tilde{z}\right)$ without the i^{th} elements. The first term on the right side is current gross profit flow accruing to the firms product portfolio less current expenditure on R&D. The second term is the expected capital gain associated with the arrival of a new product line. The third term represents the expected capital loss associated with the possibility that one among the existing product lines (chosen at random) will be destroyed. Finally, the last term is the change in value over time as a result of demand time dependence of existing products.

As one can verify by substitution, the unique solution to (9) is given by,

$$V_\tau\left(\tilde{q}^k, \tilde{z}^k, k\right) = \sum_{i=1}^{k} \frac{\tilde{z}_i \pi\left(\tilde{q}_i, \sigma\right)}{r + \delta - g(1-\sigma)} + kz\Psi_\tau, \tag{10}$$

where,

$$\begin{aligned} \Psi_\tau &= \max_{\gamma \geq 0} \frac{\gamma \nu_\tau - w\hat{c}\left(\gamma\right)}{r + \delta} \\ \nu_\tau &= \frac{\bar{\pi}_\tau(\sigma)}{r + \delta - g(1-\sigma)} + \Psi_\tau, \end{aligned}$$

where $\bar{\pi}_\tau(\sigma) = 1 - E\left[m(\tilde{q}_\tau, \sigma)^{-1}\right]$ and $\hat{c}\left(\gamma\right) \equiv c\left(\gamma\right)/z$. Ψ_τ is the type conditional innovation option value embodied in each product. ν_τ is the type conditional expected value of a product. It is the sum of the innovation option value and the discounted stream of expected profits where the effective discount rate is the sum of the the interest rate, the product destruction rate, and the rate of decline in the future demand for the product

It then follows directly from (9) that the firm's optimal choice of creation rate, γ_τ, satisfies,

$$w\hat{c}'\left(\gamma_\tau\right) = \nu_\tau, \tag{11}$$

where ν_τ is the type conditional expected value of an additional product line.

Equation (11) implies that the type contingent creation rate is size independent - a theoretical version of Gibrat's law. Also, the second order condition, $c''(\gamma) > 0$, and the fact that the marginal value of a product line is increasing in $\bar{\pi}_\tau$ imply that a firm's creation rate increases with profitability. Therefore, we obtain that $\gamma_{\tau'} \geq \gamma_\tau$ for $\tau' \geq \tau$. These results are the principal empirical implications of the model.

3.3 Firm Entry

The entry of a new firm requires innovation. Suppose that there are a constant measure m of potential entrants. The rate at which any one of them generates a new product is γ_0 and the total cost is $wc(\gamma_0)$ where the cost function is the same as that faced by an incumbent. The firm's type is unknown ex ante but is realized immediately after entry. Since the expected return to innovation is $E[\nu_\tau]$ and the aggregate entry rate is $\eta = m\gamma_0$, the entry rate satisfies the following free entry condition

$$w\hat{c}'\left(\frac{\eta}{m}\right) = \sum_\tau \nu_\tau \phi_\tau, \tag{12}$$

where ϕ_τ is the probability of being a type τ firm at entry. Of course, the second equality follows from equation (11).

3.4 The Steady State Distribution of Firm Size

A type τ firm's size is reflected in the number of product lines supplied which evolves as a birth-death process. 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 type τ firms with $k > 1$ products requires

$$\gamma_\tau(k-1)M_\tau(k-1) + \delta(k+1)M_\tau(k+1) = (\gamma_\tau + \delta)kM_\tau(k)$$

for every τ where $M_\tau(k)$ is the steady state mass of firms of type τ that supply k products. Because an incumbent dies when its last product is destroyed by assumption but entrants flow into the set of firms with a single product at rate η,

$$\phi_\tau \eta + 2\delta M_\tau(2) = (\gamma_\tau + \delta) M_\tau(1)$$

where ϕ_τ is the fraction of the new entrants of type τ. Births must equal deaths in steady state and only firms with one product are subject to death risk. Therefore, $\phi_\tau \eta = \delta M_\tau(1)$ and

$$M_\tau(k) = \frac{k-1}{k}\frac{\gamma_\tau}{\delta} M_\tau(k-1) = \frac{\eta \phi_\tau}{\delta k}\left(\frac{\gamma_\tau}{\delta}\right)^{k-1} \tag{13}$$

by induction.

The size distribution of firms conditional on type can be derived using equation (13). Specifically, the total firm mass of type τ is

$$\begin{aligned} M_\tau &= \sum_{k=1}^{\infty} M_\tau(k) = \frac{\phi_\tau \eta}{\delta} \sum_{k=1}^{\infty} \frac{1}{k}\left(\frac{\gamma_\tau}{\delta}\right)^{k-1} \\ &= \frac{\eta}{\delta} \ln\left(\frac{\delta}{\delta - \gamma_\tau}\right) \frac{\delta \phi_\tau}{\gamma_\tau}. \end{aligned} \tag{14}$$

where convergence requires that the aggregate rate of creative destruction exceed the creation rate of every incumbent type, i.e., $\delta > \gamma_\tau \ \forall \tau$. Hence, the fraction of type τ firm with k product is

$$\frac{M_\tau(k)}{M_\tau} = \frac{\frac{1}{k}\left(\frac{\gamma_\tau}{\delta}\right)^k}{\ln\left(\frac{\delta}{\delta-\gamma_\tau}\right)}. \tag{15}$$

Equation (15) is the steady state distribution of $\tilde{k}_\tau$. This is the logarithmic distribution with parameter γ_τ/δ.[7] Consistent with the observations on firm size distributions, the one implied by the model is highly skewed to the right.

By equation (15), the mean of the type conditional firm size distribution is,

$$E\left[\tilde{k}_\tau\right] = \sum_{k=1}^{\infty} \frac{k M_\tau(k)}{M_\tau} = \frac{\frac{\gamma_\tau}{\delta-\gamma_\tau}}{\ln\left(\frac{\delta}{\delta-\gamma_\tau}\right)}, \tag{16}$$

It follows that the total mass of products produced by type τ firms, K_τ, is

$$K_\tau = \sum_{k=1}^{\infty} k M_\tau(k) = \frac{\eta \phi_\tau}{\delta - \gamma_\tau}. \tag{17}$$

[7] This result is in Klette and Kortum (2004). We include the derivation here simply for completeness.

As the product creation rate increases with expected profitability, expected size does also. Formally, because $(1+a)\ln(1+a) > a > 0$ for all positive values of a, the expected number of products is increasing in expected firm profitability,

$$\frac{\partial E\left[\tilde{k}_\tau\right]}{\partial \gamma_\tau} = \left(\frac{(1+a_\tau)\ln(1+a_\tau) - a_\tau}{(1+a_\tau)\ln^2(1+a_\tau)}\right)\frac{1+a_\tau}{\delta-\gamma_\tau} > 0 \tag{18}$$

where $a_\tau = \frac{\gamma_\tau}{\delta-\gamma_\tau}$.

Finally by the condition that the total product mass $\sum_\tau K_\tau$ is constant and normalized at unity, the rate of creative-destruction is the sum of the entry rate and the aggregate creation rates of all the incumbents,

$$\delta = \eta + \sum_\tau K_\tau \gamma_\tau. \tag{19}$$

3.5 Labor Demand and Market Clearing

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.

By the normalization of labor productivity at unity, it follows from equations (5) and (6) that the production labor demand for an age a product line with initial demand realization $\tilde{z}$ and quality realization $\tilde{q}$ is given by

$$\ell^p(\tilde{z},\tilde{q},a) = \frac{\tilde{z}e^{g(1-\sigma)a}}{m(\tilde{q},\sigma)(w+\kappa)}. \tag{20}$$

Denote by ℓ^p_τ the average labor demand per product of a type τ firm for the purpose of production. Product age is going to be exponentially distributed with parameter δ. It follows that,

$$\begin{aligned}\ell^p_\tau &= \int_1^\infty \int_0^\infty \int_0^\infty m(q,\sigma)^{-1}(w+\kappa)^{-1} z' e^{g(1-\sigma)a}\delta e^{-\delta a} da dG(z') dF_\tau(q) \\ &= \frac{\delta z\big(1-\bar{\pi}_\tau(\sigma)\big)}{(w+\kappa)[\delta - g(1-\sigma)]}.\end{aligned} \tag{21}$$

Hence, the average type conditional labor demand per product is,

$$\ell_\tau = \ell^p_\tau + c(\gamma_\tau). \tag{22}$$

The expected type conditional per product labor demand, ℓ_τ is decreasing in $\bar{\pi}_\tau(\sigma)$ if the reduction in labor demand for production, ℓ^p_τ, dominates the increase in labor demand for innovation,

$c(\gamma_\tau)$. So, although more profitable firms on average supply more products, total type conditional expected employment, $\ell_\tau E[\tilde{k}_\tau]$, need not increase with $\bar{\pi}_\tau(\sigma)$, in general. Hence, the hypothesis that firms with the ability to create greater productivity improvements grow faster is consistent with dispersion in labor productivity and the correlations between value added, labor force size, and labor productivity observed in Danish data reported above.

Labor market clearing requires that the equilibrium wage solves

$$\ell = \sum_\tau K_\tau \ell_\tau + mc(\eta/m). \tag{23}$$

3.6 Growth Rate

Given stationarity of z/w, log-differentiation of equation (3) and applying the law of large numbers yield,

$$\frac{\dot{C}_t}{C_t} = \frac{\int_0^1 Z_t(j)(A_t(j)x_t(j))^{\frac{\sigma-1}{\sigma}} \frac{\dot{A}_t(j)}{A_t(j)} dj}{\int_0^1 Z_t(j)(A_t(j)x_t(j))^{\frac{\sigma-1}{\sigma}} dj} = E\left[\frac{\dot{A}_t(j)}{A_t(j)}\right].$$

As the number of innovations $J_t(j)$, $j \in [0,1]$ are independently and identically distributed Poisson random variables with common expectation δt, and the $(\ln q)$s are iid across time and product lines, the law of large number implies that

$$E\left[\frac{\dot{A}_t(j)}{A_t(j)}\right] = \delta E[\ln q].$$

Hence, the growth rate in consumption is,

$$g = \frac{\dot{C}}{C} = E\left[\frac{\dot{A}(j)}{A(j)}\right] = \delta E[\ln(q)] = \eta \sum_\tau \phi_\tau E[\ln(\tilde{q}_\tau)] + \sum_\tau K_\tau \gamma_\tau E[\ln(\tilde{q}_\tau)]. \tag{24}$$

3.7 Equilibrium

Definition 1 *A* steady state market equilibrium *is a triple (w, δ, g) together with optimally chosen entry rate $\eta = m\gamma_0$, and creation rate γ_τ, and a steady state size distribution K_τ for each type that satisfy equations (11), (12), (17), (19), (23), and (24).*

See Lentz and Mortensen (2005) for a proof of existence of a slightly simpler version of the model. In Appendix A, we describe in detail the steady state equilibrium solution algorithm used in the estimation procedure described below.

4 Estimation

If the ability to create higher quality products 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 products 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 to more profitable fast growing firms will be an important source of aggregate productivity growth because faster growing firms also contribute more to growth.

In this section, we demonstrate that firm specific differences in profitability are required to explain Danish 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.

4.1 Danish Firm Data

If more profitable firms grow faster in the sense that $\bar{\pi}_\tau > \bar{\pi}_{\tau'} \Rightarrow \gamma_\tau > \gamma_{\tau'}$, then (18) implies that fast growing firms also supply more products and sell more on average. However, because production employment per product decreases with productivity, total expected employment need not increase with $\bar{\pi}_\tau$ in general and decreases with $\bar{\pi}_\tau$ when growth is independent of a firm's past product productivity improvement realizations. These implications of the theory can be tested directly.

The model is estimated on an unbalanced panel of 4,872 firms drawn from the Danish firm panel described in Section 2. 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. In the estimation, the observed 1992 cross-section will be interpreted to reflect steady state whereas the following years generally do not reflect steady state since survival probabilities vary across firm types. Specifically, due to selection the observed cross-sections from 1993 to 1997 will have an increasing over-representation of high creation rate firm types relative to steady state. The ability to observe the gradual exit of the 1992 cross-section will be a useful source of identification. Entry in the original data set suffers from selection bias and while one can attempt to correct for the bias, we have made the choice to leave out entry altogether since it is not necessary for identification. By including the 1997 cross section in the set of moments, dynamic processes that change the cross sectional composition of survivors over time are reflected in the estimation.

The first two columns of Table 2 present a set of distribution moments with standard deviations in parenthesis. The standard deviations are obtained through bootstrapping on the original panel. Unless otherwise stated, amounts are in 1,000 real 1992 Danish Kroner where the Statistics Denmark consumer price index was used to deflate nominal amounts. It is seen that the size distributions are characterized by significant skew. The value added per worker distribution displays some skew and significant dispersion. All distributions display a right shift from 1992 to 1997. The distribution moments also include the positive correlation between firm productivity and output size and the slightly negative correlation between firm productivity and labor force size.

The last two columns of Table 2 contain the dynamic moments used in the estimation. First of all, note that empirical firm productivity displays significant persistence and some mean reversion. The dynamic moments also include the cross section distribution of growth rates that display significant dispersion. Furthermore, there is a slightly negative correlation between output size and growth rate in the data. The moments relating to firm growth rates ($\Delta Y/Y$) include firm death, specifically an exiting firm will contribute to the statistic with a -1 observation. Excluding firm deaths from the growth statistic results in a more negative correlation between firm size and growth due to the negative correlation between firm size and the firm exit hazard rate. Since the model also exhibits a negative correlation between the exit rate and size, the same will be true in the model simulations.

Finally, Table 2 also includes a standard empirical labor productivity growth decomposition. We use the preferred formulation in Foster, Haltiwanger, and Krizan (2001) which is taken from Baily, Bartelsman, and Haltiwanger (1996) and ultimately based on the Baily, Hulten, and Campbell

Table 2: Data moments (std dev in parenthesis)

	1992	1997		1992	1996
Survivors	4872.000	3628.000	$\text{Cor}[\frac{Y}{N^*}, \frac{Y_{+1}}{N^*_{+1}}]$	0.476	0.550
	–	(32.132)		(0.088)	(0.091)
$\text{E}[Y]$	26277.262	31860.850	$\text{Cor}[\frac{Y}{N^*}, \Delta\frac{Y}{N^*}]$	−0.227	−0.193
	(747.001)	(1031.252)		(0.103)	(0.057)
$\text{Med}[Y]$	13472.812	16448.965	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta Y}{Y}]$	−0.120	
	(211.851)	(329.417)		(0.016)	
$\text{Std}[Y]$	52793.105	64120.233	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta N^*}{N^*}]$	0.119	
	(5663.047)	(7741.448)		(0.032)	
$\text{E}[W]$	13294.479	15705.087	$\text{E}[\frac{\Delta Y}{Y}]$	−0.029	
	(457.466)	(609.595)		(0.008)	
$\text{Med}[W]$	7231.812	8671.939	$\text{Std}[\frac{\Delta Y}{Y}]$	0.550	
	(92.720)	(154.767)		(0.067)	
$\text{Std}[W]$	30613.801	35555.701	$\text{Cor}[\frac{\Delta Y}{Y}, Y]$	−0.061	
	(6750.399)	(8137.541)		(0.012)	
$\text{E}[\frac{Y}{N^*}]$	384.401	432.118	Within	1.015	
	(2.907)	(5.103)		(0.146)	
$\text{Med}[\frac{Y}{N^*}]$	348.148	375.739	Between	0.453	
	(1.829)	(2.139)		(0.112)	
$\text{Std}[\frac{Y}{N^*}]$	205.074	305.306	Cross	−0.551	
	(19.633)	(42.491)		(0.196)	
$\text{Cor}[Y, W]$	0.852	0.857	Exit	0.084	
	(0.035)	(0.045)		(0.066)	
$\text{Cor}[\frac{Y}{N^*}, N^*]$	−0.018	−0.026			
	(0.013)	(0.011)			
$\text{Cor}[\frac{Y}{N^*}, Y]$	0.198	0.143			
	(0.036)	(0.038)			

(1992) index (BHC).[8] The decomposition takes the form,[9]

$$\Delta P_t = \sum_{i \in C_t} s_{it-1} \Delta p_{it} + \sum_{i \in C_t} p_{it-1} \Delta s_{it} + \sum_{i \in C_t} \Delta p_{it} \Delta s_{it} + \sum_{i \in E_t} p_{it} s_{it} - \sum_{i \in X_t} p_{it-1} s_{it-1}, \qquad (25)$$

where $P_t = \sum_i s_{it} p_{it}$, $p_{it} = Y_{it}/N_{it}$, and $s_{it} = N_{it}/N_t$.

[8] Griliches and Regev (1995) present another variation on the Baily, Hulten, and Campbell (1992) decomposition. It performs much the same way as the Foster, Haltiwanger, and Krizan (2001) formulation.

[9] In the implementation of the decomposition we employ the version of (25) where the between term and the entry/exit terms are normalized by P_{t-1},

$$\Delta P_t = \sum_{i \in C_t} s_{it-1} \Delta p_{it} + \sum_{i \in C_t} \left(p_{it-1} - P_{t-1}\right) \Delta s_{it} + \sum_{i \in C_t} \Delta p_{it} \Delta s_{it} + \sum_{i \in E_t} \left(p_{it} - P_{t-1}\right) s_{it} - \sum_{i \in X_t} \left(p_{it-1} - P_{t-1}\right) s_{it-1}.$$

The identity in (25) decomposes time differences in value added per worker into 5 components in the order stated on the right hand side; within, between, a cross component, and entry and exit. As the names suggest, the BHC growth decomposition literature attaches particular significance to each term: The within component is interpreted as identifying growth in the productivity measure due to productivity improvements by incumbents. The between component is interpreted to capture productivity growth from reallocation of labor from less to more productive firms. The cross component captures a covariance between input shares and productivity growth and the last two terms capture the growth contribution of entrants and exits. The sum of the between and cross components is also sometimes referred to as gross reallocation.

We include the BHC growth decomposition in the set of data moments because it conveniently relates the estimation to the empirical growth literature. Furthermore, it does reflect a particular aspect of the dynamics in the data. As mentioned, the sample in this paper does not include entry, so there is no entry share in the decomposition. Consequently, the decomposition cannot be directly related to the results in Foster, Haltiwanger, and Krizan (2001), although a full decomposition is performed on the estimated model in section 5.

4.2 The BHC Growth Decomposition in Steady State

Despite the usual interpretations of the second and third terms, the BHC decomposition does not identify the contribution of reallocation to growth in a stochastic steady state model such as ours. Indeed, the second and third terms equal zero in any structural equilibrium model of the type studied in this paper for the following reason. In the model, all firms of the same type have the same productivity by the definition of type, and although individual firms can and do grow and contract over time, the steady state distribution of inputs over firm types is stationary by the definition of stationary stochastic equilibrium. Hence, if we let $j \in J$ represent an element of the set of firm types, let I_j denote the set of firms of type j, let s^*_{jt} represent the average share of employment per type j firm in period t, and let p^*_{jt} be the productivity of type j firms, then abstracting from entry and exit one can formulate the growth decomposition in terms of firm types as follows.[10]

$$\begin{aligned}\Delta P_t &= \sum_{j\in J}\sum_{i\in I_j} s_{it-1}\Delta p_{it} + \sum_{j\in J}\sum_{i\in I_j} p_{it-1}\Delta s_{it} + \sum_{j\in J}\sum_{i\in I_j}\Delta p_{it}\Delta s_{it} \\ &= \sum_{j\in J}|I_j|\, s^*_{jt-1}\Delta p^*_{jt} + \sum_{j\in J}|I_j|\, p^*_{jt-1}\Delta s^*_{jt} + \sum_{j\in J}|I_j|\,\Delta p^*_{jt}\Delta s^*_{jt} \\ &= \sum_{j\in J}|I_j|\, s^*_{j}\Delta p^*_{jt} \end{aligned} \tag{26}$$

where $|I_j|$ is the number of firms of type j and $s^*_{jt-1} = \frac{1}{|I_j|}\sum_{i\in I_j} s_{it-1}$. The first equality is implied by the fact that the set $\{I_1, I_2, ...I_j, ...\}$ is a partition of the set of all firms I, the second by the fact that the firms of the same type have the same productivity at any given date, and the last by the fact that the average share per firm of each type is constant (consequently, $\Delta s^*_{jt} = 0$ for all j and t) in a steady state equilibrium. The final expression in (26) is the first term in the Baily, Hulten, and Campbell (1992) index.

An interpretation of the sum of the between and cross components, $\sum_{i\in I}\Delta s_{it}p_{it}$ as the gross effect of reallocating resources across firms is incorrect because gains in employment share are exactly off set by losses in share across firms of the same type in steady state. In other words, workers are never exogenously reallocated across types in equilibrium as is implicit in the interpretation. As such, the decomposition cannot capture the steady state growth contribution from reallocation. The fact that many empirical studies based on the BHC decomposition have found little evidence of a significant contribution to growth from the gross reallocation component is not a surprise in light of the above argument.[11]

4.3 Model Estimator

An observation in the panel is given by $\psi_{it} = \{Y_{it}, W_{it}, N^*_{it}\}$, where Y_{it} is real value added, W_{it} the real wage sum, and N^*_{it} quality adjusted labor force size of firm i in year t. Let ψ_i be defined by, $\psi_i = \left\{\psi_{i1,...,}\psi_{iT}\right\}$ and finally, $\psi = \{\psi_1, \ldots, \psi_I\}$.

The model is estimated by indirect inference. The estimation procedure, as described in for example Gourieroux, Monfort, and Renault (1993), Hall and Rust (2003), and Alvarez, Browning,

[10] The general argument that includes entry and exit is presented in the appendix.

[11] Petrin and Levinsohn (2005) also reach the conclusion that the empirical measure $\sum_{i\in I}\Delta s_{it}p_{it}$ has no meaning of interest. Specifically, they argue the traditional "Solow residual" adapted to allow for market imperfections, which is the first component of the BHC index, is the correct measure for welfare comparisons. Their argument is valid for our structural model.

and Ejrnæs (2001), is as follows: First, define a vector of auxiliary data parameters, $\Gamma(\psi)$. The vector consists of all the items in Tables 2 except the number of survivors in 1992 and one of the growth decomposition components. Thus, $\Gamma(\psi)$ has length 37.

Next, produce a simulated panel $\psi^s(\omega)$ 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.[12]

The simulated auxiliary parameters are then given by,

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

where S is the number of simulation repetitions.[13]

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

$$\hat{\omega} = \arg\min_{\omega\in\Omega}\big(\Gamma^s(\omega) - \Gamma(\psi)\big)' A^{-1}\big(\Gamma^s(\omega) - \Gamma(\psi)\big), \tag{27}$$

where A is the variance-covariance matrix of the data moments $\Gamma(\psi)$. Following Horowitz (1998) it is estimated by bootstrap.

The variance of the estimator is estimated by bootstrap. In each bootstrap repetition, a new set of data auxiliary parameters $\Gamma(\psi^b)$ is produced, where ψ^b is the bootstrap data in the b^{th} bootstrap repetition. ψ^b is found by randomly selecting observations ψ_i from the original data with replacement. Thus, the sampling is random across firms but is done by block over the time dimension (if a particular firm i is selected, the entire time series for this firm is included in the sample). For the b^{th} repetition, an estimator ω^b, is found by minimizing the weighted distance between the re-centered bootstrap data auxiliary parameters $[\Gamma(\psi^b) - \Gamma(\psi)]$ and the re-centered simulated auxiliary parameters $[\Gamma^s(\omega^b) - \Gamma^s(\hat{\omega})]$,

$$\omega^b = \arg\min_{\omega\in\Omega}\Big([\Gamma^s(\omega) - \Gamma^s(\hat{\omega})] - [\Gamma(\psi^b) - \Gamma(\psi)]\Big)' A^{-1}\Big([\Gamma^s(\omega) - \Gamma^s(\hat{\omega})] - [\Gamma(\psi^b) - \Gamma(\psi)]\Big).$$

In each bootstrap repetition, a different seed is used to generate random numbers for the determination of $\Gamma^s(\omega)$. Hence, $V(\hat{\omega})$ captures both data variation and variation from the model

[12] Alternatively, one can initialize the simulation according to the observed data in the first year. This approach has the complication that a firm's number of products is not directly observed.

[13] The model estimate in the following section uses $S = 1000$.

simulation.[14]

4.4 Model Specification and Simulation

Given a set of parameter values, the model is used to generate time paths for value added (Y), the wage sum (W), and labor force size (N) for each simulated firm. The firm type distribution is specified as a 3-point discrete type distribution ϕ_τ. The type conditional productivity realization distributions are three parameter Weibull distributions that share a common shape parameter β_q and a unity point of origin. Each distribution is distinguished by its own scale parameter ξ_τ. Thus, the three productivity realization distributions are specified with 4 parameters. The demand realization distribution $G(\cdot)$ is a three parameter Weibull where o_Z is the origin, β_Z is the shape parameter, and ξ_Z is the scale parameter. The cost function is parameterized by $c(\gamma) = c_0\gamma^{(1+c_1)}$.

A type τ firm with k products characterized by $\tilde{q}^k$ and $\tilde{z}^k$ has value added,

$$Y_\tau\left(\tilde{q}^k, \tilde{z}^k\right) = \sum_{i=1}^{k} \tilde{z}_i, \tag{28}$$

and by equation (20) a wage bill of,

$$W_\tau\left(\tilde{q}^k, \tilde{z}^k\right) = \frac{w}{w+\kappa}\sum_{i=1}^{k}\frac{\tilde{z}_i}{m(\tilde{q}_i)} + wkc\left(\gamma_\tau\right). \tag{29}$$

Equations (28) and (29) provide the foundation for the model simulation.

In Appendix A, we describe the detailed procedure of how to find the steady state equilibrium for given model fundamentals. The initial characteristics of each firm are drawn from the model's steady state distributions. The steady state firm type probability distribution is

$$\phi_\tau^* = \frac{\eta\phi_\tau \ln\left(\frac{\delta}{\delta-\gamma_\tau}\right)}{M\gamma_\tau}, \tau = 1, 2, ..., N$$

where $M = \sum_\tau M_\tau$ is the total steady state mass of firms. A firm's type is drawn according to ϕ^*. Once a firm's type has been determined, its 1992 product line size is drawn from the type conditional steady state distribution of $\tilde{k}_\tau$ characterized in equation (15). Then the age realization of each product is drawn from the exponential age distribution. The age realization is used to adjust the demand realization draw for each product from $G(\cdot)$ according to equation (6).

[14] Variance estimates are obtained using 500 bootstrap repetitions

The growth rate in quality is reflected in the aggregate price index. Thus, everything else equal Y_τ and W_τ grow at rate g.

Given an initial size for a firm, its future size evolves according to the stochastic birth-death process described earlier. The forward simulation is done by dividing each annual time period into a large number of discrete sub-intervals, n. By assumption, in each sub-interval each of the stochastic creation and destruction processes can have zero or one event arrivals. Hence, a type τ firm with k products will in a given sub-interval lose a product with probability $1-e^{-k\delta/n}$ and gain a product with probability $1-e^{-k\gamma_\tau/n}$. As $n\to\infty$, the procedure will perfectly represent the continuous time processes in the model. In the simulations below, the model has been simulated with $n=104$.

The estimation allows for measurement error in both value added and the wage bill. The measurement error is introduced as a simple log-additive process,

$$
\begin{aligned}
\ln \hat{Y}_\tau\left(\tilde{q}^k,\tilde{Z}^k\right) &= \ln Y_\tau\left(\tilde{q}^k,\tilde{Z}^k\right)+\xi_Y \\
\ln \hat{W}_\tau\left(\tilde{q}^k,\tilde{Z}^k\right) &= \ln W_\tau\left(\tilde{q}^k,\tilde{Z}^k\right)+\xi_W,
\end{aligned}
$$

where $\xi_Y \sim N\left(-\frac{1}{2}\sigma_Y^2,\sigma_Y^2\right)$ and $\xi_W \sim N\left(-\frac{1}{2}\sigma_W^2,\sigma_W^2\right)$. Given this specification, the expected value of the process with noise and without are equal. The estimation is performed on the quality adjusted labor force size. Consequently, the wage bill measurement error is assumed to carry through to the labor force size, $\hat{N}_\tau\left(\tilde{q}^k,\tilde{Z}^k\right)=\hat{W}_\tau\left(\tilde{q}^k,\tilde{Z}^k\right)/w$ since by construction, $N_i^* w = W_i$ for all firms in the data.

4.5 Identification

The interest rate is set at $r=.05$. The wage w is immediately identified as the average worker wage in the sample $w=190.24$. Excluding these two, the set of structural model parameters ω has 16 parameters; $\omega=(c_0,c_1,\kappa,\sigma,m,\beta_Z,\xi_Z,o_Z,\beta_q,\xi_1,\xi_2,\xi_3,\phi_1,\phi_2,\sigma_Y^2,\sigma_W^2)$, where $\phi_3=1-\phi_1-\phi_2$ in the case of three types. In the actual implementation of the estimation, m is replaced by η as a fundamental model parameter. Of course, η is endogenous to the equilibrium, but since m is a free parameter, m can always be set to make the η estimate consistent with steady state equilibrium. The set of data moments $\Gamma(\psi)$ has size 37.

The fact that the dimension of $\Gamma(\psi)$ exceeds that of ω does not in itself guarantee separate identification of the elements in ω. To understand the identification of the model it is useful to consider a stripped down version without stochastic demand and product quality improvement realizations; that is, $V[\tilde{Z}] = 0$ and $V[\tilde{q}_\tau] = 0\,\forall\tau$. In this case, by equations (10) and (11) the optimal type conditional creation rate choice satisfies,

$$w\tilde{c}'(\gamma_\tau) = \frac{\bar{\pi}_\tau(\sigma)}{r+\delta-g(1-\sigma)} + \frac{\gamma_\tau w\tilde{c}'(\gamma_\tau) - w\tilde{c}(\gamma_\tau)}{r+\delta},\ \tau = 1,\ldots,N.$$

The value added of an average product is found by taking the expectation over the exponential age distribution of products. This yields,

$$E[Y_\tau] = \frac{\delta z}{\delta - g(1-\sigma)} E\big[\tilde{k}_\tau\big],$$

provided that $\delta > g(1-\sigma)$. By equation (22), the expected type conditional wage bill is, $E[W_\tau] = w\ell_\tau E\big[\tilde{k}_\tau\big]$. The average type conditional labor share is then given by,

$$\begin{aligned}\alpha_\tau &= E\left[\frac{W_\tau}{Y_\tau}\right] \\ &= w\left[\frac{1-\bar{\pi}_\tau(\sigma)}{w+\kappa} + \frac{\delta - g(1-\sigma)}{\delta}\tilde{c}(\gamma_\tau)\right],\ \tau = 1,\ldots,N. \qquad (30)\end{aligned}$$

Simulation of (Y, W, N) panel data absent of product demand variation due to age dispersion then follows from the expressions for the type conditional (Y, W, N) at time t,

$$\begin{aligned}Y_{\tau,t}\big(\tilde{k}_\tau\big) &= \tilde{k}_\tau z e^{gt} \\ W_{\tau,t}\big(\tilde{k}_\tau\big) &= \tilde{k}_\tau z \alpha_\tau e^{gt} \\ N_{\tau,t}\big(\tilde{k}_\tau\big) &= \frac{W_{\tau,t}\big(\tilde{k}_\tau\big)}{we^{gt}} = \tilde{k}_\tau z\frac{\alpha_\tau}{w},\end{aligned}$$

where $\tilde{k}_\tau$ is the type conditional product size random variable.

In order to solve for the type conditional dynamics of $(Y_{\tau,t}, W_{\tau,t})$, it is necessary to know (δ, γ_τ) because these two parameters govern the birth-death process of $\tilde{k}_\tau$. Thus, to simulate the full firm panel $\{Y_{jt}, W_{jt}, N_{jt}\}_{j,t}$ for N separate firm types it is necessary to know,

$$\Lambda = \{\delta, z, g, (\alpha_1, \ldots, \alpha_N), (\gamma_1, \ldots, \gamma_N), (\phi_1^*, \ldots, \phi_N^*)\}. \qquad (31)$$

This is $3N+2$ independent parameters given the restriction that $\sum_\tau \phi^*_\tau = 1$. Taking separate identification on w and r, the underlying structural parameters of the simplified model are,

$$\{c_0, c_1, \sigma, \kappa, \eta, z, (q_1, \ldots, q_N), (\phi_1, \ldots, \phi_N)\},$$

which is $2N+5$ independent parameters. Thus, fully separate identification of ω requires that the underlying true data generating process has at least three distinct types and that ω is formulated for at least three types.

Given the other parameters, $\{\delta, Z, (\phi^*_1, \ldots, \phi^*_N)\}$ and $\{\eta, Z, (\phi_1, \ldots, \phi_N)\}$ are related to each other one-to-one. The separate identification discussion can consequently be confined to the relationship between $\{g, (\alpha_1, \ldots, \alpha_N), (\gamma_1, \ldots, \gamma_N)\}$ and $\{\sigma, c_0, c_1, \kappa, (q_1, \ldots, q_N)\}$. Consider the case where $\{\hat{g}, (\hat{\alpha}_1, \ldots, \hat{\alpha}_N), (\hat{\gamma}_1, \ldots, \hat{\gamma}_N)\}$ and $\hat{\delta}$ have been identified by indirect inference such that $\hat{\delta} > \hat{\gamma}_\tau$, $\forall \tau$. There exists a unique η that is consistent with steady state for given $\hat{\delta}$ and $\hat{\gamma}_\tau$ (see appendix for proof). Denote this steady state implied aggregate entry rate by $\hat{\eta}$. Similarly, the steady state product mass distribution across types $\hat{K}_\tau$ also directly follows from the given $\hat{\delta}$ and $\hat{\gamma}_\tau$. $\{\sigma, c_0, c_1, \kappa, (q_1, \ldots, q_N)\}$ is then identified through the system,

$$wc_0(1+c_1)\hat{\gamma}_\tau^{c_1} = \frac{\bar{\pi}_\tau(\sigma)}{r+\hat{\delta}-\hat{g}(1-\sigma)} + \frac{wc_0c_1\hat{\gamma}_\tau^{1+c_1}}{r+\hat{\delta}}, \ \tau = 1, \ldots, N \tag{32}$$

$$\hat{\alpha}_\tau = w\left[\frac{1-\bar{\pi}_\tau(\sigma)}{w+\kappa} + \frac{\hat{\delta}-\hat{g}(1-\sigma)}{\hat{\delta}}c_0\hat{\gamma}_\tau^{1+c_1}\right], \ \tau = 1, \ldots, N. \tag{33}$$

$$\hat{g} = \hat{\eta}\sum_{\tau=1}^{N}\phi_\tau E\big[\ln(q_\tau)\big] + \sum_{\tau=1}^{N}\hat{\gamma}_\tau \hat{K}_\tau E\big[\ln(q_\tau)\big], \tag{34}$$

where the profits have been explicitly stated to depend on the type dependent quality improvement and σ. In the case of 3 distinct types, equations (32)-(34) are 7 equations in 7 unknowns.

Unlike other applications of the Dixit-Stiglitz demand model where σ is identified directly through an average observed markup in the data, the source of identification of σ is non-trivial in this case. Although $\sigma > 1$ imposes an upper bound, markups are first and foremost generated through quality improvements because of the Nash-Bertrand competition between producers of the same intermediate good. As it turns out the data strongly rejects values of σ much larger than one because these values severely dampen value added per worker dispersion across firms. Given the non-conventional identification of σ we provide a robustness analysis of some of the central implications of the estimation to different values of σ in section 6.

As will be shown in the following sections, the true data generating process supports at least 3 distinct types. For all estimations where the model is specified with more than one type, the estimation always yields a low type (order the types so that the low type is indexed by 1) where $\gamma_1 \simeq 0$. The higher types are estimated with $\gamma_\tau > 0\ \forall \tau > 1$. Consequently, since estimated $c_0 > 0$, the non-labor cost share is determined through equation (33) as $\kappa \simeq w(1-\alpha_1)/\alpha_1$.

The estimation of type heterogeneity is tied to three characteristics of the data: 1) The observation of substantial dispersion in value added per worker, 2) the positive relationship between value added per worker and output size, and 3) the flat relationship between value added per worker and input size. Within the framework of the model, it is possible to generate value added per worker dispersion without type heterogeneity through stochastic q realizations. However, it does not generate a positive relationship between value added per worker and output size whereas a negative relationship between labor shares and creation rates will. The positive correlation between Y/N and Y and the zero correlation between Y/N and N will result in an estimate where $\gamma_{\tau'} > \gamma_\tau \Rightarrow \alpha_{\tau'} < \alpha_\tau$. It directly follows from equation (33) that $q_{\tau'} > q_\tau$. As a caveat, it is worthwhile noting that all three characteristics could in principle be a result of simple measurement error in Y. If so, it should be clear that one would need a pretty sizeable measurement error. We have included measurement error in the estimation for this and other reasons. Measurement error is estimated to have little impact on the relationships in the data. The dynamics in the data play an important role as well. They determine magnitudes of the creation-destruction rates through the dispersion in growth rates and changes in cross section moments over time.

The estimation is performed under the assumption that the true firm population of interest coincides with the size censoring in the data. That is, the estimation does not correct for size censoring bias. While a strong assumption, it reasonably assumes that the large number of very small firms in the economy are qualitatively different from those in this analysis and are not just firms with fewer products.

Table 3: Parameter Estimates (standard deviation in parentheses)

			$\tau = 1$	$\tau = 2$	$\tau = 3$
c_0/Z	95.1898 (6.0254)	ϕ_τ	0.8683 (0.0118)	0.0661 (0.0037)	0.0655 (0.0096)
c_1	3.5069 (0.0326)	ϕ^*_τ	0.7667 (0.0153)	0.1145 (0.0054)	0.1188 (0.0147)
κ	152.1388 (1.9631)	ξ_τ	0.0000 (0.0000)	0.3262 (0.0455)	0.9914 (0.1250)
Z	16730.3381 (342.6157)	β_q	0.3779 (0.0250)	0.3779 (0.0250)	0.3779 (0.0250)
β_Z	0.9548 (0.0217)	γ_τ	0.0000 (0.0000)	0.0553 (0.0016)	0.0570 (0.0015)
o_Z	481.0943 (109.5790)	K_τ	0.5701 (0.0171)	0.2029 (0.0140)	0.2270 (0.0220)
η	0.0462 (0.0015)	υ_τ	0.0000 (0.0000)	3.1850 (0.1994)	3.5439 (0.2003)
σ^2_Y	0.0336 (0.0037)	π_τ	0.0000 (0.0000)	0.2486 (0.0137)	0.2719 (0.0137)
σ^2_W	0.0228 (0.0661)	Med[$\tilde{q}_\tau$]	1.0000 (0.0000)	1.1237 (0.0213)	1.1546 (0.0325)
δ	0.0704 (0.0017)	E[$\ln \tilde{q}_\tau$]	0.0000 (0.0000)	0.4339 (0.0352)	0.4921 (0.0361)
m	1.4660 (0.0598)	E[$\tilde{k}_\tau$]	1.0000 (0.0000)	2.3827 (0.1197)	2.5703 (0.1495)
ℓ	44.6786 (0.9175)				
g	0.0141 (0.0006)				
σ	1.0818				

Note: Equilibrium wage is estimated at $w = 190.239$. Standard errors obtained by bootstrap subject to σ estimate.

4.6 Estimation Results

The model parameter estimates are presented in Table 3. As mentioned, the standard errors are obtained through bootstrapping. The bootstrap procedure yields estimates conditional on the point estimate of σ.[15]

[15]The introduction of σ into the set of model parameters to be estimated seems to induce local minima in the criterion function in equation (27). The production of the point estimate in Table 3 consequently involves extensive use of global search methods and is quite time consuming (1 week) even with the use of parallel computation methods

The model estimates imply that at least 3 distinct types have significant representation in the steady state equilibrium.[16] The low type produces almost no improvement in quality whereas the median quality improvement of the middle and high type are 12.4% and 15.5%, respectively. The low type represents 76.7% of all firms and produces 57% of the products in steady state. This is in stark contrast to the low type's representation at entry, which is estimated at $\phi_1 = .868$. This reflects a significant selection in steady state.

The low type's creation rate is almost zero, whereas the middle and high types have creation rates at $\gamma_2 = .055$ and $\gamma_3 = 0.057$, respectively. The high and middle types are in effect crowding out the the low type through the creative destruction process. This is true in terms of firm representation because the low type has a higher exit rate, but it is particularly strong in terms of product representation because the middle and high types are substantially larger in expectation than the low type. The survival conditional size expectation is only barely above a single product for the low type whereas it is 2.4 products for the middle type and 2.6 products for the high type.

The overall creation and destruction rate is estimated at an annual rate of .07. The implied average lifespan of a product is about 14 years. The destruction rate is roughly consistent with evidence in Rosholm and Svarer (2000) that the worker flow from employment to unemployment is roughly 10% annually.

The demand distribution is estimated to be close to an exponential distribution with substantial dispersion. The low type firm will employ 49 manufacturing workers for the average demand realization. The measurement error processes are estimated to produce modest amounts of measurement error noise. The size of the labor supply is inferred for the estimated market wage w. The estimated steady state implies that 4.5% of the labor force is engaged in innovation. The remainder is employed in production.

In the appendix, we include estimates of the model by industry. The results confirm that the central qualitative features of the data persist at the disaggregate level. In particular, one observes

on a large computer cluster. This is prohibitively long for the bootstrap and a compromise was made to produce bootstrapped errors conditional on the σ estimate. One may worry that the presence of local minima could somehow signify identification problems. However, plotting out the minimized criteria function (27) for given values of sigma produces a globally concave function with a single minimum (see Appendix C). Thus, to the best of our knowledge σ is identified, but questions remain about the precision of the estimate. This is part of the motivation of the robustness study in section 6.

[16] An estimation with 4 discrete types did not improve the fit appreciably.

Table 4: Model Fit (data in top row, estimated model in bold in bottom row)

	1992	1997		1992	1996
Survivors	4872.000	3628.000	$\text{Cor}[\frac{Y}{N^*}, \frac{Y_{+1}}{N^*_{+1}}]$	0.476	0.550
	4872.000	**3594.621**		**0.720**	**0.719**
$\text{E}[Y]$	26277.262	31860.850	$\text{Cor}[\frac{Y}{N^*}, \Delta\frac{Y}{N^*}]$	-0.227	-0.193
	23119.155	**27096.128**		**-0.341**	**-0.353**
$\text{Med}[Y]$	13472.812	16448.965	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta Y}{Y}]$	-0.120	
	13327.214	**15207.558**		**-0.099**	
$\text{Std}[Y]$	52793.105	64120.233	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta N^*}{N^*}]$	0.119	
	31730.683	**37685.054**		**0.120**	
$\text{E}[W]$	13294.479	15705.087	$\text{E}[\frac{\Delta Y}{Y}]$	-0.029	
	11812.141	**13658.271**		**0.022**	
$\text{Med}[W]$	7231.812	8671.939	$\text{Std}[\frac{\Delta Y}{Y}]$	0.550	
	7115.871	**8078.232**		**0.781**	
$\text{Std}[W]$	30613.801	35555.701	$\text{Cor}[\frac{\Delta Y}{Y}, Y]$	-0.061	
	14961.133	**17571.025**		**-0.042**	
$\text{E}[\frac{Y}{N^*}]$	384.401	432.118	Within	1.015	
	379.719	**416.989**		**0.951**	
$\text{Med}[\frac{Y}{N^*}]$	348.148	375.739	Between	0.453	
	346.478	**378.693**		**0.342**	
$\text{Std}[\frac{Y}{N^*}]$	205.074	305.306	Cross	-0.551	
	200.013	**221.175**		**-0.405**	
$\text{Cor}[Y, W]$	0.852	0.857	Exit	0.084	
	0.928	**0.928**		**0.112**	
$\text{Cor}[\frac{Y}{N^*}, N^*]$	-0.018	-0.026			
	-0.029	**-0.022**			
$\text{Cor}[\frac{Y}{N^*}, Y]$	0.198	0.143			
	0.173	**0.179**			

significant firm productivity dispersion along with a positive correlation between productivity and output size, but a virtually zero correlation between productivity and input size in all industries.

5 Model Fit

Table 4 shows a comparison of the data moments and the simulated moments for the estimated model.

5.1 Size distributions.

The estimated size distributions do not quite match the heaviness of the right tail in the data. As a result, the model under-estimates the first and second moments of the distributions while matching the median. While generally performing well in terms of matching firm size distributions, the problems fitting the heavy far right tail in empirical size distributions is a well known issue associated with the Klette and Kortum (2004) model. Improvements of the model along this dimension is a topic well worth of future research.

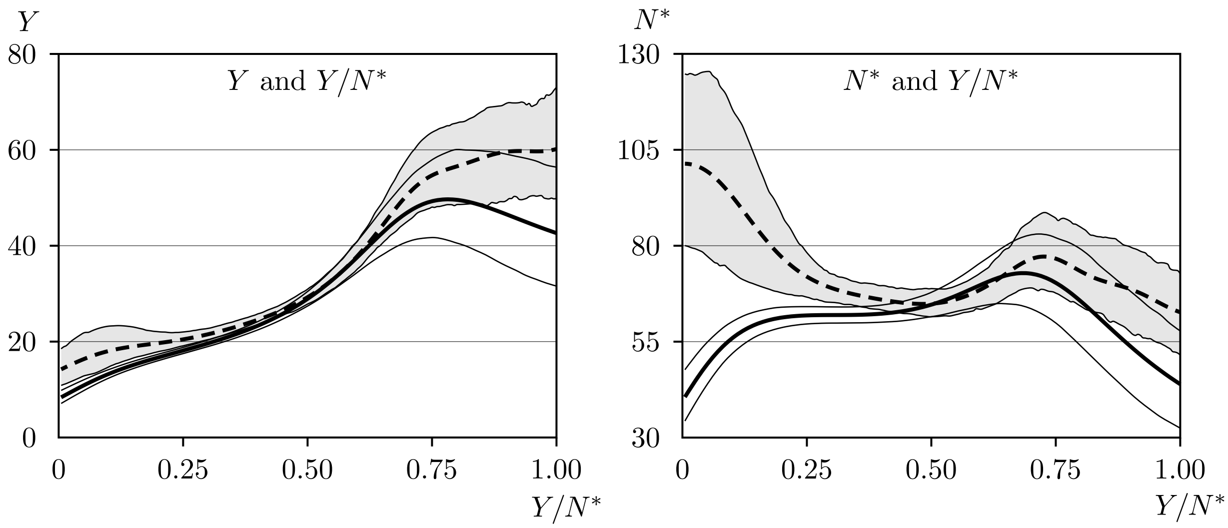
Size dispersion is impacted by the stochastic birth-death process in products, demand realization variation and potentially measurement error. Model simulation without measurement error ($\sigma_Y^2 = \sigma_W^2 = 0$) yields a reduction in the 1992 Std[Y] estimate to $30,916.42$. If in addition demand shock variation is eliminated, that is $o_z = E[\tilde{Z}]$, the 1992 Std[Y] estimate is reduced to $23,309.43$. Thus, a first order explanation of firm size dispersion is found in the relationship between the type conditional creation rates γ_τ and the rate of total creative-destruction δ. These type conditional relationships determine the distribution of $\tilde{k}_\tau$. Demand realization variation plays a secondary role and measurement error almost none.

5.2 Productivity and size correlations.

Figure 2 shows non-parametric regressions of empirical firm productivity and size for both data and the estimated model. The model performs reasonably well in explaining the relationships, at least in the central portion of the distribution of labor productivity. Table 4 shows that the model fits the correlations it has been trained to fit very well.

As mentioned in the previous section, firm type heterogeneity plays an important role in explaining the productivity and size correlations. The positive correlation between value added per worker measure and output size in the data suggests a negative relationship between the firm's wage share and its growth rate. In the model, firm type heterogeneity delivers these relationships by positively correlating high productivity types with high growth rates. The zero correlation between input size and value added per worker is delivered by balancing the labor saving feature of innovations at the product level with the greater growth rate of higher productivity firms.

Figure 2: Firm productivity and size, 1992.



Note: Value added (Y) measured in 1 million DKK. Labor force size (N^*) measured in efficiency units. Estimated model point estimate and 90% confidence bounds drawn in solid pen. Data in dashed pen. Shaded areas are 90% confidence bounds on data.

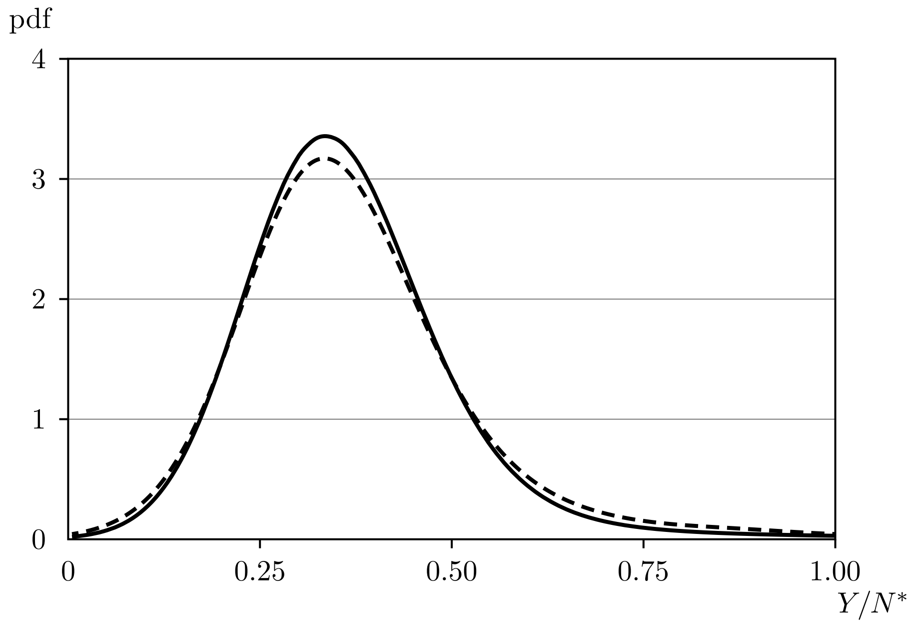
Measurement error has the potential of explaining these correlations as well. The estimation allows for both input and output measurement error which are estimated at fairly moderate amounts. If the model is simulated without the measurement error ($\sigma_Y^2 = \sigma_W^2 = 0$), the 1992 size–productivity correlations change to $\text{Cor}[Y/N, Y] = 0.144$ and $\text{Cor}[Y/N, N] = 0.006$. Thus, measurement error is estimated to have little impact on these moments in the data. Rather, they are explained as a result of the labor saving innovation process at the heart of the model combined with type heterogeneity which yields not only value added per worker dispersion across types, but also different growth rates across types.

5.3 Value added per worker distribution

Figure 3 compares the distribution of empirical firm productivity in data with the estimated model. The model does a good job of explaining this important feature of the data.

The distribution of empirical firm labor productivity Y/N is explained primarily by type heterogeneity, within type quality improvement dispersion, and the capital share. The level of value added per worker is closely linked to the estimate of κ. In particular, since the low firm type is estimated to have an almost zero creation rate and to add almost no improvement to quality, κ is

Figure 3: Firm productivity distribution fit, 1992



Note: Value added (Y) measured in 1 million DKK. Labor force size (N^*) measured in efficiency units. Estimated model in solid pen. Data in dashed pen.

estimated as $\kappa \simeq w(1-\alpha_1)/\alpha_1$. This implies a low type wage share of 55.5%. The average wage share in the data is roughly 55%.

Measurement error adds to the dispersion measure, but to a much smaller extent than firm type heterogeneity. Simulation without measurement error ($\sigma_Y^2 = \sigma_W^2 = 0$) yields a reduction in the 1992 Y/N standard deviation measure to 168.92.

In the absence of innovation labor demand, demand side shocks have no impact on the value added per worker of the firm because manufacturing labor demand and value added move proportionally in response to demand realizations. However, demand side shocks can affect value added per worker dispersion through its effect on the relative size of the manufacturing and innovation labor demands. If in addition to zero measurement error shocks, the model is also estimated without demand realization dispersion, the 1992 Std$[Y/N]$ estimate barely changes, which means that demand realization dispersion has almost no impact on the Y/N distribution.

5.4 Cross-section shifts from 1992 to 1997

Both the size and empirical productivity distributions shift right from 1992 to 1997. The model explains this through the growth rate g and as a result of a survivor bias property of the sample. The exit hazard is higher for smaller firms both in the data and in the model. As a result the mass at the lower end of the size distribution is reduced at a faster pace than elsewhere in the distribution. Furthermore, type heterogeneity also contributes to the right shift since larger firms tend to be of the high type which have lower net destruction rates. Thus, the general turnover in the model as represented by δ has an impact on the right shift as well.

The survivor bias and type heterogeneity account for a substantial part of the right shift in the size distributions. If growth is set to zero by artificially imposing a constant price index, one finds that $\mathrm{E}Y$ estimate shifts from $23,118$ in 1992 to $24,923$ in 1997. Thus, growth accounts for the remainder of the right shift in Table 4.

While growth explains a little less than half of the right shift in the size distributions, it explains almost all of the right shift of the productivity distribution. If growth is set to zero, the $\mathrm{E}[Y/N]$ estimate shifts from 379.76 in 1992 to 383.32 in 1997. Thus, growth accounts for about 90% of the estimated right shift of the productivity distribution.

In general, the estimation does not match the full right shift of the size distributions whereas it does well in capturing the right shift of the productivity distribution. Clearly, the estimation has traded off fitting the right shift of the productivity distribution with the right shift of the size distributions. If g were set higher, the estimation would have done better with the right shift of the size distributions but would have over-shot the right shift of the productivity distribution. Furthermore, it would also have taken the average firm growth rate in the wrong direction.

5.5 Firm growth rate distribution and exit hazards

The model does well in terms of capturing the amount of firm exit as well as the distribution of firm growth rates. δ is a particularly important parameter in this respect. Firm exit is directly tied to δ because it is the exit hazard of a one product firm. But, because the level of δ also determines the amount of overall turnover, it impacts the dispersion in growth rates as well. The average firm growth rate is estimated a little too high. Again, the estimation has faced a trade-off between increasing the δ estimate to reduce the average firm growth rate and reducing the number of survivors which is estimated a little below the observed number.

5.6 Value added per worker persistence and mean reversion.

The model estimates at the top of the second column of Table 4 imply too much persistence and mean reversion. The persistence in firm labor productivity can be explained directly through demand and supply shocks, the magnitudes of the creation and destruction rates γ_τ and δ, and measurement error. The given estimate of the overall creation and destruction rate implies that both the supply and the demand shock processes are quite permanent.

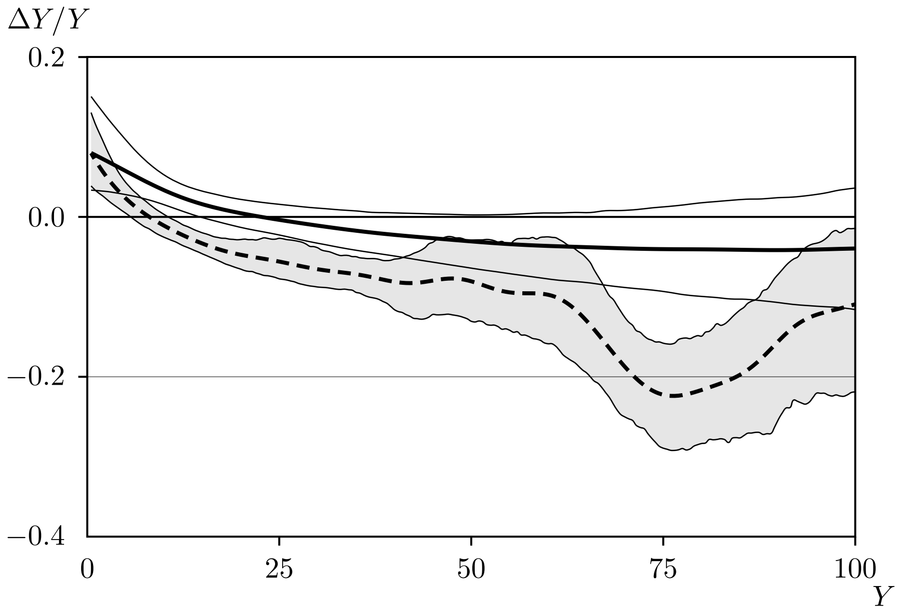
In the absence of measurement error, the model estimate implies very high persistence of value added per worker. In this case one obtains 1992 persistence and mean reversion moments of $\mathrm{Cor}\left[\frac{Y}{N}, \frac{Y_{+1}}{N_{+1}}\right] = .965$ and $\mathrm{Cor}\left[\frac{Y}{N}, \Delta\frac{Y}{N}\right] = -.019$. Adding measurement error reduces the permanence measure and increases the mean reversion measure. Given the one instrument, the estimation has traded off an under-estimate of persistence and an over-estimate of mean reversion.

It is important to note that transitory demand shocks have much the same impact as the measurement error components along this dimension. One can speculate that the introduction of an additional demand noise component of a more transitory nature will result in a lower measurement error noise estimate.

5.7 Growth rate and size (Gibrat Law)

Beginning with Gibrat (1931), much emphasis has been placed on the relationship between firm growth and firm size. Gibrat's law is interpreted to imply that a firm's growth rate is size independent and a large literature has followed testing the validity of this law. See Sutton (1997) for a survey of the literature. No real consensus seems to exist, but at least on the study of continuing establishments, a number of researchers have found a negative relationship between firm size and growth rate. For a recent example, see Rossi-Hansberg and Wright (2005). One can make the argument that Gibrat's law should not necessarily hold at the establishment level and that one must include firm death in order to correct for survivor bias. Certainly, if the underlying discussion

Figure 4: Kernel Regression of Firm Growth Rate and Size (1992).



Note: Value added (Y) measured in 1 million DKK. Model and 90% point wise confidence bounds in solid pen. Data in dashed pen. Shaded area is a 90% point wise confidence bound on data.

is about some broad notion of decreasing returns to scale in production, it is more likely to be relevant at the establishment level than at the firm level. However, as can be seen from Figure 4, in the current sample of firms where the growth rate – size regression includes firm exits, one still obtains a negative relationship.

At a theoretical level, the model satisfies Gibrat's law in the sense that each firm's expected growth is size independent. But two opposing effects will impact the unconditional size and growth relationship: First, due to selection, larger firms will tend to over-represent higher creation rate types and in isolation the selection effect will make for a positive relationship between size and the unconditional firm growth rate. Second, the mean reversion in demand shocks, measurement error, and to a smaller extend in supply shocks introduces an opposite effect: The group of small firms today will tend to over-represent firms with negative demand and measurement error shocks. Chances are that the demand realization of the next innovation will reverse the fortunes of these firms and they will experience relatively large growth rates. On a period-by-period basis, the same is true for the measurement error processes that are assumed to be iid over time. Large firms have

Table 5: FHK growth decomposition and counterfactuals.

	Data	Point Estimate	Steady state with entry: Point Estimate	Counter-factual 1	Counter-factual 2	Counter-factual 3
Within	1.0149	0.9511	1.1216	0.8872	0.9303	0.8013
Between	0.4525	0.3417	0.2919	0.0734	0.1142	0.0000
Cross	−0.5514	−0.4047	−0.6087	−0.1548	−0.2390	0.0000
Exit	0.0839	0.1119	0.1331	0.1318	0.1314	0.1279
Entry	—	—	0.0621	0.0624	0.0631	0.0709
Survivor growth rate	0.0165	0.0165	0.0165	0.0164	0.0165	0.0165
Growth rate	—	—	0.0141	0.0141	0.0141	0.0141

Note: All counterfactuals are performed on the estimated model.
Counterfactual 1 imposes zero measurement error, $\sigma_Y^2 = \sigma_W^2 = 0$.
In addition to counterfactual 1, counterfactual 2 imposes zero demand realization dispersion while maintaining $\mathrm{E}[\tilde{Z}]$ at its estimated level, $\mathrm{Var}[\tilde{Z}] = 0$.
In addition to counterfactual 2, counterfactual 3 imposes within type zero quality improvement dispersion, $\tilde{q}_\tau = (1-\kappa)/(1-\kappa-\bar{\pi}_\tau)\,\forall\tau$. Also, product age dispersion is eliminated.
The growth rate is the annual growth rate in aggregate value added per worker calculated on existing firms at a given point in time. The survivor growth rate describes the value added per worker growth rate of surviving 1992 firms.

many products and experience less overall demand variance. The demand shock and measurement error effects dominate in the estimated model as can be seen in Figure 4.[17] Note that the growth statistics include firm death. If firm deaths are excluded and the statistic is calculated only on survivors, the survival bias will steepen the negative relationship between firm size and firm growth both for the data and for the model since the model reproduces the higher exit hazard rate for small firms that is also found in data.

If the model is estimated without measurement error and demand realization dispersion, one obtains a 1992 firm size and growth correlation of $\mathrm{Cor}[\Delta Y/Y, Y] = .04$. This reflects the positive selection effect.

5.8 Labor productivity growth decomposition

Table 5 presents the decomposition results. Because there is no entry in the observed panel, no entry component is recorded in the data and point estimate columns. Furthermore, it is important to keep in mind that the growth in value added per worker is in part a result of survivor selection

[17] Figure 4 uses value added as the firm size measure. Using labor force size as the size measure instead results in a very similar looking figure and no significant change in the correlation between size and growth.

bias. Therefore, the table presents both a growth measure of the selected sample of surviving 1992 firms and the overall steady state growth rate.

The third column of Table 5 presents the decomposition performed on the estimated model where entry has been included in the simulation.[18] Consequently, each time period is a reflection of the same steady state. Thus, in the third column one can clearly see the survivor selection bias in the productivity growth rate in the data: The steady state annual growth rate is .014 whereas the annual growth rate of the survivor selected sample is .016.

The remaining 3 columns present counterfactuals for the estimated steady state model where the noise processes are gradually shut off. This will impact the decomposition results but will not affect the aggregate growth rate.

A comparison of the first two columns in Table 5 reveals that the estimated model fits the BHC type growth distribution components fairly well. The results, particularly the sign pattern, are generally consistent those found in the empirical literature. For example, the component shares for the decomposition of labor productivity growth with employment weights for U.S. Manufacturing over the period 1977-1987 reported by Foster, Haltiwanger, and Krizan (2001) are within = 0.74, between = 0.08, cross = -0.14, and net entry = 0.29.

The model has three major noise processes: Measurement error, stochastic demand realizations, and within type stochastic quality improvement realizations. In addition product age dispersion will add demand variation to the extent that $\sigma \neq 1$. The measurement error and stochastic q realizations turn out to be the most important in terms of explaining the between and cross components. Given the discussion of the decomposition in section 4.1, it is not surprising that an explanation of a non-zero between and cross components requires the existence of measurement error and transitory shocks processes. In the absence of noise processes, both terms will be zero in steady state.

The first counterfactual sets measurement error to zero. It is seen that this alone dramatically reduces the magnitude of both the between and the cross components. The second counterfactual

[18] Entry is simulated much the same way that the model is simulated forward as described in section 4.4. Each year is divided into n subperiods in which a potential entrant enters with probability $1 - e^{-\gamma_0/n}$. At the time of entry, the type of the entrant is drawn from ϕ and the subsequent life of the entrant is simulated forward just like any incumbent from that point on.

turns off both the measurement error and demand realization noise processes. It is seen that demand realization dispersion has a limited impact on the decomposition.

In addition to setting measurement error and demand noise to zero, the third counterfactual eliminates within type q realization dispersion by deterministically setting each firm's q_τ to match the estimated expected profit π_τ. In this case the theoretical result that the between and cross components equal zero almost holds. If in addition one artificially eliminates product age dispersion effects on demand one obtains that the between and cross component are exactly zero.

It is an important point that the observation of non-zero between and cross components does not necessarily imply that the data reflect an out-of-steady-state situation. The results in Table 5 show that it may simply reflect the existence of various types of noise processes. Especially those that produce noise in the productivity measure.

6 Reallocation and Growth

Aggregate productivity growth as defined in equation (24) is the sum of the contributions of entrants and incumbents where each term is equal to the sum over types of the average increases in the productivity of innovations relative to the product or service replaced, the product of the innovation frequency and quality improvement per innovation, weighted by relative sizes as reflected in the fraction of product lines supplied by each type. Note that the appropriate empirical counterpart is the traditional growth accounting measure recommended by Petrin and Levinsohn (2005), not the BHC productivity difference index.

Table 6 presents the equilibrium steady state annual growth rate implied by the estimated model, $g = 0.0141$. As reported in Table 5, the annual growth rate for the sample is 0.0165. This estimate is biased upward because of survivor selection. The traditional growth measure (the TD index) using the value added per worker for continuing firms only is 0.0148. However, the TD index will be biased upward as well in the presence of exit hazard heterogeneity that is negatively correlated with firm growth rates. The model estimate of the steady state growth rate $g = 0.0141$ provides a structural adjustment of the sample selection bias in the data.

The model also permits the identification of the contribution of survival and firm size selection,

Table 6: The Productivity Growth Rate and Its Components (std dev in parenthesis).

	$\sigma = 1.0818$	$\sigma = 0.600$	$\sigma = 1.000$	$\sigma = 1.250$
g	0.0141 (0.0006)	0.0131 —	0.0137 —	0.0143 —
Decomposition (fraction of g)				
– Entry/exit	0.2003 (0.0142)	0.2208 —	0.2122 —	0.2422 —
– Residual	0.2438 (0.0124)	0.2523 —	0.2462 —	0.2577 —
– Selection	0.5559 (0.0260)	0.5269 —	0.5415 —	0.5001 —

Note: Standard deviation estimates obtained by bootstrap. Standard deviation estimates have not been calculated for the robustness analysis.

reflected in differential firm growth rates, to aggregate growth. Specifically, because the expected productivity of the products created differ across firms and because these differences are positively associated with differences in expected profitability and consequently in creation rates, aggregate growth reflects the selection of more profitable firms by the creative-destruction process. Indeed, equation (24) can be rewritten as

$$g = \sum_{\tau} \gamma_\tau E\left[\ln \tilde{q}_\tau\right] \phi_\tau + \sum_{\tau} \gamma_\tau E\left[\ln \tilde{q}_\tau\right] (K_\tau - \phi_\tau) + \eta \sum_{\tau} E_\tau\left[\ln \tilde{q}_\tau\right] \phi_\tau \tag{35}$$

where the first term is the contribution to growth of continuing firms under the counter factual assumption that the share of products supplied by continuing firms of each type is the same as at entry, the second term accounts for type distributional impact of differential firm growth rates after entry, and the third term is the net contribution of entry and exit. Because the steady state fraction of products supplied by type τ firms is $K_\tau = \eta\phi_\tau/(\delta - \gamma_\tau)$, the selection effect is positive because firms that are expected to create higher quality products supply more product lines on average. (Formally, stochastic dominance $F_\tau \leq F_{\tau'} \Longrightarrow$ both $E\left[\ln \tilde{q}_\tau\right] \geq E\left[\ln \tilde{q}_{\tau'}\right]$ and $K_\tau - \phi_\tau \geq K_{\tau'} - \phi_{\tau'}$.)

Table 6 presents the decomposition estimate. The estimated model implies that the entry/exit component accounts for 20.03% and the selection component 55.6% of the aggregate growth rate. Hence, the dynamics of entry and firm size evolution, a process that involves continual reallocation to new and growing firms, is responsible for over three-quarters of the growth in the modelled economy.

Of course, all growth in the model is a result of worker reallocation. Every time a new innovation is created workers flow to its supplier from firms with products that have recently become obsolete. This observation does not address the issue of resource allocation across types, however. The selection effect measures the loss in productivity growth that would result if more productive firm types in any given cohort were not allowed to increase their resource share relative to that at birth. Because the more productive types in any cohort will gradually gain an ever increasing share of resources of that cohort, as a cohort ages it becomes increasingly selected while also shrinking relative to the size of the overall economy. The model's steady state is the sum of overlapping cohorts with different degrees of selection. The steady state distribution of product lines and the resources required to supply them remains constant over time because firms in the existing cohorts contract at a rate equal to the entry rate of new firms of that type. If the reallocation induced by selection is shut down, the distribution of product lines across types will gradually deteriorate to the point where it equals the entry distribution, and productivity growth will fall by 55.6%.

Finally, Table 6 also includes a check of the robustness of the growth decomposition with respect to the σ estimate. The values for σ span an interval where the model fit remains relatively good in comparison to the fit of the point estimates reported in Table 3. The span covers both the cases of complements and substitutes. It is seen that the growth estimate and the decomposition results are not particularly sensitive to the value of σ. The overall growth estimate is increasing in σ which is a result of the age effect on demand of surviving products. In the complements case, the demand of surviving products increases with age whereas the opposite is true in the substitutes case. The data moments are based on a selected sample of surviving firms. Hence, the steady state growth rate will adjust to size of the survivor bias, which is impacted by the age effect. For values of σ above 1.25, the model fit deteriorates rapidly as a result of a reduction in the dispersion of value added per worker across firms.

7 Concluding Remarks

Large and persistent differences in firm productivity and firm size exist. Worker reallocation induced by heterogeneity should be an important source of aggregate productivity growth. However, empirical studies based on the Baily, Hulten, and Campbell (1992) growth decomposition have found mixed evidence of the importance reallocation as a source of growth. We argue that the BHC growth decomposition does not correctly identify the steady state contribution of resource reallocation to productivity growth. Indeed, we show that models in which the distribution of resources across firm types is stationary imply that the "between" and "cross" firm components of the decomposition are zero in the absence of transitory noise what ever the true data generating process.

In this paper we explore a variant of the equilibrium Schumpeterian model of firm size evolution developed by Klette and Kortum (2004). In our version of the model, firms that can develop products of higher quality have an incentive to grow faster relative to less profitable firms in each cohort though a process of creative destruction. Worker reallocation from less to more profitable firms induced by the process contributes to aggregate productivity growth. Furthermore, the model is consistent with the observation that there is no correlation between employment size and labor productivity and a positive correlation between value added and labor productivity observed in Danish firm data.

We fit the model to the Danish firm panel for the 1992 − 1997 time period. The parameter estimates are sensible and the model provides a good fit to the joint size distribution and dynamic moments of the data. Although the model fits the Foster, Haltiwanger, and Krizan (2001) variant of the BHC growth decomposition well, the "between" and "cross" terms vanish in a counterfactual exercise in which purely transitory shocks and measurement errors are set to zero. Finally, the estimated model also fits the negative relationship between size and growth in the data even though at a theoretical level it satisfies Gibrat's law in the sense that a firm's innovation rate is independent of its size.

All growth in our model is attributed to reallocation in the sense that resources must flow from firms that loss markets to innovators that provide new more productive goods and services. We decompose the reallocation component into a net contribution from firm entry and exit, a firm type selection effect, and a residual. The net contribution of entry is 20% of the model's implied growth rate. The selection component, which accounts for 55% of growth, captures the contribution attributable to the fact that resources are reallocated from slow growing less productive firms to fast growing more creative ones in each cohort.

A Steady State Equilibrium Solution Algorithm

In this section, we present the steady state equilibrium solution algorithm given the parameter vector $\big(\kappa, \eta, w, \sigma, \phi, F(\cdot), c(\cdot)\big)$. Both η and w are endogenous to the equilibrium. w is estimated as the average 1992 worker wage in the data and ℓ is subsequently set to match the aggregate labor demand for the equilibrium wage. Similarly, the estimate of η maps directly to the estimate of m for the estimated model through the first order condition,

$$wc'(\eta/m) = \sum_{\tau} \phi_\tau v_\tau.$$

The core of the solution algorithm is based on the following proposition,

Proposition 1 *There exists a unique steady state (K, δ, g) for any given (η, γ, ϕ) such that $\eta > 0$, $\gamma_\tau \geq 0\,\forall\tau$, and $\phi_\tau > 0\,\forall\tau$. The steady state satisfies,*

$$\begin{aligned} K_\tau &= \frac{\eta\phi_\tau}{\delta - \gamma_\tau} > 0\,\forall\tau \\ \delta &= \eta + \sum_\tau K_\tau\gamma_\tau > 0 \\ g &= \sum_\tau (K_\tau + \eta\phi_\tau)\gamma_\tau E\big[\ln(\tilde{q}_\tau)\big] \geq 0. \end{aligned}$$

Proof. The evolution of the distribution of products across the n firm types can be written as,

$$\dot{K}_\tau = \gamma_\tau K_\tau + \eta\phi_\tau - \delta K_\tau, \tau = 1, \ldots, n,$$

where $\sum K_\tau = 1$. In steady state, this reduces to,

$$\delta = \eta + \sum_{\tau=1}^{n} \gamma_\tau K_\tau, \tag{36}$$

and,

$$K_\tau = \frac{\eta\phi_\tau}{\delta - \gamma_\tau} = \frac{\eta\phi_\tau}{\eta + \sum_{i\neq\tau}\gamma_i K_i - (1 - K_\tau)\gamma_\tau} \tag{37}$$

$$\Updownarrow$$

$$K_\tau = \frac{\eta\phi_\tau + K_\tau(1 - K_\tau)\gamma_\tau}{\eta + \sum_{i\neq\tau}\gamma_i K_i} \equiv \Gamma_\tau(K_\tau, K_{-\tau}), \tau = 1, \ldots, n. \tag{38}$$

$\Gamma_\tau(K_\tau, K_{-\tau})$ is a continuous, strictly concave function in $K_\tau \in [0, 1]$. Since $\Gamma_\tau(0, K_{-\tau}) = \Gamma_\tau(1, K_{-\tau}) \in (0, 1)$ there exists a unique fixed point $\hat{K}_\tau = \Gamma_\tau\big(\hat{K}_\tau, K_{-\tau}\big)$ where $\hat{K}_\tau \in (0, 1)$. Define the mapping

$\Omega : [0,1]^n \rightarrow (0,1)^n$ such that $\Omega(K) = \Gamma(\Omega(K), K)$. $\Omega(K)$ is continuous in K and since $[0,1]^n$ is a compact and convex set, Brouwer's fixed point theorem can be applied to prove the existence of a fixed point $K^* = \Omega(K^*)$. It follows by $K^*_\tau \in (0,1), \forall\tau$, that any fixed point has the property that $\delta > \gamma_\tau$, $\forall\tau$.

To prove uniqueness, suppose to the contrary that there exists two distinct fixed points $K^0 = \Omega\left(K^0\right)$ and $K^1 = \Omega\left(K^1\right)$, $K^0 \neq K^1$. Denote by $\delta^i = \eta + \sum_\tau \gamma_\tau K^i_\tau$, $i = 1, 2$. Since K^i is a solution to (38) it must be that $\sum_\tau K^i_\tau = \sum_\tau \frac{\eta\phi_\tau}{\delta^i - \gamma_\tau} = 1$, $i = 1, 2$. Combined with $\delta^i > \gamma_\tau, \forall\tau$, $i = 1, 2$ this implies that $\delta^0 = \delta^1$. But it then follows that $K^0_\tau = \frac{\eta\phi_\tau}{\delta^0 - \gamma_\tau} = \frac{\eta\phi_\tau}{\delta^1 - \gamma_\tau} = K^1_\tau$, $\forall\tau$, contradicting the assumption of two distinct fixed points. ■

The type conditional creation rate choice satisfies,

$$w\hat{c}'(\gamma_\tau) = \nu_\tau, \tag{39}$$

where

$$\begin{aligned} \Psi_\tau &= \max_{\gamma \geq 0} \frac{\gamma\nu_\tau - w\hat{c}(\gamma)}{r + \delta} \\ v_\tau &= \frac{\bar{\pi}_\tau}{r + \delta - g(1-\sigma)} + \Psi_\tau. \end{aligned}$$

Given the model parameters $\left(\kappa, \eta, w, \sigma, \phi, F(\cdot), c(\cdot)\right)$ where $\eta > 0$, the solution algorithm can be formulated as a fixed point search in (Ψ, δ, g) subject to the constraints

$$\begin{aligned} r + \delta &> g(1-\sigma) \\ \delta &> 0 \\ \Psi_\tau &\geq 0, \forall\tau \\ g &\geq 0. \end{aligned}$$

All four constraints are satisfied in a model equilibrium, but it is worth noting that existence of equilibrium may fail to materialize for certain model parameter combinations because of violation of the first constraint.

For a given (Ψ, δ, g) satisfying the above constraints there exists a unique solution for v_τ which directly yields $\gamma_\tau \geq 0$, $\forall\tau$. With these type conditional creation rates and the given $\eta > 0$, one can then apply the insights of Proposition 1 to yield the implied steady state values for (δ', g'). The search for the steady state reduces to solving a non-linear system of equations. Given the steady state value of the destruction rate, one obtains

$$\Psi'_\tau = \frac{\gamma_\tau v_\tau - w\hat{c}(\gamma_\tau)}{r + \delta'}, \forall \tau.$$

Denote this mapping by, $(\Psi', \delta', g') = \Upsilon(\Psi, \delta, g)$. Steady state equilibrium is a fixed point,

$$(\Psi^*, \delta^*, g^*) = \Upsilon(\Psi^*, \delta^*, g^*).$$

The search for this fixed point can be done in a number of ways. A particularly simple method is straightforward iteration on the mapping Υ. This turns out to be a robust and quick method.

B The BHC Decomposition in Stochastic Steady State

Denote by I_t the set of firms at time t. The aggregate productivity change is written by,

$$\Delta P = \sum_{i \in I_t} s_{it} p_{it} - \sum_{i \in I_{t-1}} s_{it-1} p_{it-1}.$$

Define the set of continuing firms between $t-1$ and t as the intersection of I_{t-1} and I_t, $C_t = I_{t-1} \cap I_t$. Define the set of entrants at time t as complement of I_t and C_t, $E_t = I_t \setminus C_t$. The set of exiting firms between I_{t-1} and I_t is similarly defined as $X_{t-1} = I_{t-1} \setminus C_t$. With this in hand we can write the change in aggregate productivity as,

$$\begin{aligned} \Delta P &= \sum_{i \in E_t} s_{it} p_{it} + \sum_{i \in C_t} s_{it} p_{it} - \sum_{i \in C_t} s_{it-1} p_{it-1} - \sum_{i \in X_{t-1}} s_{it-1} p_{it-1} \\ &= \sum_{i \in C_t} s_{it-1} \Delta p_{it} + \sum_{i \in C_t} \Delta s_{it} p_{it-1} + \sum_{i \in C_t} \Delta s_{it} \Delta p_{it} + \sum_{i \in E_t} s_{it} p_{it} - \sum_{i \in X_{t-1}} s_{it-1} p_{it-1}. \end{aligned}$$

This is the Foster, Haltiwanger, and Krizan (2001) decomposition.

Now, state the summations in terms of groups of identical type firms j. Furthermore, define the resource share of the group of type j firms at time t, by $\bar{s}_{jt} = \frac{1}{|I_j|} \sum_{i \in I_j} s_{it}$. Similarly, define the average productivity of type j firms at time t, by $\bar{p}_{jt} = \frac{1}{|I_j|} \sum_{i \in I_j} p_{it}$. In steady state it must be that $\bar{s}_{jt} = \bar{s}_{jt-1} = \bar{s}_j$ implying that $\Delta \bar{s}_{jt} = 0$. The change in productivity can be written as,

$$\begin{aligned}
\Delta P &= \sum_j \sum_{i \in E_t \cap I_{jt}} s_{it} p_{it} + \sum_j \sum_{i \in C_t \cap I_{jt}} s_{it} p_{it} - \sum_j \sum_{i \in C_t \cap I_{jt-1}} s_{it-1} p_{it-1} - \sum_j \sum_{i \in X_{t-1} \cap I_{jt-1}} s_{it-1} p_{it-1} \\
&= \sum_j \sum_{i \in C_t \cap I_{jt}} s_{it} p_{it} - \sum_j \sum_{i \in C_t \cap I_{jt}} s_{it} p_{it-1} + \sum_j \sum_{i \in C_t \cap I_{jt}} s_{it} p_{it-1} - \sum_j \sum_{i \in C_t \cap I_{t-1}} s_{it-1} p_{it-1} + \\
&\quad \sum_j \sum_{i \in E_t \cap I_j} s_{it} p_{it} - \sum_j \sum_{i \in X_{t-1} \cap I_{t-1}} s_{it-1} p_{it-1} \\
&= \sum_j \sum_{i \in C_t \cap I_{jt}} s_{it} \Delta p_{it} + \sum_j \sum_{i \in C_t \cap I_{jt}} s_{it} p_{it-1} + \sum_j \sum_{i \in E_t \cap I_j} s_{it} p_{it-1} - \\
&\quad \left[\sum_j \sum_{i \in C_t \cap I_{t-1}} s_{it-1} p_{it-1} + \sum_j \sum_{i \in X_{t-1} \cap I_{t-1}} s_{it-1} p_{it-1} \right] + \sum_j \sum_{i \in E_t \cap I_j} s_{it} p_{it} - \sum_j \sum_{i \in E_t \cap I_j} s_{it} p_{it-1} \\
&= \sum_j \sum_{i \in C_t \cap I_{jt}} s_{it} \Delta p_{it} + \sum_j \sum_{i \in I_{jt}} s_{it} p_{it-1} - \sum_j \sum_{i \in I_{jt-1}} s_{it-1} p_{it-1} + \sum_j \sum_{i \in E_t \cap I_j} s_{it} \Delta p_{it} \\
&= \sum_j \Delta \bar{p}_{jt} \sum_{i \in I_{jt}} s_{it} + \sum_j \bar{p}_{jt-1} \sum_{i \in I_{jt}} s_{it} - \sum_j \bar{p}_{jt-1} \sum_{i \in I_{jt-1}} s_{it-1} \\
&= \sum_j \Delta \bar{p}_{jt} \left| I_{jt} \right| \bar{s}_{jt} + \sum_j \bar{p}_{jt-1} \left| I_{jt} \right| \bar{s}_{jt} - \sum_j \bar{p}_{jt-1} \left| I_{jt-1} \right| \bar{s}_{jt-1}.
\end{aligned}$$

In steady state it must be that,

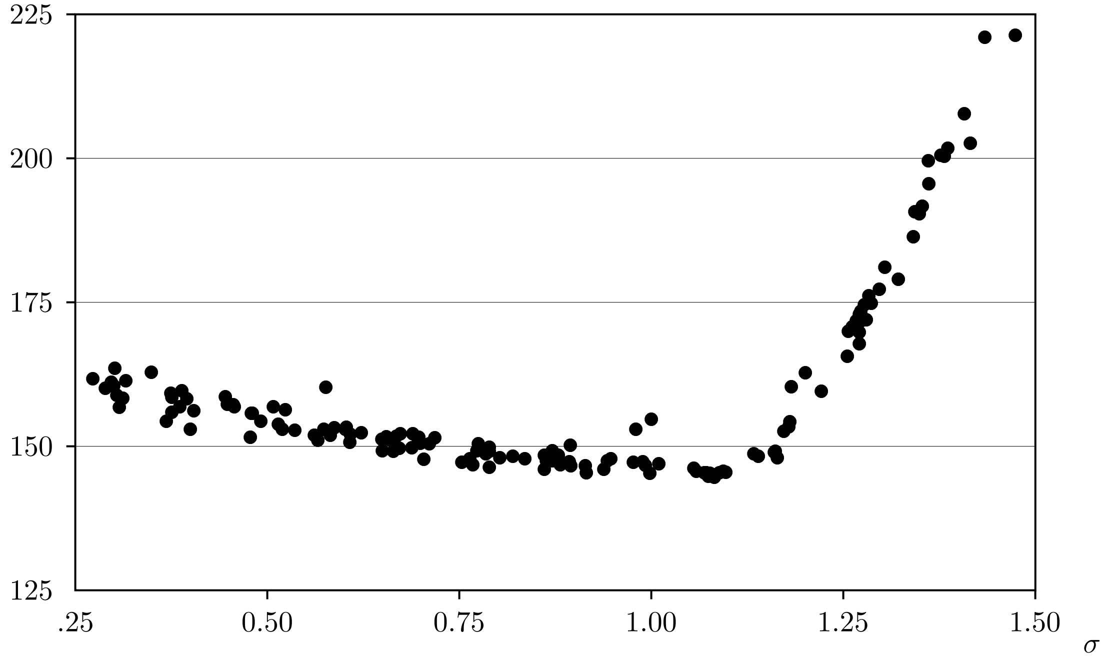
$$\begin{aligned}
\Delta P_t &= \sum_j \left| I_j \right| \bar{s}_j \Delta \bar{p}_{jt} + \sum_j \left[\left| I_j \right| \bar{s}_j - \left| I_j \right| \bar{s}_j \right] \bar{p}_{jt-1} \\
&= \sum_j \left| I_j \right| \bar{s}_j \Delta \bar{p}_{jt}.
\end{aligned}$$

If type were observable, $\Delta \bar{p}_j$ would be estimated by taking the average growth in p for the set of continuing type j firms.

C Identification of σ

Figure 5 plots part of the intermediate output of a global search routine over σ. The points in the figure are minimized values of the criteria function for given values of σ. The points off the minimum frontier reflect local minima that were abandoned by the search routine.

Figure 5: Criteria Function



D Estimation Results by Industry

This section contains model estimates by industry. The estimation is done subject to the $\sigma = 1.0818$ estimate from the full data set since this is considered common to all the industries. We are showing the results for the three largest industries. The smaller industries are too small to yield statistically interesting results.

Table 7: Manufacturing – Parameter Estimates (standard deviation in parentheses)

			$\tau = 1$	$\tau = 2$	$\tau = 3$
c_0/Z	76.2759 (6.7627)	ϕ_τ	0.8342 (0.0191)	0.0147 (0.0326)	0.1511 (0.0363)
c_1	3.2576 (0.0381)	ϕ^*_τ	0.7314 (0.0221)	0.0222 (0.0531)	0.2464 (0.0572)
κ	146.7780 (2.8209)	ξ_τ	0.0000 (0.0000)	0.0976 (0.0707)	0.5677 (0.0524)
Z	19275.7389 (537.6697)	β_q	0.6526 (0.0398)	0.6526 (0.0398)	0.6526 (0.0398)
β_Z	0.8230 (0.0267)	γ_τ	0.0000 (0.0000)	0.0471 (0.0039)	0.0508 (0.0021)
o_Z	3181.3213 (432.9361)	K_τ	0.5572 (0.0229)	0.0327 (0.0898)	0.4101 (0.0935)
η	0.0450 (0.0020)	υ_τ	0.0000 (0.0000)	2.9391 (0.6506)	3.7606 (0.2750)
σ^2_Y	0.0245 (0.0032)	π_τ	0.0000 (0.0000)	0.2412 (0.0482)	0.2978 (0.0183)
σ^2_W	0.0181 (0.1586)	$\text{Med}[\tilde{q}_\tau]$	1.0000 (0.0000)	1.2178 (0.0717)	1.3238 (0.0545)
δ	0.0673 (0.0023)	$\text{E}[\ln \tilde{q}_\tau]$	0.0000 (0.0000)	0.3328 (0.0803)	0.4394 (0.0384)
m	1.5472 (0.0815)	$\text{E}[\tilde{k}_\tau]$	1.0000 (0.0000)	1.9342 (0.1751)	2.1853 (0.1406)
ℓ	51.5503 (1.4053)				
g	0.0129 (0.0008)				
σ	1.0818				

Note: Equilibrium wage is estimated at $w = 190.660$. Point estimate and standard errors are obtained subject to overall product demand specification estimate of $\sigma = 1.0818$.

Table 8: Manufacturing – Model Fit (data in top row, estimated model in bottom row)

	1992	1997		1992	1996
Survivors	2051.000	1536.000	$\mathrm{Cor}[\frac{Y}{N^*}, \frac{Y_{+1}}{N^*_{+1}}]$	0.650	0.728
	2051.000	1536.219		0.644	0.646
$\mathrm{E}[Y]$	30149.461	35803.473	$\mathrm{Cor}[\frac{Y}{N^*}, \Delta\frac{Y}{N^*}]$	−0.024	−0.195
	25974.016	29936.366		−0.397	−0.406
$\mathrm{Med}[Y]$	15117.552	18858.445	$\mathrm{Cor}[\frac{Y}{N^*}, \frac{\Delta Y}{Y}]$	−0.133	
	14819.859	16758.498		−0.127	
$\mathrm{Std}[Y]$	56081.995	69574.991	$\mathrm{Cor}[\frac{Y}{N^*}, \frac{\Delta N^*}{N^*}]$	0.145	
	33152.398	38562.133		0.151	
$\mathrm{E}[W]$	15047.636	17318.195	$\mathrm{E}[\frac{\Delta Y}{Y}]$	−0.035	
	13258.885	15060.236		0.005	
$\mathrm{Med}[W]$	8031.273	9531.066	$\mathrm{Std}[\frac{\Delta Y}{Y}]$	0.474	
	7948.588	8943.336		0.481	
$\mathrm{Std}[W]$	24667.884	27159.439	$\mathrm{Cor}[\frac{\Delta Y}{Y}, Y]$	−0.073	
	15561.368	17813.551		−0.042	
$\mathrm{E}[\frac{Y}{N^*}]$	379.047	422.471	Within	0.863	
	374.100	407.843		0.864	
$\mathrm{Med}[\frac{Y}{N^*}]$	347.100	375.300	Between	0.365	
	347.282	377.355		0.262	
$\mathrm{Std}[\frac{Y}{N^*}]$	163.174	226.860	Cross	−0.297	
	140.725	154.685		−0.256	
$\mathrm{Cor}[Y, W]$	0.889	0.855	Exit	0.068	
	0.934	0.935		0.129	
$\mathrm{Cor}[\frac{Y}{N^*}, N^*]$	0.011	−0.003			
	−0.003	0.009			
$\mathrm{Cor}[\frac{Y}{N^*}, Y]$	0.236	0.200			
	0.220	0.232			

Table 9: Manufacturing – Growth Decomposition (standard deviation in parentheses)

	Point Estimate	Fraction of g
g	0.0129	1.0000
	(0.0008)	—
Entry/exit	0.0032	0.2492
	(0.0003)	(0.0209)
Residual	0.0036	0.2799
	(0.0003)	(0.0139)
Selection	0.0061	0.4710
	(0.0006)	(0.0338)

Table 10: Wholesale and Retail – Parameter Estimates (standard deviation in parentheses)

			$\tau = 1$	$\tau = 2$	$\tau = 3$
c_0/Z	71.9217 (6.2081)	ϕ_τ	0.9269 (0.0065)	0.0012 (0.0083)	0.0719 (0.0096)
c_1	2.9448 (0.0304)	ϕ^*_τ	0.8705 (0.0084)	0.0019 (0.0144)	0.1276 (0.0157)
κ	177.5595 (2.7765)	ξ_τ	0.0000 (0.0000)	0.0168 (0.0232)	2.1016 (0.1732)
Z	16811.6212 (515.8550)	β_q	0.8222 (0.0768)	0.8222 (0.0768)	0.8222 (0.0768)
β_Z	0.9502 (0.0363)	γ_τ	0.0000 (0.0000)	0.0421 (0.0038)	0.0477 (0.0017)
o_Z	1175.5965 (398.9608)	K_τ	0.7502 (0.0127)	0.0030 (0.0270)	0.2468 (0.0294)
η	0.0505 (0.0017)	υ_τ	0.0000 (0.0000)	4.7427 (0.9755)	6.8165 (0.2813)
σ^2_Y	0.0165 (0.0042)	π_τ	0.0000 (0.0000)	0.3878 (0.0709)	0.5289 (0.0207)
σ^2_W	0.0198 (0.3790)	$\text{Med}[\tilde{q}_\tau]$	1.0000 (0.0000)	1.6151 (0.2264)	2.3457 (0.1518)
δ	0.0623 (0.0019)	$\text{E}[\ln \tilde{q}_\tau]$	0.0000 (0.0000)	0.5912 (0.1439)	0.9502 (0.0471)
m	2.5787 (0.1033)	$\text{E}[\tilde{k}_\tau]$	1.0000 (0.0000)	1.8504 (0.1688)	2.2439 (0.1173)
ℓ	41.6051 (1.2684)				
g	0.0147 (0.0008)				
σ	1.0818				

Note: Equilibrium wage is estimated at $w = 187.720$. Point estimate and standard errors are obtained subject to overall product demand specification estimate of $\sigma = 1.0818$.

Table 11: Wholesale and Retail – Model Fit (data in top row, estimated model in bottom row)

	1992	1997		1992	1996
Survivors	1584.000	1189.000	$\text{Cor}[\frac{Y}{N^*}, \frac{Y_{+1}}{N^*_{+1}}]$	0.325	0.674
	1584.000	1185.787		0.739	0.745
$\text{E}[Y]$	22952.920	28386.719	$\text{Cor}[\frac{Y}{N^*}, \Delta\frac{Y}{N^*}]$	−0.195	−0.259
	20178.326	22850.027		−0.324	−0.335
$\text{Med}[Y]$	12757.909	15288.949	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta Y}{Y}]$	−0.088	
	12777.856	14215.385		−0.094	
$\text{Std}[Y]$	33400.313	41409.060	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta N^*}{N^*}]$	0.189	
	24007.935	27848.885		0.164	
$\text{E}[W]$	10696.683	12712.898	$\text{E}[\frac{\Delta Y}{Y}]$	−0.042	
	9403.578	10452.085		−0.005	
$\text{Med}[W]$	6423.473	7650.564	$\text{Std}[\frac{\Delta Y}{Y}]$	0.425	
	6224.264	6894.377		0.424	
$\text{Std}[W]$	15360.222	16802.715	$\text{Cor}[\frac{\Delta Y}{Y}, Y]$	−0.090	
	9887.575	11089.624		−0.034	
$\text{E}[\frac{Y}{N^*}]$	410.234	466.591	Within	1.176	
	403.265	445.119		0.798	
$\text{Med}[\frac{Y}{N^*}]$	373.928	408.244	Between	0.618	
	372.462	408.447		0.236	
$\text{Std}[\frac{Y}{N^*}]$	171.661	278.495	Cross	−0.826	
	166.805	188.294		−0.170	
$\text{Cor}[Y, W]$	0.922	0.914	Exit	0.032	
	0.903	0.900		0.135	
$\text{Cor}[\frac{Y}{N^*}, N^*]$	−0.028	−0.039			
	−0.002	0.017			
$\text{Cor}[\frac{Y}{N^*}, Y]$	0.252	0.188			
	0.288	0.311			

Table 12: Wholesale and Retail – Growth Decomposition (standard deviation in parentheses)

	Point Estimate	Fraction of g
g	0.0147	1.0000
	(0.0008)	—
Entry/exit	0.0035	0.2364
	(0.0003)	(0.0199)
Residual	0.0033	0.2229
	(0.0002)	(0.0110)
Selection	0.0080	0.5407
	(0.0007)	(0.0306)

Table 13: Construction – Parameter Estimates (standard deviation in parentheses)

			$\tau = 1$	$\tau = 2$	$\tau = 3$
c_0/Z	66.6745 (8.0458)	ϕ_τ	0.5522 (0.0441)	0.1969 (0.0150)	0.2509 (0.0425)
c_1	3.6104 (0.0640)	ϕ^*_τ	0.4373 (0.0393)	0.2291 (0.0170)	0.3336 (0.0431)
κ	81.0334 (2.3698)	ξ_τ	0.0000 (0.0000)	0.0902 (0.0153)	1.2743 (0.2217)
Z	8433.0564 (172.4217)	β_q	0.6568 (0.0308)	0.6568 (0.0308)	0.6568 (0.0308)
β_Z	1.3145 (0.1323)	γ_τ	0.0000 (0.0000)	0.0468 (0.0038)	0.0567 (0.0041)
o_Z	3619.6641 (1079.7041)	K_τ	0.3085 (0.0302)	0.2514 (0.0226)	0.4401 (0.0410)
η	0.0465 (0.0024)	υ_τ	0.0000 (0.0000)	0.9345 (0.1450)	1.8701 (0.1661)
σ^2_Y	0.0347 (0.0041)	π_τ	0.0000 (0.0000)	0.0908 (0.0133)	0.1671 (0.0118)
σ^2_W	0.0205 (0.0229)	$\mathrm{Med}[\tilde{q}_\tau]$	1.0000 (0.0000)	1.0516 (0.0091)	1.1202 (0.0139)
δ	0.0833 (0.0044)	$\mathrm{E}[\ln \tilde{q}_\tau]$	0.0000 (0.0000)	0.1039 (0.0161)	0.2106 (0.0182)
m	1.0970 (0.0820)	$\mathrm{E}[\tilde{k}_\tau]$	1.0000 (0.0000)	1.5551 (0.0592)	1.8703 (0.1128)
ℓ	28.4961 (0.6216)				
g	0.0099 (0.0010)				
σ	1.0818				

Note: Equilibrium wage is estimated at $w = 191.849$. Point estimate and standard errors are obtained subject to overall product demand specification estimate of $\sigma = 1.0818$.

Table 14: Construction – Model Fit (data in top row, estimated model in bottom row)

	1992	1997		1992	1996
Survivors	651.000 651.000	480.000 474.231	$\text{Cor}[\frac{Y}{N^*}, \frac{Y_{+1}}{N^*_{+1}}]$	0.428 0.386	0.345 0.379
E[Y]	15191.354 12116.030	16869.551 14011.735	$\text{Cor}[\frac{Y}{N^*}, \Delta\frac{Y}{N^*}]$	−0.327 −0.540	−0.560 −0.550
Med[Y]	8688.501 8844.399	10711.648 9943.091	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta Y}{Y}]$	−0.187 −0.231	
Std[Y]	31287.564 10759.536	22454.655 12657.142	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta N^*}{N^*}]$	0.089 0.192	
E[W]	9973.166 7847.026	10594.737 9012.535	$\text{E}[\frac{\Delta Y}{Y}]$	−0.025 0.004	
Med[W]	5785.053 5907.388	6838.405 6607.510	$\text{Std}[\frac{\Delta Y}{Y}]$	0.448 0.447	
Std[W]	24526.438 6560.671	14181.147 7712.691	$\text{Cor}[\frac{\Delta Y}{Y}, Y]$	−0.122 −0.079	
$\text{E}[\frac{Y}{N^*}]$	305.075 303.475	342.273 323.656	Within	0.986 1.035	
$\text{Med}[\frac{Y}{N^*}]$	286.749 287.092	311.509 306.474	Between	0.635 0.421	
$\text{Std}[\frac{Y}{N^*}]$	133.111 95.424	173.871 100.840	Cross	−0.870 −0.555	
Cor[Y, W]	0.967 0.924	0.922 0.926	Exit	0.249 0.098	
$\text{Cor}[\frac{Y}{N^*}, N^*]$	−0.040 −0.091	−0.093 −0.082			
$\text{Cor}[\frac{Y}{N^*}, Y]$	0.131 0.193	0.174 0.195			

Table 15: Construction – Growth Decomposition (standard deviation in parentheses)

	Point Estimate	Fraction of g
g	0.0099 (0.0010)	1.0000 —
Entry/exit	0.0034 (0.0004)	0.3446 (0.0302)
Residual	0.0040 (0.0005)	0.3999 (0.0191)
Selection	0.0025 (0.0005)	0.2554 (0.0435)

References

Alvarez, Javier, Martin Browning, and Mette Ejrnæs (2001). Modelling income processes with lots of heterogeneity. *Working Paper*.

Aw, Bee Yan, Xiaomin Chen, and Mark J. Roberts (1997). Firm-level evidence on productivity differentials, turnover, and exports in taiwanese manufacturing. *National Bureau of Economic Research, Inc, NBER Working Papers*, no. 6235.

Baily, M., C. Hulten, and D. Campbell (1992). Productivity dynamics in manufacturing plants. *Brookings Papers on Economic Activity, Microeconomics*: 187–249.

Baily, Martin Neil, Eric J. Bartelsman, and John Haltiwanger (1996). Downsizing and productivity growth: Myth or reality? *Small Business Economics* 8, no. 4: 259–78.

Bartelsman, Eric J. and Phoebus J. Dhrymes (1994). Productivity dynamics: U.s. manufacturing plants, 1972-1986. *Board of Governors of the Federal Reserve System (U.S.), Finance and Economics Discussion Series* 94, no. 1.

Bartelsman, Eric J. and Mark Doms (2000). Understanding productivity: Lessons from longitudinal microdata. *Journal of Economic Literature* 38, no. 3: 569–594.

Foster, Lucia, John Haltiwanger, and C.J. Krizan (2001). Aggregate productivity growth: Lessons from microeconomic evidence. In Charles R. Hulten, Edwin R. Dean, and Michael J. Harper (Eds.), *New Developments in Productivity Analysis*. Chicago: University of Chicago Press.

Gibrat, Robert (1931). *Les Inégalités Économiques; Applications: Aux Inégalités Des Richesses, À la Concentration Des Entreprises, Aux Populations Des Villes, Aux Statistiques Des Familles, Etc., D'une Loi Nouvelle, la Loi de L'effet Proportionnel*. Paris: Librairie du Recueil Sirey.

Gourieroux, Christian, Alain Monfort, and Eric Renault (1993). Indirect inference. *Journal of Applied Econometrics* 8, no. 0: S85–S118.

Griliches, Zvi and Haim Regev (1995). Firm productivity in israeli industry: 1979-1988. *Journal of Econometrics* 65, no. 1: 175–203.

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

Hall, George and John Rust (2003). Simulated minimum distance estimation of a model of optimal commodity price speculation with endogenously sampled prices. *Working Paper*.

Horowitz, Joel L. (1998). Bootstrap methods for covariance structures. *Journal of Human Resources 33*: 39–61.

Klette, Tor Jakob and Samuel Kortum (2004). Innovating firms and aggregate innovation. *Journal of Political Economy* 112, no. 5: 986–1018.

Lentz, Rasmus and Dale T. Mortensen (2005). Productivity growth and worker reallocation. *International Economic Review* 46, no. 3: 731–751.

Liu, Lili and James R. Tybout (1996). Productivity growth in chile and colombia: The role of entry, exit, and learning. In Mark J. Roberts and James R. Tybout (Eds.), *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*, pp. 73–103. Oxford and New York: Oxford University Press for the World Bank.

Olley, Steven G. and Ariel Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64, no. 6: 1263–97.

Petrin, Amil and James Levinsohn (2005). Measuring aggregate productivity growth using plant-level data. *Working paper*.

Rosholm, Michael and Michael Svarer (2000). Wage, training, and job turnover in a search-matching model. *IZA Discussion Papers 223*.

Rossi-Hansberg, Esteban and Mark L.J. Wright (2005). Firm size dynamics in the aggregate economy. *NBER Working Paper Series 11261*.

Scarpetta, Stefano, Philip Hemmings, Thierry Tressel, and Jaejoon Woo (2002). The role of policy and institutions for productivity and firm dynamics: Evidence from micro and industry data. *OECD Economics Department Working Paper 329*.

Sutton, John (1997). Gibrat's legacy. *Journal of Economic Literature* 35, no. 1: 40–59.

Tybout, James R. (1996). Heterogeneity and productivity growth: Assessing the evidence. In Mark J. Roberts and James R. Tybout (Eds.), *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*, pp. 43–72. Oxford and New York: Oxford University Press for the World Bank.