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: October 6, 2006

Abstract

Evidence suggests that productivity differences across firms are large and persistent, and worker reallocation is an important source of aggregate productivity growth. The purpose of the paper is to estimate the structure of an equilibrium model of growth through innovation that explains these facts while providing a theoretical foundation for an aggregate growth decomposition analysis. We argue that the empirical growth decomposition literature is meaningless for a steady state stochastic growth model such as ours. 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 than includes information on value added, employment, and wages. The model's fit is good and the structural parameter estimates imply that more productive firms grow faster and consequently crowd out less productive firms in steady state. This effect accounts for 58% of overall growth.

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, Robert Hall, John Kennan, Samuel Kortum, 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 as well as from exiting to entering 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] Bartelsman and Doms (2000) make the additional conclusion that the empirical growth decomposition literature shows evidence of a strong contribution to growth from reallocation.

In this paper, we take a position on the reasons for the role of reallocation by estimating a structural model of economic growth through product innovation using firm panel data.[2] As the 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, firms are monopoly suppliers of differentiated intermediate products that serve as inputs in the production of a final consumption good. More productive or higher quality intermediate products are introduced from time to time as the outcome of R&D investment by both existing firms and new entrants. 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 empirical growth decomposition approach is not meaningful in a stochastic steady state model as ours. As a simple illustration of the empirical decomposition approach, consider the

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

[2] To our knowledge, ours is the first attempt to do this.

following derivation. Let

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

represent an aggregate index of average productivity, say a measure of output per worker, where i is a firm index and t represent a particular time period, s_{it} denotes the firm i share of inputs in period t and p_{it} is input productivity in firm i in period t. The change in aggregate productivity can be written as

$$\begin{aligned} \Delta P_t &= \sum_{i \in I} s_{it} p_{it} - \sum_{i \in I} s_{it-1} p_{it-1} \\ &= \sum_{i \in I} s_{it-1} \Delta p_{it} + \sum_{i \in I} \Delta s_{it} p_{it-1} + \sum_{i \in I} \Delta s_{it} \Delta p_{it} \end{aligned}$$

where $\Delta x_t = x_t - x_{t-1}$ is the difference operator. The firm term can be interpreted as the within firm growth in productivity. The second term, the effect of changing firm shares, is generally interpreted as the effect of reallocation across firms. Finally, the last cross term is more difficult to interpret. On the face of it, the second term is positive if the firms that gain employment share in any period tend be the more productive as one would expect.

Although the interpretations of the sum of the second and third terms seem obvious, they have no meaning in an equilibrium model of the type estimated in this paper. Indeed, the sum of the second and third terms, $\sum_{i \in I} \Delta s_{it} p_{it}$, is zero in any structural equilibrium model of the type studied in this paper for the following reason. First, productivity in the model is the same for all firms of the same type by definition of type. Second, 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

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

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 I, the second by the fact that the firms of the same type have the same productivity at every date, and the last by the fact that the average share per firm of each type is constant ($s^*_{jt} = s^*_j$ for all t) in a steady state equilibrium. The expression in (1) corresponds to the first term in the Baily, Hulton, and Campbell (1992) index.

A naive interpretation of $\sum_{i\in I} \Delta s_{it} p_{it}$ as the gross effect of reallocating resources across firms is incorrect because gains in employment share are just 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. However, there will be important dynamic selection effects within any entry cohort composed of different types of firms that reflect the contribution of reallocation to aggregate growth. Specifically, resource reallocation contributes to growth when the firms of the more productive types grow faster within a given entry cohort. We will refer to this channel of growth as the selection effect.

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 that it does not necessarily measure the change in welfare that arises from additional final output holding primary inputs constant. In their view, the traditional "Solow residual," adapted to allow for market imperfections is the correct measure. Furthermore, it is the first component of the BHC index, which is then the same conclusion that we reach in equation (1). Indeed, their argument is valid for our structural model.

In our earlier paper, Lentz and Mortensen (2005), we establish the existence of a general equilibrium solution to the model. In the current 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 simulated moments. Providing a good fit to data, the model is estimated on among other moments the relationship between firm size and firm growth which is slightly negative in the data. The model satisfies a theoretical version of Gibrat's law conditional on type, but nevertheless replicates the negative relationship between size and growth found in data. The model is also estimated to fit the growth decomposition pioneered by Foster, Haltiwanger, and Krizan (2001) found in our data. According to the decomposition, all growth in the data can be attributed to the first term in the Baily, Hulton, and Campbell (1992) index which is the growth contribution from firm level productivity changes holding shares constant. Consequently, the decomposition says that the contribution from gross reallocation is zero. These are in fact exactly the numbers predicted in our discussion of the empirical decomposition approach.

We perform a structural growth decomposition on the estimated model and quantify in particular the net entry and the selection effect contributions to growth. The selection effect contributes positively to growth when more productive firms grow faster following entry than the less productive firms in their cohort. In this case, workers will reallocate to the more productive firms which will employ a greater share of productive resources in steady state relative to their share at entry.

The model estimate implies that net entry accounts for 18% of aggregate productivity growth while 58% can be attributed to the selection effect. Although the estimated model can explain the empirical Baily, Hulton, and Campbell (1992) decomposition of productivity growth in our model, in fact all of the observed growth can be attributed to reallocation.

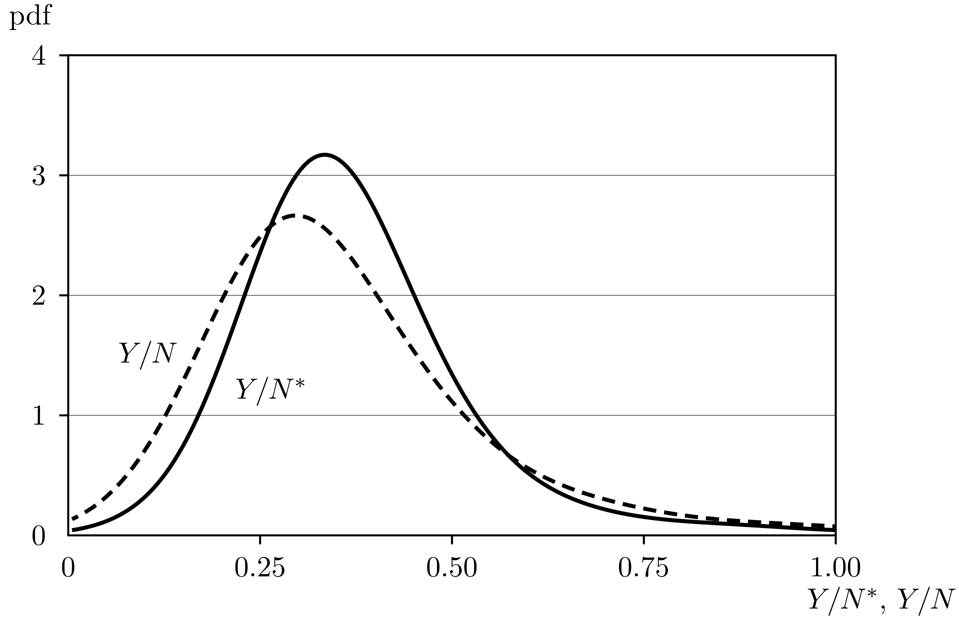
2 Danish Firm Data

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

[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.

is value added per worker (Y/N) while the second is valued added per unit of quality adjusted 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

$$w = \frac{\sum_j W_j}{\sum_j N_j} \tag{2}$$

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

[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

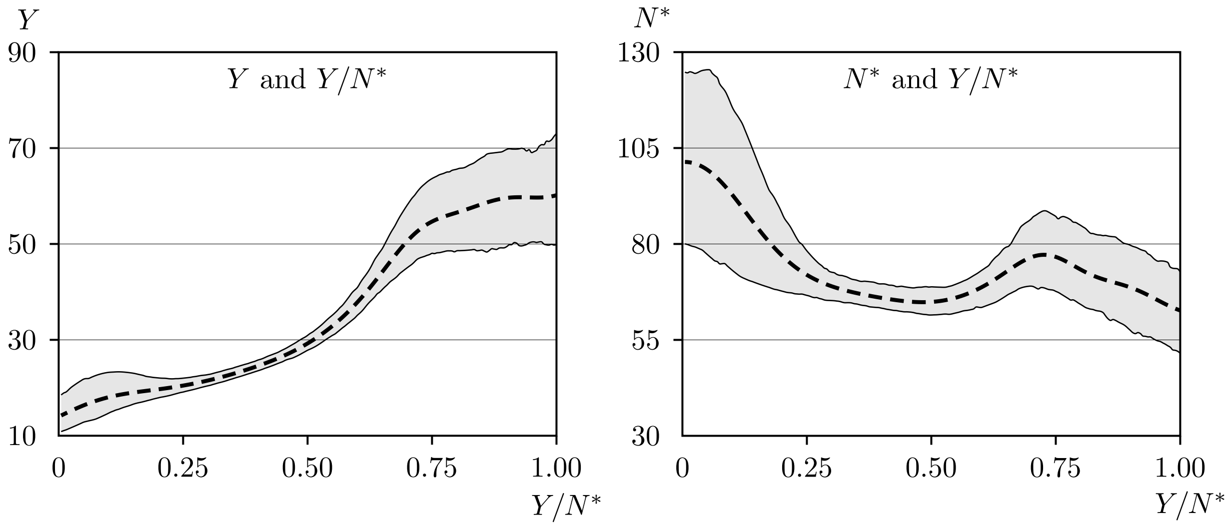
distributions are significantly dispersed and skewed to the right. In the case of the adjusted measure of productivity, the 5$^{\text{th}}$ percentile is roughly half the mode while the 95$^{\text{th}}$ percentile is approximately twice as large are the mode. The range between the two represents a four 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. Non-parametric regressions of value added and employment on the productivity measure are illustrated in Figure 2. The top and bottom curves in the figures represent a 90% confidence interval for the relationship. The positive relationship between value added and labor productivity is highly significant and no such positive relationship exists between labor force size and 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

Figure 2: Observed firm productivity and size correlation (1992).



Note: Value added (Y) measured in 1 million DKK. Labor force size (N^*) measured in efficiency units. Shaded areas are 90% confidence bounds.

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

$$U_t = \int_t^{\infty} \ln C_s e^{-r(s-t)} ds \tag{3}$$

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 $E_t = 1$ 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 large number of different inputs, N, and consumption is determined by the constant returns CES production function

$$C_t = \left[\sum_{j=1}^{N} Z\left(j\right) \left(A_t(j) x_t(j)\right)^{\rho} \right]^{\frac{1}{\rho}}, \rho \leq 1 \tag{4}$$

where $x_t(j)$ is the quantity of input $j \in [1, N]$ at time t and $A_t(j)$ is the productivity of input j at time t. $Z\left(j\right)$ 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,

$$A_t(j) = \prod_{i=1}^{J_t(j)} q_i(j), \tag{5}$$

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. 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.

A Cobb-Douglas ($\rho = 0$) production function specification is assumed for the purpose of estimation. In this case, intermediate product demand is,

$$x(j) = \frac{Z(j)}{p(j)}. \tag{6}$$

The general CES case is discussed in the appendix.

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 limited by the ability of suppliers of previous versions to provide a substitute. In Nash-Bertrand equilibrium, any innovator takes over the market for its good type by setting the price just below that at which consumers are indifferent between the more productive intermediate good supplied by the innovator and an alternative supplied by the previous supplier. The price charged is the product of relative productivity of the current most productive version of the product and the marginal cost of production.

The output of any intermediate good is proportional to labor input. However, the firm must also pay a fixed non-labor costs, $K(j) = \kappa Z(j)$ for complementary inputs, where κ is the non-labor cost share. Labor productivity is the same across all inputs and is set equal to unity without loss of generality. As the current limit price that can be charged yields zero profit to the nearest competitor, $p(j) = wq(j)/(1-\kappa)$. Thus, the supplier in market j charges a markup equal to the size of the improvement in the products productivity, $q(j)$, relative to the previous version supplied to the market.

The demand expression in equation (6) implies $p(j)x(j) = Z(j)$. Therefore, product j output and labor demand is,

$$x(j) = \frac{Z(j)}{p(j)} = \frac{(1-\kappa)Z(j)}{wq(j)}, \tag{7}$$

and the gross profit rate associated with supplying the good is

$$\pi(q(j)) = \frac{\Pi(q(j))}{Z(j)} = \frac{p(q(j))x(j) - wx(j) - K(j)}{Z(j)} = (1-\kappa)\left(1 - q(j)^{-1}\right). \tag{8}$$

It follow that $0 < \pi(q(j)) < 1-\kappa$ and that the profit rate is strictly increasing and concave in the productivity improvement $q(j)$. The labor saving nature of any productivity improvement is implicit in the fact that labor required to produce a product is decreasing in $q(j)$.

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]

By assumption firms cannot and have no incentive to 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 varies across products according to $Z(j)$, firms face

[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

demand uncertainty for a new innovation resolved only when the product type of an innovation is realized. Denote by $G(\cdot)$ the cumulative distribution function of $Z(j)$ across products. The demand for an innovation is then determined as a realization of the random variable $\tilde{Z} \sim G(\cdot)$. Denote by $Z \equiv E\big[\tilde{Z}\big]$. $\tilde{Z}$ and $\tilde{q}_\tau$ 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 $\tilde{Z}^k = \{\tilde{Z}_1, \ldots, \tilde{Z}_k\}$. Given such a state, the value of a type τ firm is accordingly given by,

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

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. Finally, the last term represents the expected capital loss associated with the possibility that one among the existing product lines (chosen at random) will be destroyed.

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

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

where,

$$\Psi_\tau = \max_{\gamma \geq 0} \frac{\gamma \bar{\pi}_\tau / \left(r+\delta\right) - w\hat{c}\left(\gamma\right)}{r+\delta-\gamma},$$

and $\bar{\pi}_\tau = E\pi\left(\tilde{q}_\tau\right)$ and $\hat{c}\left(\gamma\right) \equiv c\left(\gamma\right)/Z$. It then follows directly from (9) that the firm's optimal choice of creation rate, γ_τ, satisfies,

$$w\hat{c}'\left(\gamma_\tau\right) = \frac{\bar{\pi}_\tau}{r+\delta} + \Psi_\tau = v_\tau \equiv \max_{\gamma \geq 0} \frac{\bar{\pi}_\tau - w\tilde{c}\left(\gamma\right)}{r+\delta-\gamma} \tag{11}$$

of industries. However, if there are a large number of product variants supplied by each industry, then it is less objectionable. Later we show that similar results are obtained when estimating the model within broadly defined industries.

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 $\tau' \geq \tau$.

3.3 Firm Entry and Labor Market Clearing

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

$$wc'\left(\frac{\eta}{m}\right) = \sum_\tau \nu_\tau \phi_\tau = \sum_\tau \max_{\gamma \geq 0} \frac{\bar{\pi}_\tau - w\hat{c}(\gamma)}{r + \delta - \gamma} \phi_\tau \tag{12}$$

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

There is a fixed measure of available workers, denoted by L, seeking employment at any positive wage. In equilibrium, these are allocated across production and R&D activities, those performed by both incumbent firms and potential entrants. Since the average number of workers employed by a type τ firm for production purposes per product is $E[x_\tau] = E[\tilde{Z}_\tau(1-\kappa)/w\tilde{q}_\tau] = Z(1-\kappa-\bar{\pi}_\tau)/w$ from equations (7) and (8), the total number demanded for production activity by firms of type τ with k products is $L_\tau^x(k) = kZ(1-\kappa-\bar{\pi}_\tau)/w > 0$. The number of R&D workers employed by incumbent firms of type τ with k products is $L_\tau^R(k) = kc(\gamma_\tau)$. Because each potential entrant innovates at frequency η/m, the aggregate number of workers engaged by all m in R&D is $L_E = mc(\eta/m)$. Hence, the equilibrium wage satisfies the labor market clearing condition,

$$\begin{aligned} L &= \sum_\tau \sum_{k=1}^{\infty} \left[L_\tau^x(k) + L_\tau^R(k)\right] M_\tau(k) + L_E \\ &= \sum_\tau \left(\frac{Z}{w}(1-\kappa-\bar{\pi}_\tau) + c(\gamma_\tau)\right) \sum_{k=1}^{\infty} kM_\tau(k) + mc\left(\frac{\eta}{m}\right) \end{aligned} \tag{13}$$

where $M_\tau(k)$ represents the mass of firms of type τ that supply k products.

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

$$M_\tau(k) = \frac{k-1}{k}\frac{\gamma_\tau}{\delta}M_\tau(k-1) = \frac{\eta\phi_\tau}{\delta k}\left(\frac{\gamma_\tau}{\delta}\right)^{k-1} \tag{14}$$

by induction.

The size distribution of firms conditional on type can be derived using equation (14). Specifically, the total firm mass of type τ is

$$\begin{aligned} M_\tau &= \sum_{k=1}^{\infty} M_\tau(k) = \frac{\phi_\tau\eta}{\delta}\sum_{k=1}^{\infty}\frac{1}{k}\left(\frac{\gamma_\tau}{\delta}\right)^{k-1} \\ &= \frac{\eta}{\delta}\ln\left(\frac{\delta}{\delta-\gamma_\tau}\right)\frac{\delta\phi_\tau}{\gamma_\tau}. \end{aligned} \tag{15}$$

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

$$\frac{M_\tau(k)}{M_\tau} = \frac{\frac{1}{k}\left(\frac{\gamma_\tau}{\delta}\right)^k}{\ln\left(\frac{\delta}{\delta-\gamma_\tau}\right)}. \tag{16}$$

Equation (16) 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 (16), the mean of the type conditional firm size distribution is,

$$E\left[\tilde{k}_\tau\right] = \sum_{k=1}^{\infty} \frac{kM_\tau(k)}{M_\tau} = \frac{\frac{\gamma_\tau}{\delta-\gamma_\tau}}{\ln\left(\frac{\delta}{\delta-\gamma_\tau}\right)}, \tag{17}$$

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

$$K_\tau = \sum_{k=1}^{\infty} kM_\tau(k) = \frac{\eta\phi_\tau}{\delta-\gamma_\tau}. \tag{18}$$

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,

$$\frac{\partial E\left[\tilde{k}_\tau\right]}{\partial\gamma_\tau} = \left(\frac{(1+a_\tau)\ln(1+a_\tau)-a_\tau}{(1+a_\tau)\ln^2(1+a_\tau)}\right)\frac{1+a_\tau}{\delta-\gamma_\tau} > 0 \tag{19}$$

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

Although more profitable firms supply more products, total expected employment, $nZE\left[\tilde{k}_\tau\right]$ where $n = (1-\kappa-\bar{\pi}_\tau)/w + \hat{c}(\gamma_\tau)$, need not increase with $\bar{\pi}_\tau$ in general and decreases with $\bar{\pi}_\tau$ if innovation is not related to profitability because innovation is labor saving. 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.

Finally, the rate of creative-destruction is the sum of the entry rate and the aggregate creation rates of all the incumbents given that the total mass of products is fixed. Because the new product arrival rate of a firm of type τ with k products is $\gamma_\tau k$ and the measure of such firms is $M_\tau(k)$,

$$\delta = \eta + \sum_\tau\sum_{k=1}^{\infty}\gamma_\tau kM_\tau(k) = \eta + \sum_\tau \gamma_\tau K_\tau. \tag{20}$$

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

3.5 Equilibrium

Definition A *steady state market equilibrium* is a triple composed of a labor market clearing wage w, entry rate η, and creative destruction rate δ together with an optimal creation rate γ_τ and a steady state size distribution $M_\tau(k)$ for each type that satisfy equations (11), (12), (13), (14), and (20) provided that $\gamma_\tau < \delta$, for every τ in the support of the entry distribution.

Proposition If the cost of innovation, $c(\gamma)$, is strictly convex and $c'(0) = c(0) = 0$, then a steady state market equilibrium with positive entry exists. In the case of a single firm type, there is only one.

Proof. See Lentz and Mortensen (2005).[8] ■

4 Estimation

If the ability to create more 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 sources of aggregate productivity growth because faster growing firms create more improved products.

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 $\gamma'(\pi) > 0$, then (19) 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, equal to

[8] Although the cost of entry is linear in the paper cited while the cost is convex here, the principal argument holds in this case as well.

$[(1-\pi)/w + c(\gamma(\pi))]\, E[k]$, need not increase with π in general and decreases with π 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.

Table 2 presents a set of distribution moments with standard deviations in parenthesis. The standard deviations are obtained through bootstrapping. 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.

Table 3 contains the dynamic moments used in the estimation.[9] The moments relating to firm growth rates ($\Delta Y/Y$) include firm death, so specifically an exiting firm will contribute to the statistic with a -1 observation. There is a negative correlation between firm size and firm growth. Should one exclude firm deaths from the growth statistic, one will obtain 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

[9] Violating consistency, the dynamic moments in the 1992 column reflect changes between 1992 and 1993 whereas the 1997 column reflects changes between 1996 and 1997.

Table 2: Distribution Moments (std dev in parenthesis)

	1992	1997		1992	1997
$E[Y]$	$26,277.26$ (747.00)	$31,860.85$ $(1,031.25)$	$E\left[\frac{Y}{N^*}\right]$	384.40 (2.91)	432.12 (5.10)
$Med[Y]$	$13,471.00$ (211.35)	$16,432.10$ (329.77)	$Med\left[\frac{Y}{N^*}\right]$	347.08 (1.83)	377.09 (2.14)
$Std[Y]$	$52,798.52$ $(5,663.63)$	$64,129.07$ $(7,742.51)$	$Std\left[\frac{Y}{N^*}\right]$	205.09 (19.63)	305.35 (42.50)
$E[W]$	$13,294.48$ (457.47)	$15,705.09$ (609.60)	$Cor[Y,W]$	0.85 (0.04)	0.86 (0.04)
$Med[W]$	$7,229.70$ (92.75)	$8,670.28$ (154.90)	$Cor\left[\frac{Y}{N^*},Y\right]$	0.20 (0.04)	0.14 (0.04)
$Std[W]$	$30,616.94$ $(6,751.09)$	$35,560.60$ $(8,138.66)$	$Cor\left[\frac{Y}{N^*},N^*\right]$	-0.02 (0.01)	-0.03 (0.01)

and size, the same will be true in the model simulations. Firm productivity exhibits persistence and mean reversion.

In addition to the moments in Table 2 and 3, the model will also be asked to explain 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). The decomposition takes the form,

$$\begin{aligned}\Delta P_t \;=\; & \sum_{e\in C} s_{et-1}\Delta p_{et} + \sum_{e\in C}\left(p_{et-1}-P_{t-1}\right)\Delta s_{et} + \sum_{e\in C}\Delta p_{et}\Delta s_{et} + \sum_{e\in N}\left(p_{et}-P_{t-1}\right)s_{et} - \\ & \sum_{e\in X}\left(p_{et-1}-P_{t-1}\right)s_{et-1}, \end{aligned} \tag{21}$$

where $P_t = \sum_e s_{et}p_{et}$, $p_{et} = Y_{et}/N_{et}$, and $s_{et} = N_{et}/N_t$. Thus, (21) will be used to decompose 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. The within component is interpreted as growth in the productivity measure due to productivity improvements by incumbents, the between component is designed 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 decomposition shares in the data are shown in Table 4. As mentioned, the sample in this paper

Table 3: Dynamic Moments (std dev in parenthesis)

	1992	1997
Survivors	$4,872.000$	$3,628.000$ (32.130)
$Cor\left[\frac{Y}{N^*}, \frac{Y_{+1}}{N^*_{+1}}\right]$	0.476 (0.088)	0.550 (0.091)
$Cor\left[\frac{Y}{N^*}, \Delta\frac{Y}{N^*}\right]$	-0.227 (0.103)	-0.193 (0.057)
$Cor\left[\frac{Y}{N^*}, \frac{\Delta Y}{Y}\right]$	-0.120 (0.016)	
$Cor\left[\frac{Y}{N^*}, \frac{\Delta N^*}{N^*}\right]$	0.119 (0.032)	
$E[\frac{\Delta Y}{Y}]$	-0.029 (0.008)	
$Std\left[\frac{\Delta Y}{Y}\right]$	0.550 (0.067)	
$Cor\left[\frac{\Delta Y}{Y}, Y\right]$	-0.061 (0.012)	

does not include entry, so there is no entry share in the decomposition and the decomposition shares in Table 4. 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 4.4.2.

The decomposition provides additional information on dynamics in the data and is therefore valuable for identification purposes. But it is also a useful method of directly relating the model to the empirical growth decomposition literature.

4.2 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, $\psi_i = \left\{\psi_{i1,\ldots,}\psi_{iT}\right\}$ and finally, $\psi = \{\psi_1, \ldots, \psi_I\}$.

Simulated minimum distance estimators, as described in for example Gourieroux, Monfort, and Renault (1993), Hall and Rust (2003), and Alvarez, Browning, and Ejrnæs (2001), are computed as follows: First, define a vector of auxiliary data parameters, $\Gamma(\psi)$. The vector consists of all the items in Tables 2 and 3 except the number of survivors in 1992 and three of the moments in

Table 4: Y/N 1992 to 1997 Growth Decomposition. Std Dev in parentheses.

	Growth Shares
Within	1.015 (0.146)
Between	0.453 (0.112)
Cross	-0.551 (0.196)
Exit	0.084 (0.066)

table 4. Thus, $\Gamma(\psi)$ has length 37. Second, $\psi^s(\omega)$ is simulated from the model for a given set of model parameters ω. The model simulation is initialized by assuming that the economy is in steady state in the first year and consequently that firm observations are distributed according to the ω-implied steady state distribution. Alternatively, one can initialize the simulation according to the observed data in the first year, $\{\psi_{11}, \ldots, \psi_{1I}\}$. The assumption that the economy is initially in steady state provides additional identification in that $\{\psi_{11}, \ldots, \psi_{1I}\}$ can be compared to the model-implied steady state distribution $\{\psi^s_{11}(\omega), \ldots, \psi^s_{1I}(\omega)\}$. Initializing by the observed state in the first period has the additional problem that initialization is on each firm's number of products which is not directly observed but must be inferred.

The simulated auxiliary parameters are then given by,

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

where S is the number of simulation repetitions. The estimator is then the choice of parameters that minimizes the weighted distance between the data and simulated auxiliary parameters,

$$\hat{\omega} = \arg\min_{\omega\in\Omega}\big(\Gamma^s(\omega) - \Gamma(\psi)\big)' A^{-1}\big(\Gamma^s(\omega) - \Gamma(\psi)\big), \tag{22}$$

where A is some positive definite matrix. If A is the identity matrix, $\hat{\omega}$ is the equally weighted minimum distance estimator (EWMD). If A is the covariance matrix of the data moments $\Gamma(\psi)$, $\hat{\omega}$ is the optimal minimum distance estimator (OMD). The OMD estimator is asymptotically more efficient than the EWMD estimator. However, Altonji and Segal (1996) show that the estimate of A as the second moment matrix of $\Gamma(\cdot)$ may suffer from serious small sample bias. Horowitz (1998) suggests a bootstrap estimator of A. The estimation in this paper adopts Horowitz's bootstrap estimator of the covariance matrix A.

In addition to the $\hat{\omega}$ estimator, the analysis also presents a bootstrap estimator as in Horowitz (1998). In each bootstrap repetition, a new set of data auxiliary parameters $\Gamma\left(\psi^b\right)$ 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 $\left[\Gamma\left(\psi^b\right)-\Gamma\left(\psi\right)\right]$ and the re-centered simulated auxiliary parameters $\left[\Gamma^s\left(\omega^b\right)-\Gamma^s\left(\hat{\omega}\right)\right]$,

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

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

The bootstrap estimator of the structural parameters is then the simple average of all the ω^b estimators,

$$\hat{\omega}^{bs}=\frac{1}{B}\sum_{b=1}^{B}\omega^b, \tag{23}$$

where B is the total number of bootstrap repetitions. In the estimation below, $B=500$ and $S=600$.

4.3 Model Simulation

The model simulation produces time paths for value added (Y), the wage sum (W), and labor force size (N) for a given number of firms. The firm type distribution is estimated non-parametrically as a 3-point discrete type distribution ϕ_τ. The type conditional productivity realization distributions are assumed to be three parameter Weibull distributions that share a common shape parameter $\theta_\beta^{q_\tau}$ and a point of origin as a common fraction between 1 and the mean of the distribution. Thus, the three productivity realization distributions are estimated with 5 parameters. The demand realization distribution $G(\cdot)$ is assumed to be a three parameter Weibull where θ_γ^Z is the origin,

θ_β^Z is the shape parameter, and θ_η^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,

$$Y_\tau\left(\tilde{q}^k, \tilde{Z}^k\right) = \sum_{i=1}^{k} \tilde{Z}_i, \tag{24}$$

and a wage bill of,

$$W_\tau\left(\tilde{q}^k, \tilde{Z}^k\right) = \sum_{i=1}^{k} \tilde{Z}_i\left(1 - \kappa - \pi\left(\tilde{q}_i\right)\right) + wkc\left(\gamma_\tau\right). \tag{25}$$

Equations (24) and (25) provide the foundation for the model simulation. The simulation is initialized by the assumption of steady state. The steady state firm type distribution is given by,

$$\phi_\tau^{ss} = \frac{\eta\phi_\tau \ln\left(\frac{\delta}{\delta-\gamma_\tau}\right)}{M\gamma_\tau},$$

where M is the total steady state mass of firms. Thus, a firm's type is drawn according to ϕ^{ss}. The firm's 1992 product line size is then determined according to the type conditional steady state distribution of $\tilde{k}_\tau$ as stated in (16),

$$\Pr\left(\tilde{k}_\tau = k_1\right) = \frac{M_\tau(k)}{M_\tau} = \frac{\frac{1}{k_1}\left(\frac{\gamma_\tau}{\delta}\right)^{k_1}}{\ln\left(\frac{\delta}{\delta-\gamma_\tau}\right)}. \tag{26}$$

With a given initial product size, simulation of the subsequent time path requires knowledge of the transition probability function $\Pr\left(\tilde{k}_{\tau,2} = k|k_1\right)$. Denote by $p_{\tau,k}\left(t\right)$ the probability of a type τ firm having product size k at time t. As shown in Klette and Kortum (2004), $p_{\tau,k}\left(t\right)$ evolves according to the ordinary differential equation system,

$$\begin{aligned} \dot{p}_{\tau,k}\left(t\right) &= \left(k-1\right)\gamma_\tau p_{\tau,k-1}\left(t\right) + \left(k+1\right)\delta p_{\tau,k+1}\left(t\right) - \left(\delta+\gamma_\tau\right)p_{\tau,k}\left(t\right),\ \forall k \geq 1 \\ \dot{p}_{\tau,0}\left(t\right) &= \delta p_{\tau,1}\left(t\right). \end{aligned} \tag{27}$$

Hence, with the initial condition,

$$p_{\tau,k}\left(0\right) = \begin{cases} 1 \text{ if } k = k_1 \\ 0 \text{ otherwise.} \end{cases} \tag{28}$$

one can determine $\Pr\left(\tilde{k}_{\tau,2} = k|k_1\right)$ by solving the differential equation system in (27) for $p_{\tau,k}\left(1\right)$. Solving for $p_{\tau,k}\left(1\right)$ involves setting an upper reflective barrier to bound the differential equation system. It has been set sufficiently high so as to avoid biasing the transition probabilities. Based on the transition probabilities $\Pr\left(\tilde{k}_{\tau,t+1} = k|k_t\right)$ one can then iteratively simulate product size paths for each firm. The procedure correctly captures the evolution of k_t but it does not identify the exact evolution of $\left(\Pi^{k_t}, Z^{k_t}\right)$. The evolution of $\left(\Pi^{k_t}, Z^{k_t}\right)$ is assumed to follow the net change in products.[10]

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 $\xi_Y \sim N\left(0, \sigma_Y^2\right)$ and $\xi_W \sim N\left(0, \sigma_W^2\right)$. 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.

To account for non-stationarity, the simulation allows for an exogenous growth factor in both value added and the wage bill, denoted as $\widehat{g}$. Since the equilibrium ratio Z/w is stationary given a fixed labor force size from the labor market clearing condition, equation (13), it is independent of the endogenous productivity improvement in the intermediate goods. Finally, the interest rate will be set at $r = .05$. The wage w is immediately identified as the average worker wage in the sample $w = 190.24$. Excluding w, the estimation identifies 17 parameters.

4.4 Estimation Results

The model parameter estimates are given in table 5.[11]

[10] Suppose firm i is simulated to lose one product in a given year. In this case, $\left(\Pi^{k_{it}}, Z^{k_{it}}\right)$ is updated by randomly eliminating one element from it. This assumes that the net loss of one product took place by the gross destruction of one product and zero gross creation. This is the most likely event by which the firm loses one product. However, the net loss could also come about by the gross destruction of two products and gross creation of one product during the year. In this case, $\left(\Pi^{k_{it}}, Z^{k_{it}}\right)$ should be updated by randomly eliminating two elements and adding one. There are in principle an infinite number of ways that the firm can loose one product over the year. The estimation consequently over-estimates the persistency of $\left(\Pi^{k_{it}}, Z^{k_{it}}\right)$. The bias will go to zero as the period length is reduced, though.

[11] Note that the estimate for $\theta_\gamma^{q_\tau}$, the origin of the type conditional quality realization Weibull distribution, is estimated using a single parameter. The estimate is constrained across the types as the fraction of the distance between 1 and the average quality realization.

Given the steady state equilibrium definition, one can infer the overall entry rate, η, and the measure of potential entrant, m.[12] The implied values of these parameters are also reported in Table 6. The average incumbent creation rate, $\bar{\gamma}$, is simply the difference between the entry rate and the destruction rate. It is seen that the estimates imply that more than half of all innovation comes from entrants.

Given the estimated steady state distribution of firms, ϕ_τ^{ss} and the other parameters of the model, one can infer the ex ante type distribution, ϕ_τ. It is seen that the higher productivity type firms choose higher creation rates and consequently grow to be over-represented in steady state relative to the ex ante type distribution. The selection effect is also seen in the product mass by type, K_τ, where the highest type is estimated to supply 33.2% of the products despite a representation at birth of only 9.1%. The consequences of these facts for aggregate growth are explored more fully below.

The overall creation and destruction rate is estimated at an annual rate of 8%. The average lifespan of a product is consequently about 12.5 years. The destruction rate is roughly consistent with evidence in Rosholm and Svarer (2000) that the worker flow from employment to unemployment is around 10% annually. The non-labor share is estimated at about 43%. The average labor share in the data is roughly 55% implying a modest average profit rate in the model. The exact profit implications are explored in more detail below.

Table 7 allows a comparison of the data moments and the simulated moments associated with the model parameter estimates.

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. Furthermore, firms of size 20 or large account for almost all of private employment.

The estimation explicitly includes a number of dynamic moments. In addition, it should be

[12]The formulas used to make the calculations are presented in the appendix.

Table 5: Parameter Estimates

	SMD Estimator	Bootstrap Estimator	Std dev
c_0/Z	9.2477	9.2467	0.5059
c_1	2.8890	2.8835	0.0153
κ	0.4269	0.4279	0.0035
δ	0.0799	0.0797	0.0010
$\hat{g}$	0.0154	0.0155	0.0012
σ_Y^2	0.0353	0.0356	0.0047
σ_W^2	0.0182	0.0193	0.0039
θ_γ^Z	6,344.8418	6,300.6200	361.5880
θ_η^Z	7,311.7865	7,494.4960	433.2505
θ_β^Z	0.6722	0.6772	0.0176
ϕ_τ^{ss}			
– type 1	0.4645	0.4366	0.0403
– type 2	0.3728	0.4044	0.0427
– type 3	0.1627	0.1590	0.0178
$\theta_\gamma^{q_\tau}$			
– type 1	1.0014	1.0011	0.0005
– type 2	1.0033	1.0022	0.0009
– type 3	1.1001	1.0840	0.0226
$\theta_\eta^{q_\tau}$			
– type 1	0.0029	0.0033	0.0019
– type 2	0.0076	0.0070	0.0030
– type 3	0.7936	0.9187	0.2365
$\theta_\beta^{q_\tau}$	0.4327	0.4432	0.0253

noted that since the estimation is performed on cross-section moments not just in 1992 but also in 1997 and because of the specific sampling procedure in the data, the estimation implicitly address dynamic features of the model. The trends in the moments over time are in part interpreted as a result of systematic selection bias due to creation rate heterogeneity across types.

Size Distributions As seen in figure 3, the model fits the size distributions very well, although the figure does not reveal that the model does 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.

The dispersion estimate is a result of a combination of the stochastic nature of the birth-death

Table 6: Inferred estimates (std dev in parentheses)

		Type 1	Type 2	Type 3
$\bar{\gamma}$	0.0348 (0.0014)			
η	0.0451 (0.0017)			
M	0.6968 (0.0148)			
m	1.4172 (0.0942)			
ϕ_τ		0.5146 (0.0441)	0.3944 (0.0452)	0.0910 (0.0115)
K_τ		0.3621 (0.0331)	0.3058 (0.0371)	0.3321 (0.0276)
γ_τ		0.0158 (0.0021)	0.0217 (0.0021)	0.0675 (0.0012)
ν_τ/Z		0.0428 (0.0205)	0.1073 (0.0301)	2.8404 (0.2419)
$\bar{\pi}_\tau$		0.0051 (0.0023)	0.0122 (0.0033)	0.2264 (0.0185)
$E[\ln \tilde{q}_\tau]$		0.0091 (0.0043)	0.0225 (0.0063)	0.7100 (0.0861)

process of products, the demand shock process, and to a lesser extend the measurement error processes. Model simulation without measurement error ($\sigma_Y^2 = \sigma_W^2 = 0$) yields a reduction in the 1992 value added standard deviation estimate from $31,108.56$ to $29,761.48$. A model simulation with zero variance in the demand realization ($Var[Z] = 0$) yields a reduction in the 1992 value added standard deviation estimate to $25,208.61$.

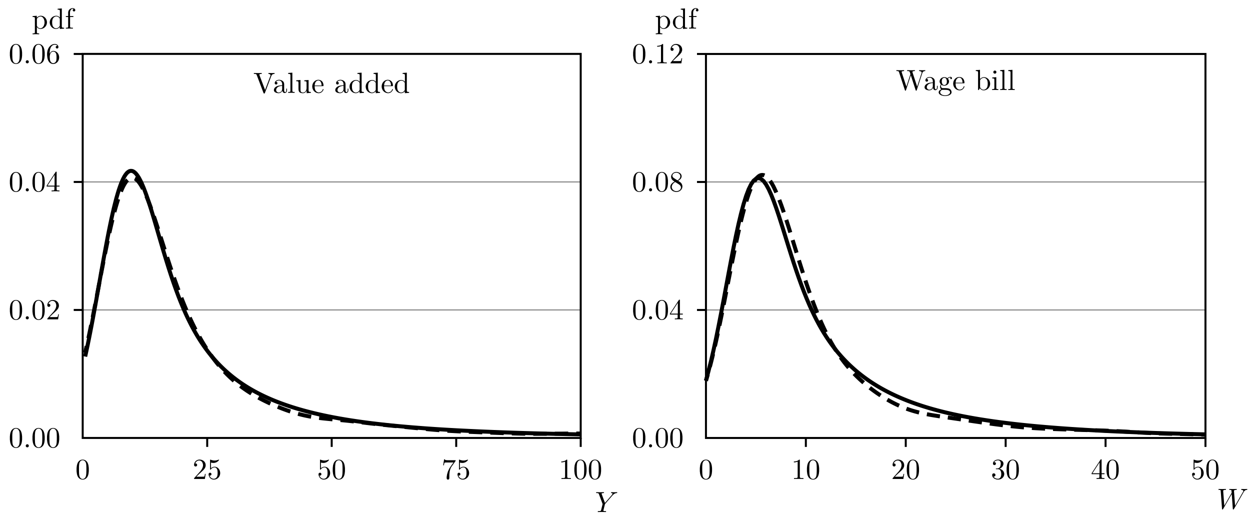
Productivity–Size Correlations Type heterogeneity and the variance of the type conditional productivity realizations play an important role in explaining the productivity – size correlations. Type heterogeneity provides the foundation for a positive correlation between productivity and output size through a greater product creation rate for firms that create higher quality innovations. The overall heterogeneity in product productivity realizations both through type heterogeneity and random productivity realizations within types explains the difference between the productivity – input size correlation and the productivity – output size correlation. Measurement error has the

Table 7: Model fit

	Data		Estimated model	
	1992	1997	1992	1997
Survivors	$4,872.000$	$3,628.000$	$4,872.000$	$3,565.850$
$E[Y]$	$26,277.262$	$31,860.851$	$23,351.933$	$27,669.629$
$Med[Y]$	$13,472.813$	$16,432.098$	$13,215.248$	$15,324.092$
$Std[Y]$	$52,793.105$	$64,129.072$	$31,202.597$	$37,502.778$
$E[W]$	$13,294.479$	$15,705.087$	$11,785.798$	$13,703.361$
$Med[W]$	$7,231.813$	$8,670.279$	$7,138.942$	$8,241.911$
$Std[W]$	$30,613.801$	$35,560.602$	$13,537.309$	$15,939.896$
$E[Y/N^*]$	384.401	432.118	383.645	418.839
$Med[Y/N^*]$	348.148	375.739	347.084	377.091
$Std[Y/N^*]$	205.074	305.348	207.974	227.394
$Cor[Y,W]$	0.852	0.857	0.918	0.919
$Cor[Y/N^*,Y]$	0.198	0.143	0.181	0.194
$Cor[Y/N^*,N^*]$	-0.018	-0.026	-0.028	-0.015
$Cor\left[Y/N^*,Y_{+1}/N^*_{+1}\right]$	0.476	0.550	0.743	0.742
$Cor[Y/N^*,\Delta(Y/N^*)]$	-0.227	-0.193	-0.336	-0.333
$Cor[Y/N^*,\Delta Y/Y]$	-0.120	–	-0.113	–
$Cor[Y/N^*,\Delta N^*/N^*]$	0.119	–	0.120	–
$E[\Delta Y/Y]$	-0.029	–	0.007	–
$Std[\Delta Y/Y]$	0.550	–	0.511	–
$Cor[\Delta Y/Y,Y]$	-0.061	–	-0.054	–
Growth decomp.				
– Within	1.015	–	0.873	–
– Between	0.453	–	0.323	–
– Cross	-0.551	–	-0.349	–
– Exit	0.084	–	0.153	–

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 $corr(Y/N,Y) = 0.147$ and $corr(Y/N,N) = 0.001$. 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.

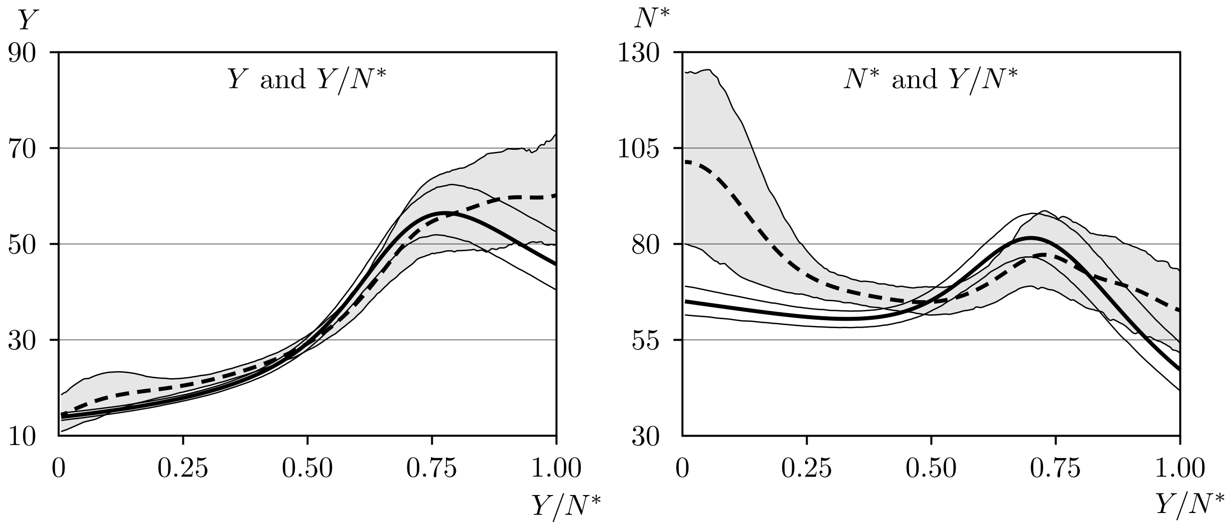
Figure 3: Size distribution fit, 1992



Note: Value added (Y) and wage bill (W) measured in 1 million DKK. Estimated model in solid pen. Data in dashed pen.

Right-Shift of Size Distributions Notice that the model successfully captures the right shift of the Y and W distributions of survivors from 1992 to 1997. There are three effects that contribute to the right shift: Generally, since the sampling eliminates the flow in of entrants, the model predicts a general decrease in mass of firms of all product sizes and types, $M_k(\pi)$, since all firms face an overall negative product growth rate. However, since entrants are assumed to flow in from the lower end of the size distribution, the reduction in mass is relatively stronger at the lower end and consequently the size distribution of survivors will begin to place relatively more weight on the upper end as time passes. Thus, the model predicts that the use of an unbalanced panel that excludes entry will itself produce a right shift of the distributions since entrants are assumed to enter as small firms. Second, the positive exogenous growth estimate directly predicts a right shift of the Y and W distributions. The third effect comes from type heterogeneity. In steady state, larger firms will over-represent firms with higher creation rates and small firms will over-represent firms with low creation rates. Thus, smaller firms face greater net product destruction than large firms. In the absence of entry, the negative correlation between size and net product destruction rate will in isolation produce a right shift of the Y and W distributions over time. Hence, this effect is also a consequence of the use of an unbalanced panel that excludes entry, but is separate

Figure 4: Firm productivity and size, 1992 (data and simulation).

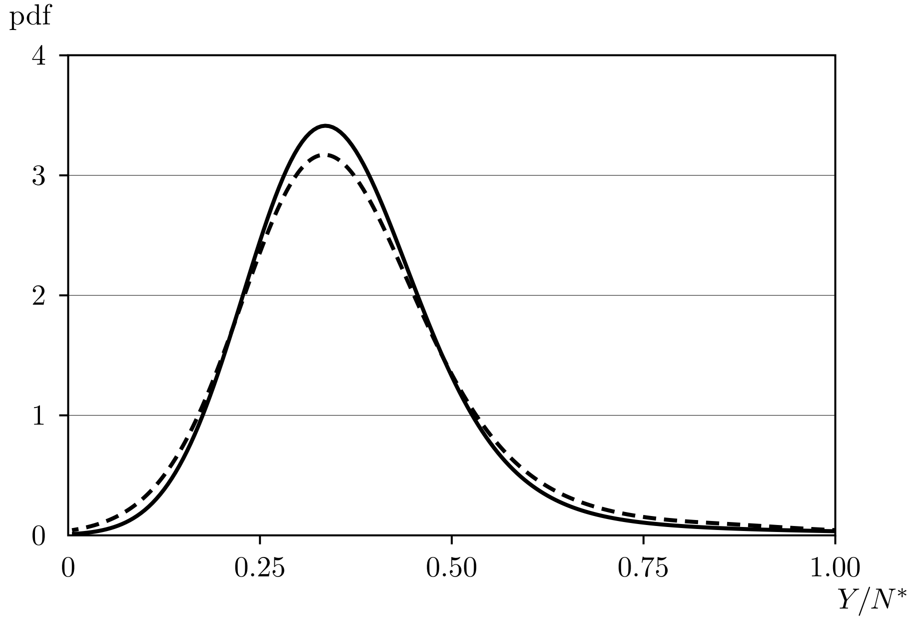


Note: Value added (Y) measured in 1 million DKK. Labor force size (N^*) measured in efficiency units. Estimated model point estimate and 90% confidence bounds drawn in solid pen. Data in dashed pen. Shaded areas are 90% confidence bounds on data.

from the first explanation which is not a result of destruction rate heterogeneity.

Value Added per Worker Distribution The distribution of firm labor productivity Y/N is explained primarily by type heterogeneity, the capital share, the structural noise processes, and measurement error. The mean level of value added per worker is closely linked to the estimate of κ given that profit levels are modest. The dispersion in Y/N across firms is explained primarily and in roughly equal parts by type heterogeneity and within type variance in productivity realizations. Measurement error adds to the dispersion measure, but to a smaller extend. Simulation without measurement error ($\sigma_Y^2 = \sigma_W^2 = 0$) yields a reduction in the 1992 Y/N standard deviation measure from 207.86 to 178.21. 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. The firm's innovation labor demand is unaffected by demand realizations so a greater demand realization will result in increased value added per worker. The effects

Figure 5: 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.

are secondary, and demand side shocks have little impact on value added per worker dispersion in the estimated model.

The right shift of the value added per worker distribution from 1992 to 1997 is explained as a combination of the exogenous growth estimate and the selection effect in that more productive firms have lower exit hazard rates. However, given the relatively low estimate of overall creative destruction, the primary effect is through the growth estimate.

Value Added per Worker Persistence and Mean Reversion The persistence in firm labor productivity $cor\left(\frac{Y}{N}, \frac{Y_{+1}}{N_{+1}}\right)$ 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 estimate of the relatively low level of overall creation and destruction implies that both the supply and the demand shock processes are fairly permanent and they turn out to contribute very little to the explanation of the persistence and mean reversion of value added per worker. Thus, it is left to the transitory nature of the measurement error processes to explain the exact persistence and mean reversion of the value added per worker measures. Simulating the model without measurement error

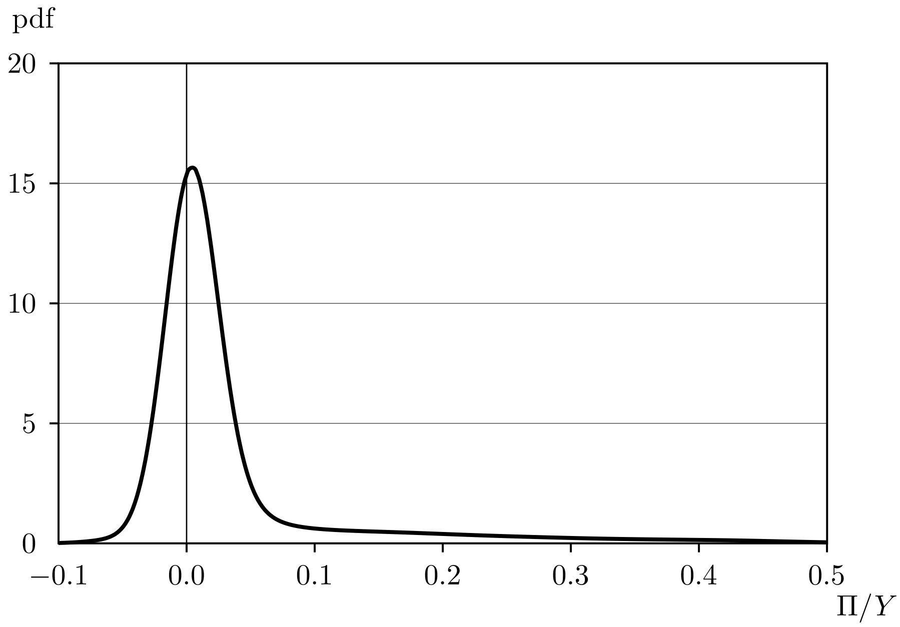
$(\sigma_Y^2 = \sigma_W^2 = 0)$ results in 1992 persistence and mean reversion moments of $cor\left(\frac{Y}{N}, \frac{Y_{+1}}{N_{+1}}\right) \approx .97$ and $cor\left(\frac{Y}{N}, \Delta\frac{Y}{N}\right) \approx -.06$. So, without the measurement error, the model implies a high level of value added per worker persistence, which is ultimately reduced by the measurement error components. 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.

Profit rates Profit rates as measured by total per period profits relative to value added vary both across and within types. Firm types that create products of higher quality have higher profit rates due to larger markups. However, profit rates vary considerably within type as a result of demand and productivity realization variance. Figure 6 shows the estimated profit rate distribution in steady state. Profits are generally estimated to be modest at an average of 3.3%. The median profit rate is 0.3%. Negative profit rates can occur in the case where past demand and productivity realizations are sufficiently low relative to the firm's innovation efforts so as to generate a negative cash flow.

4.4.1 Growth Rate and Size

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 is about issues 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 7, in the current sample of firms where the growth rate – size regression includes firm exits, one still obtains

Figure 6: Profit rate distribution, 1992



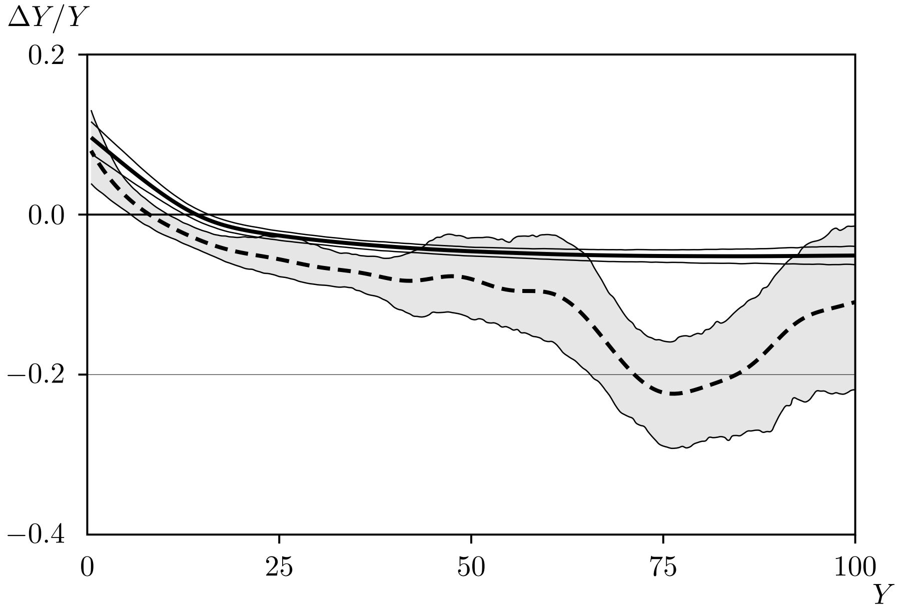
Note: Profit rate measured as profits relative to value added.

a negative relationship.

At a theoretical level, the model satisfies Gibrat's law by assumption in the sense that each firm's net innovation rate is size independent. But two opposing effects will impact the unconditional size–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 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 7.[13] Note that the growth statistics include firm death. If firm deaths are excluded and the statistic is calculated only on

[13]Figure 7 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.

Figure 7: Kernel Regression of Firm Growth Rate and Size (1992).



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

Table 8: Growth dynamics and counterfactuals

	Data	Point Estimate	$\sigma_Y^2 = 0$ $\sigma_W^2 = 0$	$Var(Z) = 0$ $\sigma_Y^2 = 0$ $\sigma_W^2 = 0$
$E\,[\Delta Y/Y]$	−0.029	0.008	−0.027	−0.037
$Std\,[\Delta Y/Y]$	0.550	0.514	0.406	0.297
$Cor\,[\Delta Y/Y, Y]$	−0.061	−0.054	−0.003	0.026

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.

Although Gibrat's law holds by design abstracting from shocks, the estimated version nevertheless exhibits a negative relationship between observed firm size and growth rate. As shown in Table 8, the model explains the negative relationship found in data through demand fluctuations and measurement error. The demand shock process is quite permanent and the measurement error process is consequently a bit stronger. In fact, without the demand shock variation, the model will exhibit an even stronger negative correlation between size and growth because the group of small firms now has a stronger selection of negative transitory shocks. However, as can be seen from table 8, in the absence of measurement error, demand side shocks clearly contribute to a negative relationship between growth and size. A more transitory demand shock process will resemble the measurement error process in the model and consequently, as above, the introduction of such a process may reduce the measurement error variance estimate.[14] Gibrat's law may at one level simply be a statement about the observed relationship between firm size and growth, and its validity is in this sense an issue that can be settled through observations such as the one in Figure 7. However, we have interpreted Gibrat's law to be a statement about a more fundamental proportionality between size and the firm's growth process, specifically innovation. In this case, the structural estimation shows that observation of the relationship between firm growth and firm size is not enough to falsify the statement.

4.4.2 Y/N Growth Decomposition

With the introduction of longitudinal micro-level data sets, a large literature has emerged with the focus on firm level determinants of aggregate productivity growth. See Bartelsman and Doms (2000) for a review of the literature. Given the observation of extensive firm level productivity dispersion, one particular area of interest has been the contribution to aggregate productivity growth from resource reallocation. The discussion has been quantified through decompositions such as (21), where productivity has been defined either as value added per worker or firm TFP. In the estimation in this paper, we have used the value added per worker measure. It should be immediately clear that value added per worker is not perfectly related to actual productivity growth in our model, so we should at the outset expect some divergence between the reduced form decomposition in (21) and the decomposition implied by the model that we present in the following section. Furthermore, Petrin and Levinsohn (2005) point out problems of interpreting the reallocation components of the standard empirical decomposition.

In the estimation and in the data sample, entry is excluded and the decomposition consequently

[14] It is important to note that identification of the demand shock and measurement error processes comes from other aspects of the data as well such as dispersion in the size distribution and a number of the dynamic moments. If the Gibrat related moments are excluded from the estimation, the estimated model still exhibits a negative relationship between observed firm size and growth rate.

has no value added per worker growth contribution from entry. The first two columns of Table 9, presents the decomposition results from the data and the simulated steady state that excludes entry. The remaining three columns in the table presents the simulated steady state with entry for the actual point estimate and for the two counterfactuals where measurement error noise and demand shocks have been eliminated.

The steady state with entry simulates not only the dynamic evolution of the sample of incumbents, which is the sample that the estimation is based on, it also simulates the entry process implied by the steady state general equilibrium. The entry process is described in section 3.3. For simulation purposes, the size of the potential entrant pool in the estimated model is $4,872m/M = 9,909$. At any point in time, each of these potential entrants will enter according to entry rate $\gamma_0 = \eta/m = 0.0318$. The entry process is simulated to fit the one year observation frequency in the data. Thus, for each entrant who starts the year in the potential entrant pool, we calculate the transition probability that after 1 year the potential entrant has k products, $\Pr(k_e = k|\tau)$, where the type conditioning refers to the firm type realization at entry. The type realization is obviously unknown to the potential entrant prior to entry, but is subsequently of importance in terms of determining the birth-death process of product lines in the remainder of the year after entry. If $k_e > 0$ the firm is registered as an entrant with k_e products and the subsequent life of the entrant is simulated through the incumbent transition probability described in section 4.3.

The type τ conditional potential entrant transition probability, $\Pr(k_e = k|\tau)$, is calculated in a similar fashion to the incumbent transition probability as described in section 4.3. However, in this case, the differential equation system that describes the probability that the potential entrant has product size n at time t, takes the form,

$$\begin{aligned}
\dot{p}_e(t) &= -\gamma_0 p_e(t) \\
\dot{p}_{\tau,1}(t) &= \gamma_0 p_e(t) - (\delta + \gamma_\tau) p_{\tau,1}(t) \\
\dot{p}_{\tau,n}(t) &= (n-1)\gamma_\tau p_{\tau,n-1}(t) + (n+1)\delta p_{\tau,n+1}(t) - (\delta + \gamma_\tau) p_{\tau,n}(t), \ \forall n \geq 2 \\
\dot{p}_{\tau,0}(t) &= \delta p_{\tau,1}(t),
\end{aligned}$$

Table 9: FHK growth decomposition and counterfactuals.

	Data	Point Estimate	Steady state with entry: Point Estimate	Steady state with entry: $\sigma_Y^2 = 0$, $\sigma_W^2 = 0$	Steady state with entry: $Var[Z] = 0$, $\sigma_Y^2 = 0$, $\sigma_W^2 = 0$
Within	1.015	0.873	1.025	0.835	0.854
Between	0.453	0.323	0.261	0.082	0.100
Cross	−0.551	−0.349	−0.530	−0.160	−0.198
Exit	0.084	0.153	0.179	0.178	0.177
Entry	—	—	0.069	0.066	0.066

where the notation follows the notation in section 4.3 with the addition that $p_e(t)$ refers to the probability that the potential entrant is still a potential entrant at time t (and has product size 0). Given the initial condition $p_e(0) = 1$, the potential entrant transition probability is found by solving the above differential equation system for $p_e(1)$ and $p_{\tau,k}(1)$. Thus, the probability that the potential entrant will not have entered after one year is $p_e(1) + p_{\tau,0}(1)$. The latter term reflects the event that a firm enters but exits again before the year's end, in which case the firm is not included in the pool of entrants. It is also seen that the discrete observation frequency implies that entry with more than one product is a positive likelihood event.

The empirical decomposition results suggest a significant contribution to productivity growth from reallocation, roughly 45%, which is a bit higher than results in Foster, Haltiwanger, and Krizan (2001), but still within the general range of their results. Part of this could have been interpreted to be a result of a missing entry component. The model does well in capturing the decomposition. The third column introduces the model implied steady state entry to the decomposition and does confirm the notion that the somewhat high reallocation contribution in the data could be a result of missing entry observations.

The fourth column in Table 9 illustrates the model decomposition results without the measurement error. Both the cross-term and reallocation contribution components drop to close to zero magnitude. As will be shown in detail in section 5, true productivity growth is unaffected by both measurement error and demand shock variance. Nevertheless, the reallocation components in the FHK decomposition are shown to be highly sensitive to in particular measurement error. Foster, Haltiwanger, and Krizan (2001) mention this sensitivity and propose alternative reallocation contribution measures. However, the discussion of the BHC index in Petrin and Levinsohn (2005) suggests that there will still be a substantial disconnect between the structural growth decomposition laid out in section 5 and alternative variations on the reallocation component in the BHC decomposition.

In terms of identification, the cross-term component turns out to be of particular importance for the input measurement error parameter. If the model is estimated subject to $\sigma_W^2 = 0$, the remaining model parameters change modestly in the direction of more estimated type dispersion, but leaves the estimated cross term component close to zero. Allowing for input measurement error results in the fairly good fit of the cross-term component as shown in Table 9.

4.5 Estimation by Industry

Firm heterogeneity across industries may obscure the true picture of more homogenous subgroups of firms. This turns out not to be the case. Data moments by industry reveal the same qualitative picture as in Table 2. Table 10 presents data moments for the 3 largest industries (by firm count). All industries show evidence of significant firm productivity dispersion, a roughly zero correlation between productivity and firm input size and a positive correlation between productivity and firm output size (roughly 0.2). All industries also display significant productivity persistence and mean reversion. Finally, both the value added and wage bill distributions are characterized by a strong right shift over time across industries.

The estimates by industry are reported in Table 11. The model estimates by industry are not qualitatively different from the full sample estimate but it is worth noting a consistent drop in the estimated type dispersion in the industry estimates. This is likely a result of effectively allowing for more heterogeneity in other model parameters. The construction industry consists of generally significantly smaller firms and is estimated to have a significantly smaller non-labor cost share than the two other industries.

5 Reallocation and Growth

In this section we study the implications of the estimated model for aggregate growth and the role of reallocation in the growth process. From equations (4), (5) and (7), the normalized log of

Table 10: Data moments by industry

	Manufacturing		Wholesale and retail		Construction	
	1992	1997	1992	1997	1992	1997
Survivors	$2,051.000$	$1,536.000$	$1,584.000$	$1,189.000$	651.000	480.000
$E[Y]$	$30,149.461$	$35,803.473$	$22,952.920$	$28,386.719$	$15,191.354$	$16,869.551$
$Med[Y]$	$15,117.552$	$18,858.445$	$12,757.909$	$15,288.949$	$8,688.501$	$10,711.648$
$Std[Y]$	$56,081.995$	$69,574.991$	$33,400.313$	$41,409.060$	$31,287.564$	$22,454.655$
$E[W]$	$15,047.636$	$17,318.195$	$10,696.683$	$12,712.898$	$9,973.166$	$10,594.737$
$Med[W]$	$8,031.273$	$9,531.066$	$6,423.473$	$7,650.564$	$5,785.053$	$6,838.405$
$Std[W]$	$24,667.884$	$27,159.439$	$15,360.222$	$16,802.715$	$24,526.438$	$14,181.147$
$E\left[\frac{Y}{N^*}\right]$	379.047	422.471	410.234	466.591	305.075	342.273
$Med\left[\frac{Y}{N^*}\right]$	347.100	375.300	373.928	408.244	286.749	311.509
$Std\left[\frac{Y}{N^*}\right]$	163.174	226.860	171.661	278.495	133.111	173.871
$Cor[Y,W]$	0.889	0.855	0.922	0.914	0.967	0.922
$Cor\left[\frac{Y}{N^*},Y\right]$	0.236	0.200	0.252	0.188	0.131	0.174
$Cor\left[\frac{Y}{N^*},N^*\right]$	0.011	-0.003	-0.028	-0.039	-0.040	-0.093
$Cor\left[\frac{Y}{N^*},\frac{Y_{+1}}{N^*_{+1}}\right]$	0.650	0.728	0.325	0.674	0.428	0.345
$Cor\left[\frac{Y}{N^*},\Delta\frac{Y}{N^*}\right]$	-0.024	-0.195	-0.195	-0.259	-0.327	-0.560
$Cor\left[\frac{Y}{N^*},\frac{\Delta Y}{Y}\right]$	-0.133	–	-0.088	–	-0.187	–
$Cor\left[\frac{Y}{N^*},\frac{\Delta N^*}{N^*}\right]$	0.145	–	0.189	–	0.089	–
$E[\Delta Y/Y]$	-0.035	–	-0.042	–	-0.025	–
$Std[\Delta Y/Y]$	0.474	–	0.425	–	0.448	–
$Cor[\Delta Y/Y,Y]$	-0.073	–	-0.090	–	-0.122	–
Growth decomp.						
– Within	0.863	–	1.176	–	0.986	–
– Between	0.365	–	0.618	–	0.635	–
– Cross	-0.297	–	-0.826	–	-0.870	–
– Exit	0.068	–	0.032	–	0.249	–

consumption per product line can be written as

$$\frac{\ln C_t}{N} = \frac{1}{N}\sum_{j=1}^{N}\alpha_j\left(\sum_{i=1}^{J_t(j)}\ln q_i(j) + \ln\left(\frac{(1-\kappa)\,Z}{wq_{J_t(j)}(j)}\right)\right)$$

in the Cobb-Douglas case where $Z = \frac{1}{N}\sum_j^N Z_t(j)$ and $\alpha_j = NZ_t(j)/Z$ represent real expenditure shares. As the number of innovations $J_t(j),\ j = 1,...,N$ are independently and identically distributed Poisson random variables with common expectation δt, and the $\ln qs$ are iid across time and product lines, the law of large number implies that

$$\frac{\ln C_t}{N} = \delta E\{\ln q\}t + \ln(1-\kappa) - E\{\ln q\} + \ln(Z_t/w_t)$$

Table 11: Parameter estimates by industry

	Manufacturing	Wholesale and retail	Construction
c_0/Z	18.1369	15.3557	9.8525
c_1	2.8347	2.6580	2.8162
κ	0.4338	0.4853	0.3076
δ	0.0664	0.0640	0.0752
$\hat{g}$	0.0157	0.0182	0.0143
σ_Y^2	0.0246	0.0201	0.0339
σ_W^2	0.0160	0.0134	0.0202
θ_γ^Z	3,162.0872	1,343.5249	4,842.5514
θ_η^Z	16,350.0027	15,819.3413	4,219.2477
θ_β^Z	0.7739	0.9267	0.8508
ϕ_τ^{ss}			
– type 1	0.3752	0.8875	0.4106
– type 2	0.4327	0.0035	0.4742
– type 3	0.1921	0.1089	0.1152
$\theta_\gamma^{q_\tau}$			
– type 1	1.0000	1.0000	1.0002
– type 2	1.0000	1.0000	1.0014
– type 3	1.0533	1.0001	1.0118
$\theta_\eta^{q_\tau}$			
– type 1	0.0000	0.0000	0.0046
– type 2	0.0000	0.0034	0.0272
– type 3	0.6633	1.8416	0.3597
$\theta_\beta^{q_\tau}$	0.7171	0.9806	0.6295

holds as an approximation when N is large. Because the total Danish employment is roughly constant over the sample period, Z/w is also constant from the labor market clearing condition, equation (13). Hence, the growth rate in consumption is

$$\frac{\dot{C}}{C} = g = \delta E\{\ln q\} = \eta \sum_\tau E\left[\ln \tilde{q}_\tau\right] \phi_\tau + \sum_\tau \gamma_\tau E\left[\ln \tilde{q}_\tau\right] K_\tau \tag{29}$$

Note that the appropriate empirical counterpart is the traditional growth accounting measure recommended by Petrin and Levinsohn (2005), not the BHC index. However, an accurate measurement requires an appropriately constructed cost of living index of the type discussed by Hausman (2003).

5.1 The Growth Rate and Its Components

The equilibrium steady state growth rate implied by the model and our parameter estimates is $g = 1.96\%$, per year. Under the assumption that a true cost of living index were used to measure the aggregate real value added and wage bill in our data, namely an index of consumption goods prices that accounts for productivity improvements and product substitution, the traditional empirical growth measure (the TD index) using the value added per worker for our sample of continuing firms is 1.48% percent per year. Although we do not have an empirical measure for the contribution of entry and exit to growth in productivity, the model implies that the net effect of entry is very close to the difference between these two numbers. Hence, the implications of the structural model for the growth rate are consistent with the growth in value added observed in our data.

The model also permits the identification of the contribution survival and firm size selection, reflected in differential firm growth rates, to aggregate growth in consumption. Specifically, because the expected productivity of the products created differ across firms and because these differences are positively associated with differences in expected profitability, aggregate growth reflects the selection of more profitable firms by the creative-destruction process. Indeed, equation (29) can be rewritten as

$$g = \sum_{\tau} \gamma_\tau E\left[\ln \tilde{q}_\tau\right] \phi_\tau + \sum_{\tau} \gamma_\tau E\left[\ln \tilde{q}_\tau\right] (K_\tau - \phi_\tau) + \eta \sum_{\tau} E_\tau \left[\ln \tilde{q}_\tau\right] \phi_\tau. \tag{30}$$

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 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 $F_\tau \leq F_{\tau'} \implies$ both $E\left[\ln \tilde{q}_\tau\right] \geq E\left[\ln \tilde{q}_{\tau'}\right]$ and $K_\tau - \phi_\tau \geq K_{\tau'} - \phi_{\tau'}$.) The parameter estimates imply that the entry component accounts for 17.9% and the selection component 58.5% 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.

6 Concluding Remarks

Large and persistent differences in firm productivity and firm size exist. Worker reallocation induced by heterogeneity can be an important source of aggregate productivity growth. Previous work using Baily, Hulton, and Campbell (1992) type growth decomposition measures has found large contributions from reallocation. We argue however that the reallocation component of the BHC measure has no meaning in stochastic steady state model as ours. Thus, we arrive at the same conclusion as Petrin and Levinsohn (2005) who argue that the reallocation component does not capture the welfare implications of growth.

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 grow larger at the expense of less profitable firms 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. Furthermore, the model also fits the Foster, Haltiwanger, and Krizan (2001) variant of the BHC growth decomposition well. Notably, the model 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.

In the absence of other sources of growth such as dis-embodied growth, all growth in our model is attributed to reallocation. We decompose the reallocation component into a net contribution from firm entry and exit, a firm type selection effect, and a residual. The selection component captures the contribution to growth from the mechanism that more productive firms are estimated to grow larger at the expense of less productive firms. The selection component is estimated to contribute 58 percent of overall growth.

A Model simulation notes

In this section, we present the algorithm used to compute the values of model parameters implied by the estimates and the equilibrium and optimal growth rates, all reported in the text. To do so, one must account for the two parameters not explicitly used in the initial presentation of the model, the average demand per product, Z, which was normalized to unity in the model, and the cost of capital per product line, denoted κZ. Hence, profit per product line can be represented as πZ for a firm of type π where

$$\pi = (1-\kappa)(1-q^{-1}) \tag{31}$$

is now profit express as a fraction of value average sales.

Since the parametric form of the steady state distribution of firms over profit, denoted $p(\pi)$ in the text, is specified in the model estimated, one needs to derive its relationship to the initial density of entering firms over profit, $\phi(\pi)$, by inverting the steady state relationship implied by the model. Specifically,

$$p\left(\pi\right) = M\left(\pi\right)/M$$

where $M(\pi)$ is the steady state mass of firms of type π and $M = \int_\pi M(\pi)d\pi$ is the total mass of firms. Since

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

from equation (14), it follows that

$$\eta\phi(\pi) = \frac{\gamma(\pi)M\left(\pi\right)}{\ln\left(\frac{\delta}{\delta-\gamma(\pi)}\right)} = \frac{\gamma(\pi)p\left(\pi\right)M}{\ln\left(\frac{\delta}{\delta-\gamma(\pi)}\right)},$$

At this stage, the aggregate entry rate η and the total mass of firms M have yet to be separately identified. But by $\int_\pi \phi\left(\pi\right)d\pi = 1$, it follows that,

$$\eta = \eta\int_\pi \phi(\pi)d\pi = M\int_\pi \frac{\gamma(\pi)p\left(\pi\right)}{\ln\left(\frac{\delta}{\delta-\gamma(\pi)}\right)}d\pi. \tag{32}$$

Consequently, the profit density at entry is

$$\phi(\pi) = \frac{\frac{\gamma(\pi)p(\pi)}{\ln\left(\frac{\delta}{\delta-\gamma(\pi)}\right)}}{\int_x \frac{\gamma(x)p(x)}{\ln\left(\frac{\delta}{\delta-\gamma(x)}\right)}dx}. \tag{33}$$

Equation (16) and the assumption that the measure of products is unity, the steady state measure of continuing firms in the market solves

$$\begin{aligned} 1 &= \int_\pi \sum_{k=1}^{\infty} kM_k(\pi)d\pi = \int_\pi M(\pi) \sum_{k=1}^{\infty} \frac{kM_k(\pi)}{M(\pi)} d\pi \\ &= \int_\pi \frac{\gamma(\pi)M(\pi)}{(\delta-\gamma(\pi))\ln\left(\frac{\delta}{\delta-\gamma(\pi)}\right)} d\pi = M \int_\pi \frac{\gamma(\pi)p(\pi)}{(\delta-\gamma(\pi))\ln\left(\frac{\delta}{\delta-\gamma(\pi)}\right)} d\pi. \end{aligned} \tag{34}$$

Hence,

$$\eta = \frac{\int_\pi \frac{\gamma(\pi)p(\pi)}{\ln\left(\frac{\delta}{\delta-\gamma(\pi)}\right)} d\pi}{\int_\pi \frac{\gamma(\pi)p(\pi)}{(\delta-\gamma(\pi))\ln\left(\frac{\delta}{\delta-\gamma(\pi)}\right)} d\pi}. \tag{35}$$

from by equations (32) and (34).

To solve the planner's problem, one also needs the size of the aggregate labor force, L, and the measure of potential entrants, m. Because one can show that the limit price charged by the current supplier of each product solves $p(1-\kappa) = wq$ when a capital cost exists, the demand for production workers is $Zx(\pi) = 1/p = Z(1-\kappa)/wq = Z(1-\kappa-\pi)/w$ from (31). Hence, equations (13) and (14) imply

$$L = Z\left[\int_\pi \left(\frac{1-\kappa-\pi}{w} + \widetilde{c}(\gamma(\pi))\right) \frac{\eta\phi(\pi)d\pi}{\delta-\gamma(\pi)} + m\widetilde{c}(\eta/m)\right] \tag{36}$$

where, as specified in the text, $\widetilde{c}(x) = c_0 x^{1+c_1}$. Finally, one can obtain the value of m by using the fact that the marginal cost of entry must equal the expected marginal cost of innovation by incumbents. Specifically, equations (12) and (11) imply require that m solves

$$\widetilde{c}'\left(\frac{\eta}{m}\right) = \int_\pi \widetilde{c}'(\gamma(\pi))\,\phi(\pi)d\pi \tag{37}$$

Finally, the parametric specification of heterogeneity in product productivity is

$$q(z) = 1 + e^{\mu_\pi + \sigma_\pi z} \tag{38}$$

where z is the standard normal random variable. Hence, one can use the fact that $f(z)dz = p(\pi(z))d\pi(z)$, where $f(z)$ is the standard normal pdf and $\pi(z) = (1-\kappa)(1-q(z)^{-1})$ by (31), to compute all the necessary integrals in the equations above and those that define the components of the growth rate found in the text.

B Identification

Consider a simplified setup of Lentz and Mortensen (2006). Specifically, disregard stochastic demand shocks as well as stochastic quality realizations. In this case, the value function of a type τ firm with k products is,

$$V_\tau(k) = kZv_\tau,$$

where

$$v_\tau \equiv \max_{\gamma} \frac{\pi(q_\tau) - w\tilde{c}(\gamma)}{r + \delta - \gamma},$$

where $\tilde{c}(\gamma) = c(\gamma)/Z$. Hence, the optimal creation rate choice is,

$$w\tilde{c}'(\gamma_\tau) = v_\tau.$$

Finally, we have that $\pi(q) = (1-\kappa)\left(1 - q^{-1}\right)$.

Simulation of (Y, W, N) data from the model then follows from,

$$\begin{aligned}
Y_\tau(k) &= kZ \\
W_\tau(k) &= kZ\left(\frac{1-\kappa}{q_\tau} + w\tilde{c}(\gamma_\tau)\right) \\
N_\tau(k) &= W_\tau(k)/w = kZ\left(\frac{1-\kappa}{wq_\tau} + \tilde{c}(\gamma_\tau)\right).
\end{aligned}$$

We can define the type conditional labor share by

$$\alpha_\tau = \frac{1-\kappa}{q_\tau} + w\tilde{c}(\gamma_\tau).$$

In this case, the expression simplify to,

$$\begin{aligned}
Y_\tau(k) &= kZ \\
W_\tau(k) &= kZ\alpha_\tau.
\end{aligned}$$

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

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

This is $3M+1$ independent parameters given the restriction that $\sum p_i = 1$. Taking separate identification on w and r, the underlying structural parameters of the model are,

$$\{c_0, c_1, \kappa, \delta, Z, (q_1, \ldots, q_M), (p_1, \ldots, p_M)\},$$

which is $2M+4$ independent parameters. This suggests identification problems for $M \leq 2$.

Identification requires that the underlying true data generating process has at least three distinct types. For the three type case, identification of $(c_0, c_1, \kappa, q_1, q_2, q_3)$ comes from the system,

$$\begin{aligned}
wc_0 (1+c_1) (\gamma_1^*)^{c_1} &= \frac{1-\kappa-\alpha_1^*}{r+\delta^*-\gamma_1^*} \\
wc_0 (1+c_1) (\gamma_2^*)^{c_1} &= \frac{1-\kappa-\alpha_2^*}{r+\delta^*-\gamma_2^*} \\
wc_0 (1+c_1) (\gamma_3^*)^{c_1} &= \frac{1-\kappa-\alpha_3^*}{r+\delta^*-\gamma_3^*} \\
\alpha_1^* &= \frac{1-\kappa}{q_1} + wc_0 (\gamma_1^*)^{1+c_1} \\
\alpha_2^* &= \frac{1-\kappa}{q_1} + wc_0 (\gamma_2^*)^{1+c_1} \\
\alpha_3^* &= \frac{1-\kappa}{q_1} + wc_0 (\gamma_3^*)^{1+c_1}.
\end{aligned}$$

where the $\alpha^* s$ and $\gamma^* s$ are determined by the data. The parameters (c_0, c_1, κ) are identified from the first three equations and, given these (q_1, q_2, q_3) are obtained from the last three.

C The General CES Case

C.1 Derivations

Consider a simplified version the model estimated by Lentz and Mortensen (2006). Let types be permanent. Disregard stochastic demand shocks as well as stochastic quality realizations. The consumption good is supplied by many competitive providers at a price expressed in terms of the numeraire as P_t and the aggregate quantity produced is determined by the quantity and quality of the economy's intermediate inputs. Specifically, there is a large number of intermediate products indexed by j and consumption is determined by the following constant returns CES production function

$$C_t = \left[\sum_j (A_t(j) x_t(j))^{\frac{\sigma}{\sigma-1}} \right]^{\frac{\sigma-1}{\sigma}} \tag{39}$$

where $\sigma \geq 0$ is the elasticity of substitution between any two good, $x_t(j)$ is the quantity of input j at time t and $A_t(j)$ is quality or productivity of input j at time t. Let

$$X_t(j) = A_t(j)x_t(j) \text{ and } P_t(j) = p(j)A_t^{-1}(j)$$

denote the quantity and price of intermediate input j expressed in quality units. Given profit maximization in the competitive final goods market and a choice of the numeraire so that $P_tC_t = 1$ for all t , the share of expenditure on input j is

$$\frac{P_t(j)X_t(t)}{P_tC_t} = P_t(j)X_t(t) = \left(\frac{P_t(j)}{P_t}\right)^{1-\sigma}, \tag{40}$$

where $\sum_j \left(\frac{P_t(j)}{P_t}\right)^{1-\sigma} = 1$ and the definition of $P_t(j)$ above imply

$$P_t = \left(\sum_j (P_t(j))^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = A_t^{-1}\left(\sum_j p(j)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}. \tag{41}$$

Under the assumption that the capital requirement and labor productivity is the same across intermediate inputs, the common marginal cost of production is $w/(1-\kappa)$. As a new innovation is q times better than the product it will replace, the innovator can charge no more than the limit price $qw/(1-\kappa)$ in order to take over the market. Of course, this limit price is bounded above by the profit maximizing monopoly price. In sum,

$$p(j) = \frac{m(q)w}{1-\kappa} \text{ where the mark up is } m(q) = \begin{cases} \frac{\sigma}{1-\sigma} \text{ if } q > \frac{\sigma}{1-\sigma} \\ q \text{ otherwise} \end{cases} \tag{42}$$

Note that the markup is generally independent of the product type but does depend on the quality improvement relative to the last version of the type.

The demand for a given product will not generally be stationary over its life cycle. Indeed, because the price of a product does not change over its life time but the prices of substitutes and output do, the revenue from any product of age a created at date t

$$P_t(j)X_{t+a}(j) = \left(\frac{P_t(j)}{P_{t+a}}\right)^{1-\sigma}$$

has a trend except in the case of $\sigma = 1$. Indeed, the expected profit of and the demand for product

j created by a type i firm at future date $t + a$ expressed in terms of the numeraire is

$$\begin{aligned}
\pi_i(t+a) &= \sum_j \left[(1-\kappa)p(j)x_{t+a}(j) - wx_{t+a}(j)\right] \\
&= (1-\kappa)\sum_j \left[p_t(j) - \frac{w}{1-\kappa}\right] A_t^{-1} P_t(j)^{-\sigma} P_{t+a}^{\sigma-1} \\
&= (1-\kappa)w\left(1-m(q_i)^{-1}\right)\sum_j \left(\frac{P_t(j)}{P_{t+a}}\right)^{1-\sigma} \\
&= (1-\kappa)w\left(1-m(q_i)^{-1}\right)\left(\frac{P_t}{P_{t+a}}\right)^{1-\sigma} \\
&= (1-\kappa)w\left(1-m(q_i)^{-1}\right)e^{(1-\sigma)ga}
\end{aligned}$$

and

$$\begin{aligned}
x_i(t+a) &= \sum_j A_t^{-1} P_t(j)^{-\sigma} P_{t+a}^{\sigma-1} = \sum_j p(j)\left(\frac{P_{t+a}}{P_t(j)}\right)^{\sigma-1} \\
&= \frac{m(q_i)w}{1-\kappa}\left(\frac{P_{t+a}}{P_t}\right)^{\sigma-1} = \frac{m(q_i)w}{1-\kappa}e^{(1-\sigma)ga}
\end{aligned}$$

where the sum can be eliminated using equation (41) and the last equality is implied by $P_{t+a} = P_t e^{-ga}$ in steady state in each case. Hence, profits and demand grows with age, represented by a, at a constant rate equal to $(1-\sigma)g$ in steady state, which is negative if and only if, the elasticity of substitution σ exceeds unity.

C.2 Identification Revisited

In general, $m(q_i)$ replaced q_i and profit and demand have a trend equal to $(1-\sigma)g$ relative to the Cobb-Douglas ($\sigma = 1$) case. However, the first order condition can be written as

$$\begin{aligned}
wc'(\gamma) &= \int_0^\infty \pi(a)e^{-(r+\delta)a}da + \frac{\gamma w\widetilde{c}'(\gamma) - w\widetilde{c}(\gamma)}{r+\delta} \\
&= \frac{(1-\kappa)\left(1-m(q_i)^{-1}\right)}{r+\delta-(1-\sigma)g} + \frac{\gamma w\widetilde{c}'(\gamma) - w\widetilde{c}(\gamma)}{r+\delta}
\end{aligned}$$

given convergence of the integral.

To compute the actual current revenue, one needs the ages of all products, another random variable. However, the average demand per product, $\overline{x}$, can be computed as follows: As the age on any product is exponentially distributed with parameter δ,

$$\begin{aligned}\overline{x} &= \int_0^\infty \delta x(a) e^{-\delta a} da = \frac{m(q_i)w}{1-\kappa}\int_0^\infty \delta e^{-(\delta-(1-\sigma)g)a} da \\ &= \frac{\delta}{\delta-(1-\sigma)\,g}\frac{wm(q_i)}{1-\kappa} \text{ if } \delta-(1-\sigma)\,g>0.\end{aligned}$$

The wage bill as a fraction of average value added is

$$\alpha = \frac{1-\kappa}{m(q)} + \frac{\delta-(1-\sigma)\,g}{\delta} w\widetilde{c}(\gamma).$$

In sum, we have added another parameter σ. Given three types and observations on the three $\alpha^* s$ and three $\gamma^* s$ and the growth rate g, we have six equations and seven unknowns: $m(q_1), m(q_2), m(q_3), c_0, c_1, \kappa$,and σ given the power specification of the cost function $\widetilde{c}(\gamma) = c_0\gamma^{1+c_1}$. Of course, the growth equation

$$g = \eta\sum_i \ln(q_i)\phi_i + \sum_i \gamma_i \ln(q_i)\frac{\eta_i}{\delta-\gamma_i}$$

ties down g. From equation (42), one needs at least four types to identify the qs when the constraint $m(q) \leq \frac{\sigma}{1-\sigma}$ is slack. If it binds, then σ is identified by the constraint, g is determined instead by the other equations, and the growth equation can be solved for the highest q. In sum, four firm types are necessary and sufficient for identification in the general CES case.

References

Altonji, J. G. and L. M. Segal (1996). Small-sample bias in GMM estimation of covariance structures. *Journal of Business and Economic Statistics 14*: 353–366.

Alvarez, 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. Hulton, 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, John C Haltiwanger, and Stefano Scarpetta (2004). Microeconomic evidence of creative destruction in industrial and developing countries. *The World Bank, Policy Research Working Paper Series 3464*.

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.

Boskin, Michael J., Eller R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale Jorgenson (1996). *Toward a More Accurate Measure of the Cost of Living*. Final Report to the Senate Finance Comittee from the Advisory Commisson to Study the Consumer Price Index. Washington: Senate Finance Committee.

Christensen, Bent Jesper, Rasmus Lentz, Dale T. Mortensen, George Neumann, and Axel Werwatz (2005). On the job search and the wage distribution. *Journal of Labor Eco-*

nomics 23, no. 1: 31–58.

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

Fallick, Bruce C. and Charles A. Fleischman (2001). The importance of employer-to-employer flows in the u.s. labor market. *Federal Reserve Board Finance and Economics Discussion Paper 2001-18*.

Foster, 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.

Frederiksen, A. and N. Westergaard-Nielsen (2002). Where did they go? *Århus School of Business Working Paper*.

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 Regev Haim (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*.

Hausman, Jerry (2003). Sources of bias and solutions to bias in the consumer price index. *Journal of Economic Perspectives* 17, no. 1: 23–44.

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.

Nagypál, Éva (2005). On the estend of job-to-job transitions. *Working paper*.

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*.

Stewart, Jay (2002). Recent trends in job stability and job security: Evidence from the march CPS. *U.S. BLS Working Paper 356*.

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.