ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

http://www.econometricsociety.org/

Econometrica, Vol. 76, No. 6 (November, 2008), 1317–1373

AN EMPIRICAL MODEL OF GROWTH THROUGH PRODUCT INNOVATION

RASMUS LENTZ
University of Wisconsin–Madison, Madison, WI 53706-1393, U.S.A. and Centre for Applied Microeconometrics, University of Copenhagen, Copenhagen, Denmark

DALE T. MORTENSEN
Northwestern University, Evanston, IL 60208, U.S.A. and NBER and Institute for the Study of Labor, Bonn, Germany

Econometrica, Vol. 76, No. 6 (November, 2008), 1317–1373

AN EMPIRICAL MODEL OF GROWTH THROUGH PRODUCT INNOVATION

BY RASMUS LENTZ AND DALE T. MORTENSEN[1]

Productivity differences across firms are large and persistent, but the evidence for worker reallocation as an important source of aggregate productivity growth is mixed. The purpose of this paper is to estimate the structure of an equilibrium model of growth through innovation designed to identify and quantify the role of resource reallocation in the growth process. The model is a version of the Schumpeterian theory of firm evolution and growth developed by Klette and Kortum (2004) extended to allow for firm heterogeneity. The data set is a panel of Danish firms that includes information on value added, employment, and wages. The model's fit is good. The estimated model implies that more productive firms in each cohort grow faster and consequently crowd out less productive firms in steady state. This selection effect accounts for 53% of aggregate growth in the estimated version of the model.

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

1. INTRODUCTION

IN THEIR REVIEW ARTICLE of empirical productivity studies based on longitudinal plant and firm data, Bartelsman and Doms (2000) concluded 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.

In this paper, we quantify the effects of worker reallocation on productivity growth. Our model is a richer version of that studied in an earlier paper (Lentz and Mortensen (2005)), where we established the existence of a general equilibrium solution. The model is an extension of that proposed by Klette and Kortum (2004), which itself builds on the endogenous growth models of Grossman and Helpman (1991) and Aghion and Howitt (1992). It is designed to capture the implications for growth through reallocation induced by the creative–destruction process.

[1]This research was supported by collaborative grant SES-0452863 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, Gadi Barlevy, Michele Boldrin, Jonathan Eaton, Robert Hall, Hugo Hopenhayn, John Kennan, Samuel Kortum, Giuseppe Moscarini, Jean-Marc Robin, Rob Shimer, three anonymous referees, and the editor for many useful comments and suggestions. All remaining errors are ours. Supplemental materials to this paper are provided in Lentz and Mortensen (2008).

 DOI: 10.3982/ECTA5997

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

Using the equilibrium relationships of the model, the parameters are estimated by the method of indirect inference based on information on value added, employment, and wage payments drawn from a Danish panel of firms over the period 1992–1997. The model is estimated on a number of cross-section and dynamic moments, including size, productivity, and firm growth distribution moments. The data do not contain any direct observation on a firm's innovation activity. In the model, innovation activity determines the stochastic processes that govern firm size and productivity dynamics. These panel moments and the model's structure allow inference regarding innovation. Using patent data, Balasubramanian and Sivadasan (2008) provided evidence in favor of the Klette and Kortum (2004) model of innovation and firm dynamics by investigating the direct impact of patenting activity on firm size and measured productivity. Although patent data are a narrow reflection of investment in firm growth, patent activity can be viewed as a proxy.

In our model, the aggregate growth rate in final good consumption is equal to the sum of the expected percentage increase in the productivity of the intermediate inputs weighted by their contributions to final consumption output. This term can be decomposed by type of firm into the net contribution of entrants and incumbents. The net contribution of entry is the average increase in productivity of the entrants relative to those that exit the market within each period. The second term—that associated with continuing firms—can be decomposed into two parts designed to reveal the consequences of the selection process associated with differences in firm growth and survival rates: The first is the contribution of incumbents if counterfactually the share of value added supplied by each firm type were to remain equal to that at entry; the second, the selection effect, is the contribution of the difference between the steady state share and the share at entry. Because a more productive firm type grows faster, its share in steady state exceeds that at entry which implies that selection induced by differential growth contributes positively to the aggregate growth rate. Indeed, our estimated model implies that net entry accounts for 21% of the aggregate growth rate, while 53% can be attributed to the selection effect.

Although all productivity growth in the model is associated with reallocation, we emphasize the selection effect as the contribution to growth that results from the extent to which more productive firms increase their share of the economy at the expense of less productive firms through reallocation. In the model, firm types differ with respect to the product quality improvement distribution. The selection effect works through differential innovation rates across more or less productive firm types ordered by stochastic dominance. If all firms face the same product quality improvement distribution, hence they choose the same innovation rate, then the selection effect is zero. Innovative activity has a positive spillover in that all firm innovate on the same quality frontier and innovations push the frontier forward. Hence, while more productive firms on average produce greater quality improvements in our model, the fact that all firms innovate on the same frontier implies a constant average productivity difference between high and low productivity firms in steady state. The selection impacts on the steady state growth rate rather than the productivity level because more productive firms contribute more to the quality frontier.

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 4900 firms is restricted to those with 20 or more employees. The sample does not include entrants.[2] 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 real Danish kroner (DKK), while N is a body count.

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

[2]The full panel of roughly 6700 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.

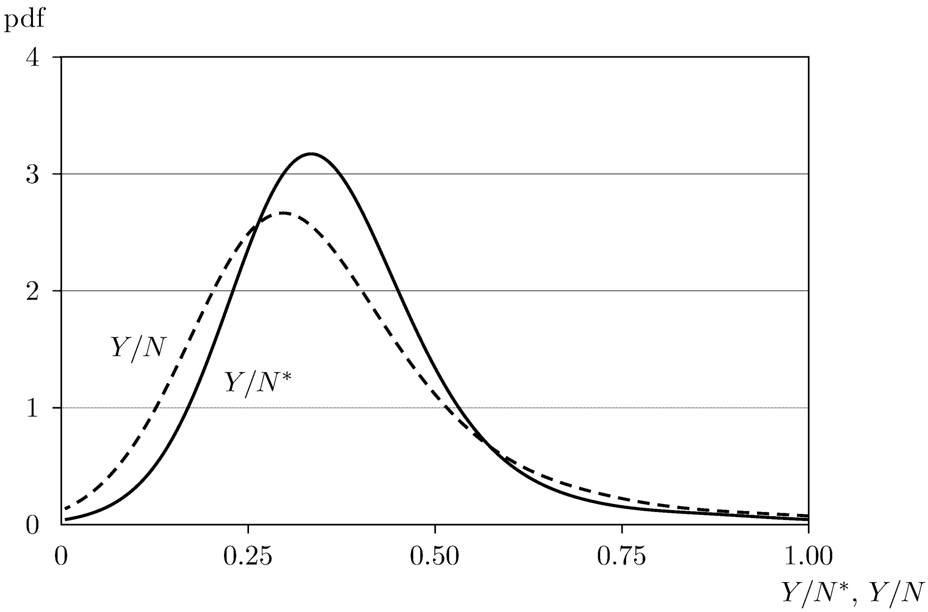

FIGURE 1.—Observed firm productivity distribution, 1992. Value added (Y) measured in 1 million DKK. N is the raw labor force size measure. N^* is the quality adjusted labor force size.

for this source of cross-firm differences in productive efficiency. Formally, the constructed quality adjusted employment of firm j is defined as $N_j^* = W_j/w$, where

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

is the average wage paid per worker in the market.[3] Although correcting for wage differences across firms in this manner does reduce the spread and skew of the implied productivity distribution somewhat, both distributions have high variance and skew, and are essentially the same general shape.

Both distributions are consistent with those found in other data sets. For example, productivity distributions are significantly dispersed and skewed to the right. In the case of the adjusted measure of productivity, the 5th percentile is roughly half the mode, while the 95th percentile is approximately twice as large as the mode. The range between the two represents a fourfold differ-

[3]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 I
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

ence in value added per worker across firms. These facts are similar to those reported by Bartelsman and Doms (2000) for the United States.

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 essential. 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 I. As documented in the table, the correlation between labor force size and productivity using either the raw or the adjusted employment measure is zero. However, note the strong positive association between value added and both measures of labor productivity.

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

3. AN EQUILIBRIUM MODEL OF CREATIVE–DESTRUCTION

As is well known, firms come is an amazing range of shapes and sizes. This fact cannot be ignored in any analysis of the relationship between firm size and productivity. Furthermore, an adequate theory must account for entry, exit, and firm evolution to explain the size distributions observed. Klette and Kortum (2004) constructed 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) allowed for productive heterogeneity, firm productivity and growth are unrelated because costs and benefits of growth are both proportional to firm productivity in their model. Allowing for a positive relationship between firm growth and productivity is necessary for consistency with the relationships found in the Danish firm data studied in this paper.

3.1. *Preferences and Technology*

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

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

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

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

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

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

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

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

Aggregate profit from supplying the consumption good at any point in time t is

$$P_tC_t - \int_0^1 P_t(j)X_t(j)$$
$$= P_t\left[\int_0^1 \alpha(j)X_t(j)^{(\sigma-1)/\sigma}\,dj\right]^{\sigma/(\sigma-1)} - \int_0^1 P_t(j)X_t(j)\,dj,$$

where $X_j = A_t(j)x_t(j)$ is the total input of intermediate good j express in quality units and $P_j(t)$ is the price of a quality unit of intermediate good j. Hence, the final goods profit maximizing expenditures on each input are given by

$$P_t(j)X_t(j) = \left(\frac{P_t}{P_t(j)}\right)^{\sigma-1}\alpha(j)^\sigma Z, \quad j \in (0,1). \tag{5}$$

Because the production function is homogenous of degree 1, maximal final good profit is zero,

$$P_tC_t = \int_0^1 P_t(j)X_t(j)\,dj = \int_0^1 \left(\frac{P_t}{P_t(j)}\right)^{\sigma-1}\alpha(j)^\sigma P_tC_t\,dj,$$

which implies that the price of the consumption good is

$$P_t = \left(\int_0^1 P_t(j)^{1-\sigma}\alpha(j)^\sigma\,dj\right)^{1/(1-\sigma)}. \tag{6}$$

3.2. *Intermediate Good Prices*

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

The output of any intermediate good requires labor and capital input in fixed proportions. Total factor productivity is the same across all goods and is set equal to unity without loss of generality. Denote by w the wage of a unit of labor and by κ the cost of capital per unit of output. The price per quality unit of intermediate good j is $P_t(j) = p_t(j)/A_t(j)$, where $p_t(j)$ is the price of a unit of the good at date t. As the innovator of a new version of any good takes over the market by either setting a Bertrand price equal to $q > 1$ times the marginal cost of production of producing the existing version of the good or the profit maximizing monopoly price, whichever is smaller. As the monopoly price is $1 - 1/\sigma$ if $\sigma > 1$ and is undefined otherwise,

$$p_t(j) = \min\left(q_t(j), \max\left(\frac{\sigma}{\sigma - 1}, 1\right)\right)(w + \kappa) = m(q_t(j), \sigma)(w + \kappa), \tag{7}$$

where

$$m(q, \sigma) = \min\left(q, \max\left(\frac{\sigma}{\sigma - 1}, 1\right)\right) \tag{8}$$

represents the price markup over cost. Let

$$x_t(j) = \frac{X_t(j)}{A_t(j)} = \frac{\alpha(j)^\sigma Z}{A_t(j)P_t(j)}\left(\frac{P_t}{P_t(j)}\right)^{\sigma-1} = \frac{z_t(j)}{p_t(j)}$$

represent the unit demand for input j and let

$$z_t(j) = \alpha(j)^\sigma Z\left(\frac{P_t}{P_t(j)}\right)^{\sigma-1} \tag{9}$$

represent the revenue per unit expressed in term of the numeraire. The profit flow at date t is

$$\begin{aligned} P_t(j)X_t(j) - (w + \kappa)x_t(j) &= (p_t(j) - w - \kappa)x_t(j) \\ &= \pi(q_t(j), \sigma)z_t(j), \end{aligned}$$

where profit per unit

$$\pi(q_t(j), \sigma) = (p_t(j) - w - \kappa)/p_t(j) = 1 - m(q_t(j), \sigma)^{-1}. \tag{10}$$

3.3. *Aggregate Growth Rate*

Note that the normalization used to define the numeraire, $P_tC_t = Z$, implies that consumption grows at the rate of deflation in the price of the consumption good in terms of the numeraire, that is,

$$C_t = \frac{Z}{P_t} = Z\left(\int_0^1 \left(\frac{A_t(j)}{p_t(j)}\right)^{\sigma-1} \alpha(j)^\sigma \, dj\right)^{1/(\sigma-1)} \tag{11}$$

from equation (6) and $p_t(j) = A_t(j)P_t(j)$. A new version of any product variety arrives at the common rate of creative–destruction, denoted by δ. In other words, the number of innovations to date in product j, denoted $J_t(j)$, is a Poisson random variable with arrival frequency δ. The improvement in productivity embodied in an innovation, denoted as q, is also a random variable independent across intermediate good types and arrival orders. In Appendix A, we show that the rate of growth in aggregate consumption is

$$g \equiv \frac{dC_t/dt}{C_t} = \delta\left(\frac{E\left\{\dfrac{q''m(q',\sigma)}{m(q'',\sigma)}\right\}^{\sigma-1} - 1}{\sigma - 1}\right), \tag{12}$$

the product of the innovation arrival rate δ and the average contribution of innovations to aggregate consumption.[4] In the latter expression, q'' is the improvement in quality embodied in the last innovation in any intermediate good as of time t, q' is the quality step size of the previous innovation in the same good, and $E\{\cdot\}$ represents the expectation taken with respect to the joint distribution of both of them.

By taking the limit, $g = \delta E\{\ln q\}$ in the Cobb–Douglas ($\sigma = 1$) case. In general, however, the growth rate in consumption is the product of the rate at which innovations arrive and an index of the average improvement in consumption embodied in the last innovation, which depends on the elasticity of substitution between any two intermediate inputs. The latter dependence takes account of the reallocation of demand for inputs by the suppliers of the consumption good induced by innovation.

The net contribution of an innovation to consumption is also sensitive on the pricing rule. As the markup is $m(q,\sigma) = \sigma/(\sigma - 1)$ in the pure monopoly case, the expected contribution term is $E\{q''\}^{\sigma-1}$ because the entire improvement in productivity is passed on to the consumer. However, in the pure Nash–Bertrand case $m(q,\sigma) = q$, which implies that the term is $E\{q'\}^{\sigma-1}$. Only the quality improvement of the previous innovation is passed to the consumer in this case because the last innovator captures the difference.

3.4. *The Value of a Firm*

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 that reflects product creation and destruction. A firm enters with one product and a firm exits when it no longer has any 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

[4]This derivation of the growth rate in the CES case under Bertrand competition appears to be an orginal contribution of this paper.

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, $C(\gamma k, k) = kc(\gamma)$, is linearly homogenous in the new product arrival rate and the number of existing products "... 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 of creative–destruction, represented by δ. Below we refer to γ as the firm's innovation rate. 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 $F_{\tau'}(\tilde{q}) \le F_\tau(\tilde{q})$ for all $\tilde{q} \ge 1$.[5] Assume that the lower bound of the support of $\tilde{q}_\tau$ is 1 for all τ.

By assumption firms cannot direct their innovation activity toward a particular market. Furthermore, their ability to create new products is not specific to any one or subset of product types.[6] Since product demands vary across product varieties, firms face demand uncertainty for a new innovation that is resolved only when the variety is realized. By equation (10), the initial product line demand realization of a product created at date s is $z_s(j) = \alpha(j)^\sigma Z(P_s(j)/P_s)^{1-\sigma}$, a result of a random draw over the product space $j \in [0, 1]$. Denote by $G(\cdot)$ the cumulative distribution function of $z_s(j)$ across products. As the price per quality unit increases at the same rate as the price of consumption goods, on average, the initial demand for an innovation is stationary and independent of both product type j and the innovation's increment to product quality conditional on type $\tilde{q}_\tau$. Let $\tilde{z} \sim G(\cdot)$ represent a realization of $z_s(j)$. By definition of $z_s(j)$ in equation (9) it follows that $Z = E[\tilde{z}] = \int \tilde{z}\, dG(\tilde{z})$.

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

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

By $P_t C_t = Z$, the growth rate in the final consumption good price is $\dot{P}_t/P_t = -g$. Conditional on no innovation in product j at time t, the price per quality unit $P_t(j)$ does not change. Hence, in this case, by equation (9), the demand follows the law of motion, $\dot{z}_t(j) = (1-\sigma)g z_t(j)$.

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

$$\begin{aligned} (13)\qquad rV_\tau(\tilde{q}^k, \tilde{z}^k, k) = \max_{\gamma \ge 0} \Bigg\{ & \sum_{i=1}^{k} \tilde{z}_i \pi(\tilde{q}_i, \sigma) - kwc(\gamma) \\ & + k\gamma\big[E_\tau[V_\tau(\tilde{q}^{k+1}, \tilde{z}^{k+1}, k+1)] - V_\tau(\tilde{q}^k, \tilde{z}^k, k)\big] \\ & + k\delta\Bigg[\frac{1}{k}\sum_{i=1}^{k} V_\tau\big(\tilde{q}^{k-1}_{\langle i\rangle}, \tilde{z}^{k-1}_{\langle i\rangle}, k-1\big) - V_\tau(\tilde{q}^k, \tilde{z}^k, k)\Bigg] \\ & + \frac{\partial V_\tau(\tilde{q}^k, \tilde{z}^k, k)}{\partial \tilde{z}^k}\dot{\tilde{z}}^k \Bigg\}, \end{aligned}$$

where conditional on survival of product i,

$$(14)\qquad \dot{\tilde{z}}_i = (1-\sigma)g\tilde{z}_i$$

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

As discussed in detail in Appendix A, the unique solution to (13) is

$$(15)\qquad V_\tau(\tilde{q}^k, \tilde{z}^k, k) = \sum_{i=1}^{k} \frac{\tilde{z}_i \pi(\tilde{q}_i, \sigma)}{r + \delta + g(\sigma - 1)} + kZ\Psi_\tau,$$

where

$$\Psi_\tau = \max_{\gamma \ge 0} \frac{\gamma \nu_\tau - w\hat{c}(\gamma)}{r+\delta},$$

$$\nu_\tau = \frac{\bar{\pi}_\tau(\sigma)}{r+\delta+g(\sigma-1)} + \Psi_\tau,$$

$\bar{\pi}_\tau(\sigma) = 1 - E[m(\tilde{q}_\tau, \sigma)^{-1}]$, and $\hat{c}(\gamma) \equiv c(\gamma)/Z$. The first term on the right side of (15) is the expected present value of the future profits of the firm's portfolio at date t. Specifically, because the remaining length of a product's life is exponential with constant hazard δ, the price of the good does not change during its lifetime, and the price level deflates at rate g the present value of future profit expected in the case of product i is given by:

$$\begin{aligned}
&\pi(\tilde{q}_i, \sigma) \int_t^\infty \tilde{z}_i \left(\frac{P_{t'}}{P_t}\right)^{\sigma-1} e^{-\delta(t'-t)} e^{-r(t'-t)}\, dt' \\
&\quad = \pi(\tilde{q}_i, \sigma)\tilde{z}_i \int_t^\infty e^{-(r+\delta+g(\sigma-1))(t'-t)}\, dt' \\
&\quad = \frac{\tilde{z}_i \pi(\tilde{q}_i, \sigma)}{r+\delta+g(\sigma-1)}.
\end{aligned}$$

Ψ_τ is the type conditional innovation option value embodied in each product. ν_τ is the type conditional expected value of a product. It is the sum of the innovation option value and the discounted stream of expected profits attributable to the innovation, where the effective discount rate is the sum of the interest rate, the product destruction rate, and the rate of decline in future demand for the product induced by the decline in the price of the consumption good during the product's lifetime.

It then follows directly from (13) that the firm's optimal choice of creation rate, γ_τ, satisfies

$$w\hat{c}'(\gamma_\tau) = \nu_\tau, \tag{16}$$

where ν_τ is the type conditional expected value of an additional product line. Equation (16) implies that the type contingent creation rate is size independent—a theoretical version of Gibrat's law. Also, the second order condition, $c''(\gamma) > 0$, and the fact that the marginal value of a product line is increasing in $\bar{\pi}_\tau$ imply that a firm's creation rate increases with profitability. Therefore, we obtain that $\gamma_{\tau'} \geq \gamma_\tau$ for $\tau' \geq \tau$. These results are the principal empirical implications of the model.

3.5. *Firm Entry*

The entry of a new firm requires innovation. Suppose that there is a constant measure μ 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 $\eta = \mu\gamma_0$, the entry rate satisfies the free entry condition

$$w\hat{c}'\left(\frac{\eta}{\mu}\right) = \sum_{\tau} \nu_\tau \phi_\tau, \tag{17}$$

where ϕ_τ is the probability of being a type τ firm at entry. Of course, the second equality follows from equation (16).

3.6. *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. This number is a birth–death process with birth frequency γk and death frequency δk. Kette and Kortum (2004) showed that steady state mass of firms of type τ converges in steady state to

$$M_\tau(k) = \frac{k-1}{k}\frac{\gamma_\tau}{\delta} M_\tau(k-1) = \frac{\eta\phi_\tau}{\delta k}\left(\frac{\gamma_\tau}{\delta}\right)^{k-1}. \tag{18}$$

The size distribution of firms conditional on type can be derived using equation (18). Specifically, the total firm mass of type τ is

$$\begin{aligned} M_\tau &= \sum_{k=1}^{\infty} M_\tau(k) = \frac{\phi_\tau \eta}{\delta} \sum_{k=1}^{\infty} \frac{1}{k}\left(\frac{\gamma_\tau}{\delta}\right)^{k-1} \\ &= \frac{\eta}{\delta} \ln\left(\frac{\delta}{\delta - \gamma_\tau}\right)\frac{\delta\phi_\tau}{\gamma_\tau}, \end{aligned} \tag{19}$$

where convergence requires that the aggregate rate of creative–destruction exceed the creation rate of every incumbent type, that is, $\delta > \gamma_\tau$ $\forall\tau$. Hence, the steady state fraction of type τ firms with k products is defined by

$$\frac{M_\tau(k)}{M_\tau} = \frac{\dfrac{1}{k}\left(\dfrac{\gamma_\tau}{\delta}\right)^k}{\ln\left(\dfrac{\delta}{\delta - \gamma_\tau}\right)}. \tag{20}$$

Equation (20) is the steady state distribution of size as reflected in the number of product lines, $\tilde{k}_\tau$. This is the logarithmic distribution with parameter γ_τ/δ. Consistent with the observations on firm size distributions, this distribution is highly skewed to the right.

The expected product size of a type τ firm is

$$E[\tilde{k}_\tau] = \sum_{k=1}^{\infty} \frac{kM_\tau(k)}{M_\tau} = \frac{\dfrac{\gamma_\tau}{\delta - \gamma_\tau}}{\ln\left(\dfrac{\delta}{\delta - \gamma_\tau}\right)}. \tag{21}$$

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

$$\frac{\partial E[\tilde{k}_\tau]}{\partial \gamma_\tau} = \left(\frac{(1 + \zeta_\tau)\ln(1 + \zeta_\tau) - \zeta_\tau}{(1 + \zeta_\tau)\ln^2(1 + \zeta_\tau)}\right)\frac{1 + \zeta_\tau}{\delta - \gamma_\tau} > 0, \tag{22}$$

where $\zeta_\tau = \gamma_\tau/(\delta - \gamma_\tau)$.

For future reference, the total mass of products produced by type τ firms, K_τ, is

$$K_\tau = \sum_{k=1}^{\infty} kM_\tau(k) = \frac{\eta\phi_\tau}{\delta - \gamma_\tau}. \tag{23}$$

It too is larger for more innovative types.

3.7. *Labor Market Clearing*

There is a fixed measure of available workers, denoted by ℓ, seeking employment at any positive wage. In equilibrium, these workers are allocated across production and R&D activities performed by both incumbent firms and potential entrants.

By the normalization of labor productivity at unity, it follows from equations (7) and (9) that the production labor demand for an age a product line with initial demand realization $\tilde{z}$ and quality realization $\tilde{q}$ is given by

$$\ell^d(\tilde{z}, \tilde{q}, a) = \frac{\tilde{z}e^{g(1-\sigma)a}}{m(\tilde{q}, \sigma)(w + \kappa)}, \tag{24}$$

where the term $e^{g(1-\sigma)a}$ captures the reduction in demand (if $\sigma > 1$) attributed to the rising relative price of the product during its lifetime. Denote by ℓ_τ^d the average labor demand per product of a type τ firm for the purpose of production. Product age is going to be exponentially distributed with parameter δ. It follows that

$$\ell_\tau^d = \int_1^\infty \int_0^\infty \int_0^\infty \frac{\tilde{z}e^{g(1-\sigma)\tilde{a}}\delta e^{-\delta\tilde{a}}}{m(\tilde{q}, \sigma)(w + \kappa)}\, d\tilde{a}\, dG(\tilde{z})\, dF_\tau(\tilde{q}) \tag{25}$$

$$= \frac{\delta Z(1 - \bar{\pi}_\tau(\sigma))}{(w+\kappa)[\delta + g(\sigma - 1)]},$$

where the equality reflects the definition $\bar{\pi}_\tau(\sigma) = 1 - E[m(\tilde{q}_\tau, \sigma)^{-1}]$. Hence, the average type conditional labor demand per product is

$$\ell_\tau = \ell_\tau^d + c(\gamma_\tau). \tag{26}$$

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

Labor market clearing requires that the equilibrium wage solves

$$\ell = \sum_\tau K_\tau \ell_\tau + \mu c(\eta/\mu). \tag{27}$$

3.8. *The Components of the Growth Rate*

The rate of creative–destruction is the sum of the rate of entry and the innovation frequency of the firm types,

$$\delta = \eta + \sum_\tau K_\tau \gamma_\tau, \tag{28}$$

where K_τ is the mass of products supplied by firms of type τ. This is the aggregate innovation frequency in the expression for aggregate growth, equation (12).

The expression for the expected contribution to growth of an innovation in equation (12) is unconditional with respect to firm type. However, it is composed of the sum of the increments to consumption attributable to the various firm types. Given that the innovator is of type τ, then the term q'' in (12), the quality increment induced by the last innovation, is distributed $F_\tau(q)$. However, the distribution of the previous innovation's quality step q' depends on the type of the incumbent supplier at the time the last version was created. As that can be any of the types, its distribution is the convolution of the distribution of q given type and the innovation frequency distribution across types. Since the frequency of innovation attributable to type τ' firms in steady state is

$\phi_{\tau'}\eta + K_{\tau'}\gamma_{\tau'}$, where $\phi_{\tau'}$ is the fraction of entrants of type τ', the steady state distribution q' can be regarded as the unconditional distribution

$$F(q) = \sum_{\tau'} \frac{\phi_{\tau'}\eta + K_{\tau'}\gamma_{\tau'}}{\eta + \sum_\tau K_\tau \gamma_\tau} F_{\tau'}(q). \tag{29}$$

Given this definition and the fact that the two innovator types are independent, the net contribution to consumption growth of the most recent innovation conditional on firm type is

$$E_\tau\left(\frac{q'' m(q',\sigma)}{m(q'',\sigma)}\right)^{\sigma-1} = \int\int\left(\frac{\tilde{q}'' m(\tilde{q}',\sigma)}{m(\tilde{q}'',\sigma)}\right)^{\sigma-1} dF(\tilde{q}')\, dF_\tau(\tilde{q}''). \tag{30}$$

Finally, taking into account innovator firm type, the growth rate in consumption can be written as

$$\begin{aligned} g \equiv \frac{dC_t/dt}{C_t} &= \delta\left(\frac{E\left(\frac{q'' m(q',\sigma)}{m(q'',\sigma)}\right)^{\sigma-1} - 1}{\sigma - 1}\right) \\ &= \delta \sum_\tau \left(\frac{\phi_\tau \eta + K_\tau \gamma_\tau}{\eta + \sum K_\tau \gamma_\tau}\right) \frac{E_\tau\left(\frac{q'' m(q',\sigma)}{m(q'',\sigma)}\right)^{\sigma-1} - 1}{\sigma - 1} \\ &= \sum_\tau (\phi_\tau \eta + K_\tau \gamma_\tau) \frac{E_\tau\left(\frac{q'' m(q',\sigma)}{m(q'',\sigma)}\right)^{\sigma-1} - 1}{\sigma - 1}. \end{aligned} \tag{31}$$

In other words, the aggregate growth rate is equal to the product of the innovation arrival rate and the net contribution of an innovation to aggregate consumption summed over firm types. As both components are larger for more productive types, more productive firms contribute more to the aggregate growth rate for both reasons.

3.9. *Equilibrium*

DEFINITION 1: A *steady state market equilibrium* is a triple (w, δ, g) together with optimally chosen entry rate $\eta = \mu\gamma_0$, creation rate γ_τ, and a steady state size distribution K_τ for each type that satisfy equations (16), (17), (23), (27), (28), and (31).

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

4. ESTIMATION

If the ability to create higher quality products is a permanent firm characteristic, then differences in firm profitability are associated with differences in the product creation rates chosen by firms. Specifically, more profitable firms grow faster, are more likely to survive in the future, and supply a larger number of products on average. Hence, a positive cross-firm correlation between current gross profit per product and sales volume should exist. Furthermore, worker reallocation from slow growing firms to more profitable fast growing firms will be an important source of aggregate productivity growth because faster growing firms also contribute more to growth.

In this section, we demonstrate that firm specific differences in profitability are required to explain Danish interfirm relationships between value added, employment, and wages paid. In the process of fitting the model to the data, we also obtain estimates of the investment cost of innovation function that all firms face as well as the sampling distribution of firm productivity at entry.

4.1. *Danish Firm Data*

If more profitable firms grow faster in the sense that $\bar{\pi}_\tau > \bar{\pi}_{\tau'} \Rightarrow \gamma_\tau > \gamma_{\tau'}$, then (22) implies that fast growing firms also supply more products and sell more on average. However, because production employment per product decreases with productivity, total expected employment need not increase with $\bar{\pi}_\tau$ in general and decreases with $\bar{\pi}_\tau$ when growth is independent of a firm's past product productivity improvement realizations. These implications of the theory can be tested directly.

The model is estimated on an unbalanced panel of 4872 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 overrepresentation of high creation rate firm types relative to steady state. The ability to observe the gradual exit of the 1992 cross section will be a useful source of identification. Entry in the original data set suffers from selection bias and while one can attempt to correct for the bias, we have made the choice to leave out entry altogether since it is not necessary for identification. By including the 1997 cross section in the set of moments, dynamic processes that change the cross-sectional composition of survivors over time are reflected in the estimation.

The first two columns of Table II present a set of distribution moments with standard deviations in parentheses. The standard deviations are obtained through bootstrapping on the original panel. Unless otherwise stated, amounts are in 1000 real 1992 Danish kroner where the Statistics Denmark consumer

TABLE II

DATA MOMENTS (STANDARD ERRORS IN PARENTHESES)[a]

	1992	1997		1992	1996
Survivors	4872.000	3628.000	$\text{Cor}[\frac{Y}{N^*}, \frac{Y_{+1}}{N^*_{+1}}]$	0.476	0.550
	—	(32.132)		(0.088)	(0.091)
$E[Y]$	26,277.262	31,860.850	$\text{Cor}[\frac{Y}{N^*}, \Delta\frac{Y}{N^*}]$	−0.227	−0.193
	(747.001)	(1031.252)		(0.103)	(0.057)
$\text{Med}[Y]$	13,472.812	16,448.965	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta Y}{Y}]$	−0.120	
	(211.851)	(329.417)		(0.016)	
$\text{Std}[Y]$	52,793.105	64,120.233	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta N^*}{N^*}]$	0.119	
	(5663.047)	(7741.448)		(0.032)	
$E[W]$	13,294.479	15,705.087	$E[\frac{\Delta Y}{Y}]$	−0.029	
	(457.466)	(609.595)		(0.008)	
$\text{Med}[W]$	7231.812	8671.939	$\text{Std}[\frac{\Delta Y}{Y}]$	0.550	
	(92.720)	(154.767)		(0.067)	
$\text{Std}[W]$	30,613.801	35,555.701	$\text{Cor}[\frac{\Delta Y}{Y}, Y]$	−0.061	
	(6750.399)	(8137.541)		(0.012)	
$E[\frac{Y}{N^*}]$	384.401	432.118	Within	1.015	
	(2.907)	(5.103)		(0.146)	
$\text{Med}[\frac{Y}{N^*}]$	348.148	375.739	Between	0.453	
	(1.829)	(2.139)		(0.112)	
$\text{Std}[\frac{Y}{N^*}]$	205.074	305.306	Cross	−0.551	
	(19.633)	(42.491)		(0.196)	
$\text{Cor}[Y, W]$	0.852	0.857	Exit	0.084	
	(0.035)	(0.045)		(0.066)	
$\text{Cor}[\frac{Y}{N^*}, N^*]$	−0.018	−0.026			
	(0.013)	(0.011)			
$\text{Cor}[\frac{Y}{N^*}, Y]$	0.198	0.143			
	(0.036)	(0.038)			

[a]Unit of measurement is 1000 DKK. The BHC growth decomposition moments are stated as fractions of total growth.

price index was used to deflate nominal amounts. It is seen that the size distributions are characterized by significant skew. The value added per worker distribution displays some skew and significant dispersion. All distributions display a right shift from 1992 to 1997. The distribution moments also include the positive correlation between firm productivity and output size, and the slightly negative correlation between firm productivity and labor force size.

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

Finally, Table II also includes a standard empirical labor productivity growth decomposition. We use the preferred formulation in Foster, Haltiwanger, and Krizan (2001) which is taken from Baily, Bartelsman, and Haltiwanger (1996) and ultimately based on the Baily, Hulten, and Campbell (1992) index (BHC).[7] The decomposition takes the form[8]

$$\Delta\Theta_t = \sum_{i\in\widehat{C}_t} s_{it-1}\Delta\theta_{it} + \sum_{i\in\widehat{C}_t}\theta_{it-1}\Delta s_{it} + \sum_{i\in\widehat{C}_t}\Delta\theta_{it}\Delta s_{it} + \sum_{i\in\widehat{E}_t}\theta_{it}s_{it} - \sum_{i\in\widehat{X}_t}\theta_{it-1}s_{it-1}, \tag{32}$$

where $\Theta_t = \sum_i s_{it}\theta_{it}$, $\theta_{it} = Y_{it}/N_{it}$, and $s_{it} = N_{it}/N_t$; $\widehat{C}_t$ is the set of continuing firms, $\widehat{E}_t$ is the set of entering firms and $\widehat{X}_t$ is the set of exiting firms.

The identity in (32) decomposes time differences in value added per worker into five components in the order stated on the right hand side: within, between, cross, and entry/exit. As the names suggest, the BHC growth decomposition literature attaches particular significance to each term: The within component is interpreted as identifying growth in the productivity measure due to productivity improvements by incumbents. The between component is interpreted to capture productivity growth from reallocation of labor from less to more productive firms. The cross component captures a covariance between input shares and productivity growth, and the last two terms capture the growth contributions from entry and exit. The sum of the between and cross components is also sometimes referred to as gross reallocation.

We include the BHC growth decomposition in the set of data moments because it conveniently relates the estimation to the empirical growth literature. Furthermore, it does reflect a particular aspect of the dynamics in the data. As

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

[8]In the implementation of the decomposition we employ the version of (32) where the between term and the entry/exit terms are normalized by Θ_{t-1}:

$$\Delta\Theta_t = \sum_{i\in\widehat{C}_t} s_{it-1}\Delta\theta_{it} + \sum_{i\in\widehat{C}_t}(\theta_{it-1} - \Theta_{t-1})\Delta s_{it} + \sum_{i\in\widehat{C}_t}\Delta\theta_{it}\Delta s_{it} + \sum_{i\in\widehat{E}_t}(\theta_{it} - \Theta_{t-1})s_{it} - \sum_{i\in\widehat{X}_t}(\theta_{it-1} - \Theta_{t-1})s_{it-1}.$$

mentioned, the sample in this paper does not include entry, so there is no entry share in the decomposition. Consequently, the decomposition cannot be directly related to the results in Foster, Haltiwanger, and Krizan (2001), although a full decomposition is performed on the estimated model in Section 4.7.

4.2. *The BHC Growth Decomposition in Steady State*

Studies based on the BHC decomposition identities provide mixed evidence of the importance of reallocation as a source of aggregate productivity growth.[9] Based on data on U.S. manufacturing firms, Bartelsman and Doms (2000) found that roughly one-quarter of growth can be attributed to gross reallocation, another quarter to net entry, and roughly half of all growth to within firm growth. However, based on the same data Foster, Haltiwanger, and Krizan (2001) found that "...much of the increase in labor productivity would have occurred even if labor share had been held constant at their initial levels."[10] In a study of a number of different OECD (Organization for Economic Cooperation and Development) countries, Scarpetta, Hemmings, Tressel, and Woo (2002) also found that the majority of growth can be attributed to within firm growth. It is seen from Table II that the Danish data support the finding of a small gross reallocation effect.

In this section we argue that the between firm component does not capture the role of reallocation in the growth process in our model. Despite the fact that all growth is associated with worker reallocation in our model, the term typically interpreted as the contribution of gross reallocation is close to zero in our data. We show that this result is to be expected in a stochastic equilibrium model such as ours. Although in the model, more profitable firms in each cohort grow faster on average as a consequence of more frequent innovation, the aggregate share of products supplied and inputs required by each firm type are constant in the model's ergodic steady state by definition. As a consequence, the "between" and "cross" terms in the Foster, Haltiwanger, and Krizan (2001) decomposition should be zero in the absence of measurement error and transitory shocks.

In the model, all firms of the same type have the same productivity by the definition of type, and although individual firms can and do grow and contract over time, the steady state distribution of inputs over firm types is stationary by the definition of stationary stochastic equilibrium. Hence, if we let $j \in J$ represent an element of the set of firm types, let I_j denote the set of firms

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

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

of type j, let s^*_{jt} represent the average share of employment per type j firm in period t, and let θ^*_{jt} be the productivity of type j firms, then by abstracting from entry and exit we can formulate the growth decomposition in terms of firm types as[11]

$$\begin{aligned}(33)\qquad \Delta\Theta_t &= \sum_{j\in J}\sum_{i\in I_j} s_{it-1}\Delta\theta_{it} + \sum_{j\in J}\sum_{i\in I_j}\theta_{it-1}\Delta s_{it} + \sum_{j\in J}\sum_{i\in I_j}\Delta\theta_{it}\Delta s_{it} \\ &= \sum_{j\in J}|I_j|s^*_{jt-1}\Delta\theta^*_{jt} + \sum_{j\in J}|I_j|\theta^*_{jt-1}\Delta s^*_{jt} + \sum_{j\in J}|I_j|\Delta\theta^*_{jt}\Delta s^*_{jt} \\ &= \sum_{j\in J}|I_j|s^*_j\Delta\theta^*_{jt},\end{aligned}$$

where $|I_j|$ is the number of firms of type j and $s^*_{jt-1} = (1/|I_j|)\sum_{i\in I_j} s_{it-1}$. The first equality is implied by the fact that the set $\{I_1, I_2, \ldots, I_j, \ldots\}$ is a partition of the set of all firms I, the second by the fact that the firms of the same type have the same productivity at any given date, and the last by the fact that the average share per firm of each type is constant (consequently, $\Delta s^*_{jt} = 0$ for all j and t) in a steady state equilibrium. The final expression in (33) is the first term in the Baily, Hulten, and Campbell (1992) index.

An interpretation of the sum of the between and cross components, $\sum_{i\in I}\Delta s_{it}\theta_{it}$, as the gross effect of reallocating resources across firms is incorrect because gains in employment share are exactly offset by losses in share across firms of the same type in steady state. In other words, workers are never exogenously reallocated across types in equilibrium as is implicit in the interpretation. As such, the decomposition cannot capture the steady state growth contribution from reallocation. The fact that many empirical studies based on the BHC decomposition have found little evidence of a significant contribution to growth from the gross reallocation component is not a surprise in light of the above argument.[12]

4.3. *Model Estimator*

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

[11]The general argument that includes entry and exit is presented in the Appendix.

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

The model is estimated by indirect inference. The estimation procedure, as described in, for example, Gourieroux, Monfort, and Renault (1993), Hall and Rust (2003), and Browning, Ejrnæs, and Alvarez (2006), is as follows: First, define a vector of auxiliary data parameters $\Gamma(\psi)$. The vector consists of all the items in Table II except the number of survivors in 1992 and one of the growth decomposition components. Thus, $\Gamma(\psi)$ has length 37.

Next, produce a simulated panel $\psi^s(\omega)$ for a given set of model parameters ω. The model simulation is initialized by assuming that the economy is in steady state in the first year and consequently that firm observations are distributed according to the ω implied steady state distribution.[13]

The simulated auxiliary parameters are then given by

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

where S is the number of simulation repetitions.[14]

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

$$\hat{\omega} = \arg\min_{\omega\in\Omega}(\Gamma^s(\omega) - \Gamma(\psi))' A^{-1}(\Gamma^s(\omega) - \Gamma(\psi)), \tag{34}$$

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

The variance of the estimator is estimated by bootstrap. In each bootstrap repetition, a new set of data auxiliary parameters $\Gamma(\psi^b)$ is produced, where ψ^b is the bootstrap data in the bth 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 bth repetition, an estimator ω^b is found by minimizing the weighted distance between the recentered bootstrap data auxiliary parameters $[\Gamma(\psi^b) - \Gamma(\psi)]$ and the recentered simulated auxiliary parameters $[\Gamma^s(\omega^b) - \Gamma^s(\hat{\omega})]$:

$$\begin{aligned}\omega^b = \arg\min_{\omega\in\Omega}&\big([\Gamma^s(\omega) - \Gamma^s(\hat{\omega})] - [\Gamma(\psi^b) - \Gamma(\psi)]\big)' \\ &\times A^{-1}\big([\Gamma^s(\omega) - \Gamma^s(\hat{\omega})] - [\Gamma(\psi^b) - \Gamma(\psi)]\big).\end{aligned}$$

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

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

In each bootstrap repetition, a different seed is used to generate random numbers for the determination of $\Gamma^s(\omega)$. Hence, $V(\hat{\omega})$ captures both data variation and variation from the model simulation.[15]

4.4. *Model Specification and Simulation*

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

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

$$Y_\tau(\tilde{q}^k, \tilde{z}^k) = \sum_{i=1}^{k} \tilde{z}_i, \tag{35}$$

and, by equation (24), a wage bill of

$$W_\tau(\tilde{q}^k, \tilde{z}^k) = \frac{w}{w+\kappa} \sum_{i=1}^{k} \frac{\tilde{z}_i}{m(\tilde{q}_i, \sigma)} + wkc(\gamma_\tau). \tag{36}$$

Equations (35) and (36) provide the foundation for the model simulation.

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

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

where $M = \sum_\tau M_\tau$ is the total steady state mass of firms. A firm's type is drawn according to ϕ^*. Once a firm's type has been determined, its 1992 product line size is drawn from the type conditional steady state distribution of $\tilde{k}_\tau$ characterized in equation (20). Then the age realization of each product is drawn from

[15]Variance estimates are obtained using 500 bootstrap repetitions.

the exponential age distribution. The age realization is used to adjust the demand realization draw for each product from $G(\cdot)$ according to equation (10).

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

Given an initial size for a firm, its future size evolves according to the stochastic birth–death process described earlier. The forward simulation is done by dividing each annual time period into a large number of discrete subintervals, n. By assumption, in each subinterval each of the stochastic creation and destruction processes can have zero or one event arrival. Hence, a type τ firm with k products will, in a given subinterval, lose a product with probability $1 - e^{-k\delta/n}$ and gain a product with probability $1 - e^{-k\gamma_\tau/n}$. As $n \to \infty$, the procedure will perfectly represent the continuous time processes in the model. In the simulations below, the model has been simulated with $n = 26$.

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,

$$\ln \hat{Y}_\tau(\tilde{q}^k, \tilde{z}^k) = \ln Y_\tau(\tilde{q}^k, \tilde{z}^k) + \xi_Y,$$
$$\ln \hat{W}_\tau(\tilde{q}^k, \tilde{z}^k) = \ln W_\tau(\tilde{q}^k, \tilde{z}^k) + \xi_W,$$

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

4.5. *Identification*

The interest rate is set at $r = 0.05$. The wage w is immediately identified as the average worker wage in the sample $w = 190.24$. Excluding these two, the set of structural model parameters ω has 16 parameters, $\omega = (c_0, c_1, \kappa, \sigma, \mu, \beta_Z, \xi_Z, o_Z, \beta_q, \xi_1, \xi_2, \xi_3, \phi_1, \phi_2, \sigma_Y^2, \sigma_W^2)$, where $\phi_3 = 1 - \phi_1 - \phi_2$ in the case of three types. In the actual implementation of the estimation, μ is replaced by η as a fundamental model parameter. Of course, η is endogenous to the equilibrium, but since μ is a free parameter, μ can always be set to make the η estimate consistent with steady state equilibrium. The set of data moments $\Gamma(\psi)$ has size 37.

The estimation identifies the aggregate entry rate η, yet the data provide no direct observation of entrants. Rather, the data follow the life and death of a given cross section of firms. As will be argued below, given the structure of the model, this allows direct inference on the aggregate destruction rate δ and the creation rates of the incumbents γ_τ. The equilibrium condition (28), stating that aggregate creation equals aggregate destruction, is the key part of the model structure that then identifies the entry rate η. The assumption that entrants face the same innovation cost function as the incumbents and have yet to learn their type then identifies the mass of potential entrants μ through equation (17). Given the wage estimate, one can then immediately obtain the total labor demand, which in turn provides the equilibrium estimate of the total labor supply ℓ from equation (26). The choice of directly estimating the η and w, which are endogenous to the equilibrium, allows us to avoid a lengthy equilibrium fixed point search for each choice of model parameter candidates in the estimation described by equation (34). For any given estimate of η and w, one can always find a combination of μ and ℓ that makes the η and w estimates consistent with equilibrium.

To understand the identification of the model, it is useful to consider a stripped down version without stochastic demand and within firm type product quality improvement realizations; that is, $V[\tilde{z}] = 0$ and $V[\tilde{q}_\tau] = 0\ \forall\tau$. Furthermore, average out demand variation due to product age dispersion. This allows a reduced form view of the model where output size is proportional to the stochastic realization of the latent firm type dependent random variable $\tilde{k}_\tau$, that is, the number of product lines. The simplified model implies a constant type dependent labor share λ_τ. Hence, the firm's wage bill and labor demand are also directly proportional to the latent random variable $\tilde{k}_\tau$. The following identification discussion is based on the argument that the data identify this reduced form and that the fundamental model parameters are identified only insofar as one can establish a mapping from the reduced form back to the underlying structural model.

Given the simplifying assumptions of no demand and within type quality realization variation, by equations (15) and (16) the optimal type conditional creation rate choice satisfies

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

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

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

provided that $\delta > g(1-\sigma)$. By equation (26), the expected type conditional wage bill is $E[W_\tau] = w\ell_\tau E[\tilde{k}_\tau]$. The average type conditional labor share is then given by

$$\lambda_\tau = E\left[\frac{W_\tau}{Y_\tau}\right] = w\left[\frac{1 - \bar{\pi}_\tau(\sigma)}{w + \kappa} + \frac{\delta - g(1-\sigma)}{\delta}\tilde{c}(\gamma_\tau)\right], \quad \tau = 1, \dots, N. \tag{37}$$

The reduced form data generating process for (Y, W, N) panel data then follows from the expressions

$$Y_{\tau,t}(\tilde{k}_\tau) = \tilde{k}_\tau Z e^{gt}, \tag{38}$$

$$W_{\tau,t}(\tilde{k}_\tau) = \tilde{k}_\tau Z \lambda_\tau e^{gt}, \tag{39}$$

$$N_{\tau,t}(\tilde{k}_\tau) = \frac{W_{\tau,t}(\tilde{k}_\tau)}{we^{gt}} = \tilde{k}_\tau Z \frac{\lambda_\tau}{w}, \tag{40}$$

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

To solve for the type conditional dynamics of $(Y_{\tau,t}, W_{\tau,t}, N_{\tau,t})$, it is necessary to know (δ, γ_τ) because these two parameters govern the birth–death process of $\tilde{k}_\tau$. Thus, to simulate the full firm panel $\{Y_{jt}, W_{jt}, N_{jt}\}_{j,t}$ for N separate firm types based on the reduced form in equations (38)–(40), it is necessary and sufficient to know

$$\Lambda = \{\delta, z, g, (\lambda_1, \dots, \lambda_N), (\gamma_1, \dots, \gamma_N), (\phi'_1, \dots, \phi'_N)\}. \tag{41}$$

This is $3N + 2$ independent parameters given the restriction that $\sum_\tau \phi^*_\tau = 1$. Given the reduced form in equations (38)–(40) and the specification of the stochastic birth–death process for the latent variable, $\tilde{k}_\tau$, the firm panel data straightforwardly identify the parameters in equation (41). The identification argument on the set of structural model parameters relies on establishing a unique mapping between the identified parameters in equation (41) and the model parameters.

Taking separate identification on w and r, the underlying structural parameters of the simplified model are

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

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

Given the other parameters, $\{\delta, z, (\phi'_1, \dots, \phi'_N)\}$ and $\{\eta, z, (\phi_1, \dots, \phi_N)\}$ are related to each other one-to-one. The separate identification discussion can consequently be confined to the relationship between $\{g, (\lambda_1, \dots, \lambda_N), (\gamma_1, \dots, \gamma_N)\}$ and $\{\sigma, c_0, c_1, \kappa, (q_1, \dots, q_N)\}$. Thus, let

$$\{\hat{g}, (\hat{\lambda}_1, \dots, \hat{\lambda}_N), (\hat{\gamma}_1, \dots, \hat{\gamma}_N)\}$$

be the estimated set of parameters for the reduced form in (38)–(40). It must be that $\hat{\delta} > \hat{\gamma}_\tau$, $\forall \tau$. There exists a unique η that is consistent with steady state for given $\hat{\delta}$ and $\hat{\gamma}_\tau$ (see the Appendix for proof). Denote this steady state implied aggregate entry rate by $\hat{\eta}$. Similarly, the steady state product mass distribution across types $\hat{K}_\tau$ also directly follows from the given $\hat{\delta}$ and $\hat{\gamma}_\tau$. $\{\sigma, c_0, c_1, \kappa, (q_1, \dots, q_N)\}$ is then identified through the system

$$wc_0(1+c_1)\hat{\gamma}_\tau^{c_1} = \frac{\bar{\pi}_\tau(\sigma)}{r+\hat{\delta}-\hat{g}(1-\sigma)} + \frac{wc_0c_1\hat{\gamma}_\tau^{1+c_1}}{r+\hat{\delta}}, \quad \tau = 1, \dots, N, \tag{42}$$

$$\hat{\lambda}_\tau = w\left[\frac{1-\bar{\pi}_\tau(\sigma)}{w+\kappa} + \frac{\hat{\delta}-\hat{g}(1-\sigma)}{\hat{\delta}}c_0\hat{\gamma}_\tau^{1+c_1}\right], \quad \tau = 1, \dots, N, \tag{43}$$

$$\hat{g} = \sum_\tau (\phi'_\tau \hat{\eta} + \hat{K}_\tau \hat{\gamma}_\tau) \frac{E_\tau\left(\dfrac{q'' m(q', \sigma)}{m(q'', \sigma)}\right)^{\sigma-1} - 1}{\sigma - 1}, \tag{44}$$

where the profits have been explicitly stated to depend on the type dependent quality improvement and σ. In the case of three distinct types, equations (42)–(44) are seven equations in seven unknowns.[16]

The estimation of type heterogeneity is tied to three characteristics of the data: (i) the observation of substantial dispersion in value added per worker, (ii) the positive relationship between value added per worker and output size, and (iii) the flat relationship between value added per worker and input size. Within the framework of the model, it is possible to generate value added per worker dispersion without type heterogeneity through stochastic q realizations. However, it does not generate a positive relationship between value added per worker and output size, whereas a negative relationship between labor shares and creation rates will. The positive correlation between Y/N and Y, and the zero correlation between Y/N and N will result in an estimate where $\gamma_{\tau'} > \gamma_\tau \Rightarrow \lambda_{\tau'} < \lambda_\tau$. It directly follows from equation (43) that $q_{\tau'} > q_\tau$. As a caveat, it is worthwhile noting that all three characteristics could, in principle, be a result of simple measurement error in Y. If so, it should be clear that one would need a pretty sizeable measurement error. We have included measurement error in the estimation for this and other reasons. As it turns out, measurement error is estimated to have little impact on the relationship between value added per worker and firm size as well as value added per worker dispersion. Finally, the dynamics in the data play an important role as well. They determine magnitudes of the creative–destruction rates through firm death rates, the dispersion in growth rates, and changes in cross-section moments over time.

Unlike other applications of the Dixit–Stiglitz demand model, where σ is identified directly through an average observed markup in the data, the source

[16]Estimation of the model with only two firm types produces a low type with a creation rate equal to zero and a single high type. The type dependent equilibrium creation rates remain well identified, but one can show that the estimation in this case lacks separate identification of c_0 and c_1. For the purpose of this paper, this would not be an issue since the results do not rely on counterfactual creation rates.

of identification of σ is nontrivial in this case. Although $\sigma > 1$ imposes an upper bound, markups are first and foremost generated through quality improvements because of the Nash–Bertrand competition between producers of the same intermediate good. The data strongly reject values of σ larger than 1 because they diminish value added per worker dispersion across firms. In Figure 5 in the Appendix we plot the minimized criterion value for different values of σ. The figure implies an estimate of $\hat{\sigma} \approx 0.75$, but it is also clear that the estimation criterion as a function of σ is very flat for $\sigma \leq 1$. This suggests weak identification of σ in the range of values less than 1. We have chosen to present the central results in this paper subject to the assumption of $\sigma = 1$. We then provide a robustness analysis of some of the central implications of the estimation to different values of σ in Section 5.

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 this is a strong assumption, it reasonably assumes that the numerous very small firms in the economy are qualitatively different from those in this analysis and are not just firms with fewer products.

4.6. *Estimation Results*

The model parameter estimates are presented in Table III. As mentioned, the standard errors are obtained through bootstrapping. The estimation is performed subject to $\sigma = 1$.

The overall productivity growth rate is estimated at 1.39% annually. We discuss the growth rate estimate in greater detail in Sections 4.7.8 and 5.

The model estimates imply that at least three distinct types have significant representation in the steady state equilibrium.[17] The low type produces no improvement in quality, whereas the median quality improvement of the middle and high types is 15% and 18%, respectively. The low type represents 74% of all firms and produces 54% of the products in steady state. This is in stark contrast to the low type's representation at entry, which is estimated at 85%. This reflects a significant selection in steady state which is attributed to creation rate dispersion across the firm types. The low type's creation rate is zero, whereas the middle and high types have creation rates at $\gamma_2 = 0.055$ and $\gamma_3 = 0.057$, respectively. The high and middle types are in effect crowding out the the low type through the creative–destruction process. This is true in terms of firm representation because the low type has a higher exit rate, but it is particularly strong in terms of product representation because the middle and high types are on average substantially larger than the low type. The survival conditional size expectation is equal to a single product for the low type, whereas it is 2.35 products for the middle type and 2.48 products for the high type.

[17] An estimation with four discrete types did not improve the fit appreciably.

TABLE III

PARAMETER ESTIMATES (STANDARD ERROR IN PARENTHESES)[a]

			$\tau=1$	$\tau=2$	$\tau=3$
c_0/Z	175.8159	ϕ_τ	0.8478	0.0952	0.0570
	(12.9903)		(0.0133)	(0.0239)	(0.0230)
c_1	3.7281	ϕ^*_τ	0.7387	0.1614	0.0999
	(0.0411)		(0.0165)	(0.0403)	(0.0394)
κ	150.0126	ξ_τ	0.0000	0.3524	0.4168
	(2.0581)		(0.0000)	(0.0471)	(0.0750)
Z	16,859.4212	β_q	0.4275	0.4275	0.4275
	(329.3350)		(0.0199)	(0.0199)	(0.0199)
β_Z	0.9577	γ_τ	0.0000	0.0553	0.0566
	(0.0227)		(0.0000)	(0.0017)	(0.0016)
o_Z	608.2725	K_τ	0.5408	0.2776	0.1816
	(113.1480)		(0.0170)	(0.0703)	(0.0700)
η	0.0451	ν_τ	0.0000	3.2392	3.5332
	(0.0016)		(0.0000)	(0.1779)	(0.2540)
σ^2_Y	0.0323	$\bar{\pi}_\tau(\sigma)$	0.0000	0.2499	0.2690
	(0.0037)		(0.0000)	(0.0121)	(0.0166)
σ^2_W	0.0254	$\text{Med}[\tilde{q}_\tau]$	1.0000	1.1495	1.1769
	(0.0000)		(0.0000)	(0.0221)	(0.0339)
δ	0.0707	$E[\tilde{k}_\tau]$	1.0000	2.3503	2.4833
	(0.0019)		(0.0000)	(0.1062)	(0.1470)
μ	1.3406				
	(0.0519)				
ℓ	45.7344				
	(0.8958)				
g	0.0139				
	(0.0006)				

[a]Equilibrium wage is estimated at $w=190.239$. Estimation is performed for given $\sigma=1$.

The overall creation and destruction rate is estimated at an annual rate of $\delta=0.07$. The implied average lifespan of a product is consequently about 14 years. The destruction rate is roughly consistent with evidence in Rosholm and Svarer (2000) that the worker flow from employment to unemployment is roughly 10% annually.

The demand distribution is estimated to be close to an exponential distribution with substantial dispersion. The measurement error processes are estimated to produce modest amounts of measurement error noise.

The low type firm will employ 50 manufacturing workers for the average demand realization and its measured wage share is 56%. The higher type firms have lower wage shares. The estimated steady state implies that 4.7% of the labor force is engaged in innovation. The remainder is employed in production.

In the Appendix, we include estimates of the model by major industries. The results confirm that the central qualitative features of the data persist at the disaggregate level. In particular, one observes significant firm productivity

TABLE IV

MODEL FIT (DATA IN TOP ROW, ESTIMATED MODEL IN BOTTOM ROW)

	1992	1997		1992	1996
Survivors	4872.000	3628.000	$\text{Cor}[\frac{Y}{N^*}, \frac{Y_{+1}}{N^*_{+1}}]$	0.476	0.550
	4872.000	3604.315		0.716	0.718
$E[Y]$	26,277.262	31,860.850	$\text{Cor}[\frac{Y}{N^*}, \Delta\frac{Y}{N^*}]$	−0.227	−0.193
	23,023.834	27,252.981		−0.342	−0.352
$\text{Med}[Y]$	13,472.812	16,448.965	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta Y}{Y}]$	−0.120	
	13,352.797	15,382.112		−0.094	
$\text{Std}[Y]$	52,793.105	64,120.233	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta N^*}{N^*}]$	0.119	
	31,013.660	37,224.359		0.123	
$E[W]$	13,294.479	15,705.087	$E[\frac{\Delta Y}{Y}]$	−0.029	
	11,772.341	13,735.062		0.024	
$\text{Med}[W]$	7231.812	8671.939	$\text{Std}[\frac{\Delta Y}{Y}]$	0.550	
	7122.721	8154.757		0.771	
$\text{Std}[W]$	30,613.801	35,555.701	$\text{Cor}[\frac{\Delta Y}{Y}, Y]$	−0.061	
	14,716.584	17,431.300		−0.042	
$E[\frac{Y}{N^*}]$	384.401	432.118	Within	1.015	
	379.930	417.041		0.969	
$\text{Med}[\frac{Y}{N^*}]$	348.148	375.739	Between	0.453	
	346.456	378.623		0.364	
$\text{Std}[\frac{Y}{N^*}]$	205.074	305.306	Cross	−0.551	
	202.134	223.173		−0.446	
$\text{Cor}[Y, W]$	0.852	0.857	Exit	0.084	
	0.927	0.928		0.113	
$\text{Cor}[\frac{Y}{N^*}, N^*]$	−0.018	−0.026			
	−0.031	−0.024			
$\text{Cor}[\frac{Y}{N^*}, Y]$	0.198	0.143			
	0.170	0.176			

dispersion along with a positive correlation between productivity and output size, but a virtually zero correlation between productivity and input size in each industry.

4.7. *Model Fit*

Table IV shows a comparison of the data moments and the simulated moments for the estimated model. The simulated moments are calculated as the average moment realization over all the bootstrap simulations.

4.7.1. *Size Distributions*

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

Size dispersion is impacted by the stochastic birth–death process in products, demand realization variation, and, potentially, measurement error. Model simulation without measurement error ($\sigma_Y^2 = \sigma_W^2 = 0$) yields a reduction in the 1992 Std[Y] estimate to 30,250.77. If, in addition, demand shock variation is eliminated, that is, $o_z = E[\tilde{z}]$, the 1992 Std[Y] estimate is reduced to 22,829.55. Thus, a first order explanation of firm size dispersion is found in the relationship between the type conditional creation rates γ_τ and the rate of total creative–destruction δ. These type conditional relationships determine the distribution of $\tilde{k}_\tau$. Demand realization variation has some impact and measurement error has almost none.

4.7.2. *Productivity and Size Correlations*

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

As mentioned in the previous section, firm type heterogeneity plays an important role in explaining the productivity and size correlations. The positive correlation between value added per worker measure and output size in the data suggests a negative relationship between the firm's wage share and its growth rate. In the model, firm type heterogeneity delivers these relationships

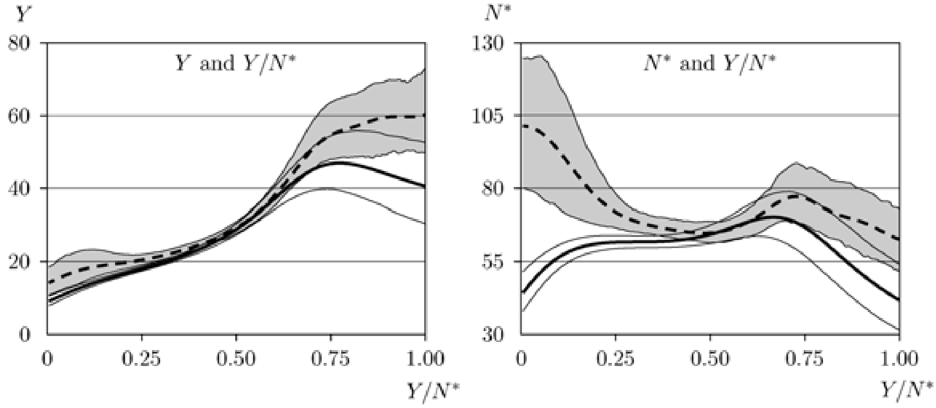

FIGURE 2.—Firm productivity and size, 1992. Value added (Y) measured in 1 million DKK. Labor force size (N^*) measured in efficiency units. Estimated model point estimate and 90% confidence bounds are drawn solid; data are dashed; shaded areas are 90% confidence bounds on data.

by positively correlating high productivity types with high growth rates. The zero correlation between input size and value added per worker is delivered by balancing the labor saving feature of innovations at the product level with the greater growth rate of higher productivity firms.

Measurement error has the potential to explain these correlations as well. The estimation allows for both input and output measurement errors which are estimated at fairly moderate amounts. If the model is simulated without the measurement error ($\sigma_Y^2 = \sigma_W^2 = 0$), the 1992 size–productivity correlations change to $\text{Cor}[Y/N, Y] = 0.142$ and $\text{Cor}[Y/N, N] = 0.009$. Thus, measurement error is estimated to have little contribution to the fit to 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.

4.7.3. *Value Added per Worker Distribution*

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

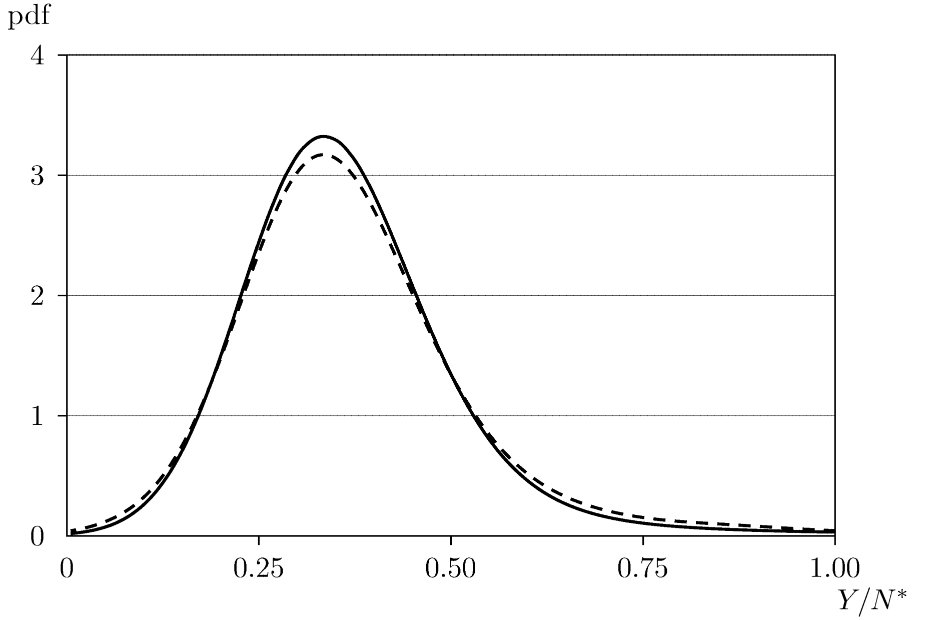

FIGURE 3.—Firm productivity distribution fit, 1992. Value added (Y) measured in 1 million DKK. Labor force size (N^*) measured in efficiency units. Estimated model is solid; data are dashed.

The distribution of empirical firm labor productivity Y/N is explained primarily by type heterogeneity, within type quality improvement dispersion, and the capital share. The κ estimate has a first order impact on the level of Y/N. In particular, for the low type where the profit and creation rates are virtually zero, the κ estimate directly determines the labor share, which then for the low type is 56%.

Measurement error adds to the dispersion measure, but to a much smaller extent than firm type heterogeneity. Simulation without measurement error ($\sigma_Y^2 = \sigma_W^2 = 0$) yields a reduction in the 1992 Y/N standard deviation measure to 169.71.

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

4.7.4. *Cross-Section Shifts From 1992 to 1997*

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

The survivor bias and type heterogeneity account for a substantial part of the right shift in the size distributions. If growth is set to zero by artificially imposing a constant price index, one finds that the $E[Y]$ estimate shifts from 23,023 in 1992 to 25,079 in 1997. Thus, growth accounts for the remainder of the right shift in Table IV.

While growth explains a little less than half of the right shift in the size distributions, it explains almost all of the right shift of the productivity distribution. If growth is set to zero, the $E[Y/N]$ estimate shifts from 379.93 in 1992 to 383.77 in 1997. Thus, growth accounts for about 90% of the estimated right shift of the productivity distribution.

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

4.7.5. *Firm Growth Rate Distribution and Exit Hazards*

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

4.7.6. *Value Added per Worker Persistence and Mean Reversion*

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

In the absence of measurement error, the model estimate implies very high persistence of value added per worker. In this case one obtains 1992 persistence and mean reversion moments of $\mathrm{Cor}[Y/N, Y_{+1}/N_{+1}] = 0.964$ and $\mathrm{Cor}[Y/N, \Delta Y/N] = -0.005$. Adding measurement error reduces the permanence measure and increases the mean reversion measure. Given the one instrument, the estimation has traded off an underestimate of persistence and an overestimate of mean reversion.

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

4.7.7. *Growth Rate and Size (Gibrat's Law)*

Beginning with Gibrat (1931), much emphasis has been placed on the empirical 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 that tests 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

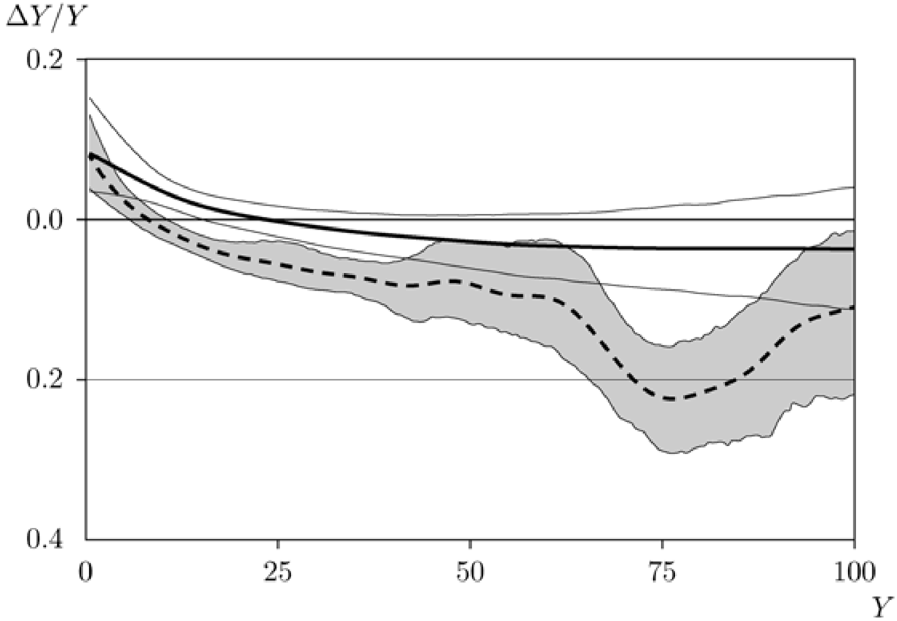

FIGURE 4.—Kernel regression of firm growth rate and size (1992). Value added (Y) measured in 1 million DKK. Model and 90% pointwise confidence bounds are solid; data are dashed. The shaded area is a 90% pointwise confidence bound on the data.

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

At a theoretical level, the model satisfies Gibrat's law in the sense that each firm's expected growth is size independent. But two opposing effects will impact the unconditional size and growth relationship: First, due to selection, larger firms will tend to overrepresent 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 extent, supply shocks introduces an opposite effect: The group of small firms today will tend to overrepresent 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 independent and identically distributed (i.i.d.) over time.

TABLE V

FHK GROWTH DECOMPOSITION AND COUNTERFACTUALS[a]

			Steady State With Entry			
	Data	Point Estimate	Point Estimate	Counter-factual 1	Counter-factual 2	Counter-factual 3
Within	1.0149	0.9676	1.1449	0.8807	0.9268	0.8053
Between	0.4525	0.3665	0.3134	0.0695	0.1124	0.0003
Cross	−0.5514	−0.4465	−0.6531	−0.1442	−0.2334	0.0003
Exit	0.0839	0.1124	0.1338	0.1321	0.1321	0.1321
Entry	—	—	0.0610	0.0618	0.0620	0.0620
Survivor growth rate	0.0165	0.0165	0.0164	0.0164	0.0164	0.0164
Growth rate	—	0.0139	0.0139	0.0139	0.0139	0.0139

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

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 4.[18] Note that the growth statistics include firm death. If firm deaths are excluded and the statistic is calculated only on survivors, the survival bias will steepen the negative relationship between firm size and firm growth both for the data and for the model since the model reproduces the higher exit hazard rate for small firms that is also found in the data.

If the model is estimated without measurement error and demand realization dispersion, one obtains a 1992 firm size and growth correlation of $\text{Cor}[\Delta Y/Y, Y] = 0.04$. This reflects the positive selection effect.

4.7.8. *Foster–Haltiwanger–Krizan Labor Productivity Growth Decomposition*

Table V presents the decomposition results. Because there is no entry in the observed panel, no entry component is recorded in the data and point estimate columns. Furthermore, it is important to keep in mind that the growth in value added per worker is in part a result of survivor selection bias. Therefore, the table presents both a growth measure of the selected sample of surviving 1992 firms and the overall steady state growth rate.

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

The third column of Table V presents the decomposition performed on the estimated model where entry has been included in the simulation.[19] Consequently, each time period is a reflection of the same steady state. Thus, in the third column one can clearly see the survivor selection bias in the productivity growth rate in the data: The steady state annual growth rate is 0.0139, whereas the annual growth rate of the survivor selected sample is 0.0164.

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

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

The model has three major noise processes: measurement error, stochastic demand realizations, and within type stochastic quality improvement realizations. In addition, product age dispersion will add demand variation to the extent that $\sigma \neq 1$. The measurement error and stochastic q realizations turn out to be the most important in terms of explaining the between and cross components. Given the discussion of the decomposition in Section 4.1, it is not surprising that an explanation of nonzero between and cross components requires the existence of measurement error and transitory shocks processes. In the absence of noise processes, both terms will be zero in steady state. The substantial sensitivity of the decomposition to these types of noise was also emphasized by Foster, Haltiwanger, and Krizan (2001) and Griliches and Regev (1995).

The first counterfactual sets measurement error to zero. It is seen that this alone dramatically reduces the magnitude of both the between and the cross components. The second counterfactual turns off both the measurement error and the demand realization noise processes. It is seen that demand realization dispersion has a limited impact on the decomposition.

In addition to setting measurement error and demand noise to zero, the third counterfactual eliminates within type q realization dispersion by deterministically setting each firm's q_τ to match the estimated expected profit π_τ. In this case the theoretical result that the between and cross components equal zero holds.

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

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

5. REALLOCATION AND GROWTH

Aggregate productivity growth as defined in equation (31) is the sum of the contributions of entrants and incumbents, where each term is equal to the sum over types of the average increases in the productivity of innovations relative to the product or service replaced—the product of the innovation frequency and quality improvement per innovation—weighted by relative sizes as reflected in the fraction of product lines supplied by each type. Note that the appropriate empirical counterpart is the traditional growth accounting measure recommended by Petrin and Levinsohn (2005), not the BHC productivity difference index.

Table VI presents the equilibrium steady state annual growth rate implied by the estimated model, $g = 0.0139$. As reported in Table V, the annual growth rate for the sample is 0.0165. This estimate is biased upward because of survivor selection. The traditional growth measure (the TD index) using the value added per worker for continuing firms only is 0.0148. However, the TD index will be biased upward as well in the presence of exit hazard heterogeneity that is negatively correlated with firm growth rates. The model estimate of the steady state growth rate $g = 0.0139$ provides a structural adjustment of the sample selection bias in the data.

The model also permits the identification of the contribution of survival and firm size selection, reflected in differential firm growth rates, to aggre-

TABLE VI

THE PRODUCTIVITY GROWTH RATE AND ITS COMPONENTS (STD ERROR IN PARENTHESES)[a]

	$\sigma = 1.000$	$\sigma = 0.500$	$\sigma = 0.750$	$\sigma = 1.250$
g	0.0139	0.0123	0.0131	0.0144
	(0.0006)	—	—	—
Decomposition (fraction of g)				
Entry/exit	0.2110	0.1876	0.1959	0.2498
	(0.0149)	—	—	—
Residual	0.2608	0.2316	0.2425	0.2788
	(0.0128)	—	—	—
Selection	0.5282	0.5809	0.5615	0.4715
	(0.0270)	—	—	—

[a]Standard deviation estimates obtained by bootstrap. Standard errors have not been calculated for the robustness analysis.

gate growth. Specifically, because the expected productivity of the products created differs across firms, and because these differences are positively associated with differences in expected profitability and, consequently, in creation rates, aggregate growth reflects the selection of more profitable firms by the creative–destruction process. Indeed, equation (31) can be written for the $\sigma = 1$ case as

$$g = \sum_{\tau} \gamma_\tau E[\ln \tilde{q}_\tau]\phi_\tau + \sum_{\tau} \gamma_\tau E[\ln \tilde{q}_\tau](K_\tau - \phi_\tau) + \eta \sum_{\tau} E_\tau[\ln \tilde{q}_\tau]\phi_\tau, \tag{45}$$

where the first term is the contribution to growth of continuing firms under the counterfactual assumption that the share of products supplied by continuing firms of each type is the same as at entry, the second term accounts for type distributional impact of differential firm growth rates after entry, and the third term is the net contribution of entry and exit. In the Appendix we present the growth decomposition for any value of σ. 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 $E[\ln \tilde{q}_\tau] \geq E[\ln \tilde{q}_{\tau'}]$ and $K_\tau - \phi_\tau \geq K_{\tau'} - \phi_{\tau'}$.)

Table VI presents the decomposition estimate. The estimated model implies that the entry/exit component accounts for 21%, the selection component accounts for 53%, and the residual component contributes 26% 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 almost three-quarters of the growth in the modelled economy.

Of course, all growth in the model is a result of worker reallocation. Every time a new innovation is created, workers flow to its supplier from firms with products that have recently become obsolete. This observation does not address the issue of resource allocation across types, however. The selection effect measures the loss in productivity growth that would result if more productive firm types in any given cohort were counterfactually not allowed to increase their resource share relative to that at birth. Because the more productive types in any cohort will gradually gain an ever increasing share of resources of that cohort, as a cohort ages it becomes increasingly selected while also shrinking relative to the size of the overall economy. The model's steady state is the sum of overlapping cohorts with different degrees of selection. The steady state distribution of product lines and the resources required to supply them remains constant over time because firms in the existing cohorts contract at a rate equal to the entry rate of new firms of that type. Holding everything else equal, if the reallocation induced by selection is shut down, the distribution of product lines across types will gradually deteriorate to the point where it equals the entry distribution, and productivity growth will fall by 53%.

Finally, Table VI also includes a check of the robustness of the growth decomposition with respect to the σ estimate. The values for σ cover both the cases of complements and substitutes. It is seen that the growth estimate and the decomposition results are not sensitive to the value of σ. The overall growth estimate is increasing in σ which is a result of the age effect on demand of surviving products. In the complements case, the demand of surviving products increases with age, whereas the opposite is true in the substitutes case. The data moments are based on a selected sample of surviving firms. Hence, the steady state growth rate will adjust to the size of the survivor bias, which is impacted by the age effect.

6. CONCLUDING REMARKS

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

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

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

All growth in our model is attributed to reallocation in the sense that resources must flow from firms that lose markets to innovators that provide new more productive goods and services. We decompose the reallocation component into a net contribution from firm entry and exit, a firm type selection effect, and a residual. The net contribution of entry is 21% of the model's implied growth rate. The selection component, which accounts for 53% of growth, captures the contribution attributable to the fact that resources are reallocated from slow growing less productive firms to fast growing more creative ones in each cohort.

APPENDIX A: DETAILED DERIVATIONS FOR SECTIONS 3.3 AND 3.4

A.1. *The Aggregate Growth Rate*

The sequence $\{A_t(j)/p_t(j)\}$ is a random jump process that is independent of j and characterized by

$$\frac{A_{t+h}(j)}{p_{t+h}(j)} = \begin{cases} \left(\dfrac{q_{J_t(j)+1} m(q_{J_t(j)}, \sigma)}{m(q_{J_t(j)+1}, \sigma)}\right)\left(\dfrac{A_t(j)}{p_t(j)}\right) \\ \qquad \text{with probability } \delta h, \\ \dfrac{A_t}{p_t} \quad \text{with probability } 1 - \delta h \end{cases}$$

for all $h > 0$ sufficiently small, where $A_{t+h}(j) = q_{J_t(j)+1} A_t(j)$ from equation (4) and $p_{t+h}(j) = m(q_{J_t(j)}, \sigma) p_t(j)/m(q_{J_t(j)+1}, \sigma)$ from equation (7) given an innovation. That is, both the quality level and the price of good j change if and only if a new innovation arrives, and the change in both determines the new price per quality unit. By equation (11), the law of large numbers, and the fact that $A_t(j)/p_t(j)$ is identically and independently distributed over the intermediate product continuum, one obtains

$$\begin{aligned} \frac{C_{t+h}}{C_t} &\equiv 1 + \frac{\Delta C_t}{C_t} = \left(\frac{\int_0^1 \left(\frac{A_{t+h}(j)}{p_{t+h}(j)}\right)^{\sigma-1} \alpha(j)^\sigma \, dj}{\int_0^1 \left(\frac{A_t(j)}{p_t(j)}\right)^{\sigma-1} \alpha(j)^\sigma \, dj} \right)^{1/(\sigma-1)} \\ &= \left(\frac{\int_0^1 \left(1 + \frac{\Delta(A_t(j)/p_t(j))}{A_t(j)/p_t(j)}\right)^{\sigma-1} \left(\frac{A_t(j)}{p_t(j)}\right)^{\sigma-1} \alpha(j)^\sigma \, dj}{\int_0^1 \left(\frac{A_t(j)}{p_t(j)}\right)^{\sigma-1} \alpha(j)^\sigma \, dj} \right)^{1/(\sigma-1)} \\ &= \left(E\left\{ \left(1 + \frac{\Delta(A_t(j)/p_t(j))}{A_t(j)/p_t(j)}\right)^{\sigma-1} \right\} \right)^{1/(\sigma-1)} \\ &= \left(1 + h\delta\left[E\left\{\left(\frac{q_{J_t(j)+1} m(q_{J_t(j)}, \sigma)}{m(q_{J(j)+1}, \sigma)}\right)^{\sigma-1}\right\} - 1\right]\right)^{1/(\sigma-1)}, \end{aligned}$$

where $\Delta x_t = x_{t+h} - x_t$ represents the difference operator. Finally,

$$g \equiv \frac{dC_t/dt}{C_t} = \lim_{h\to 0}\left\{\frac{\Delta C_t}{hC_t}\right\} = \lim_{h\to 0}\left\{\frac{(1+h\delta(E\{a^{\sigma-1}\}-1))^{1/(\sigma-1)}-1}{h}\right\}$$
$$= \frac{\delta}{\sigma-1}(E\{a^{\sigma-1}\}-1), \quad \text{where} \quad a = \frac{q_{J_t(j)+1}m(q_{J_t(j)},\sigma)}{m(q_{J_t(j)+1},\sigma)}.$$

A.2. *The Value of the Firm*

The value function can be written as

$$(46) \qquad rV_\tau(\tilde{q}^k,\tilde{z}^k,k) = \max_{\gamma\geq 0}\Bigg\{\sum_{i=1}^{k}\tilde{z}_i\pi(\tilde{q}_i,\sigma) - kwc(\gamma)$$
$$+ k\gamma\big[E_\tau[V_\tau(\tilde{q}^{k+1},\tilde{z}^{k+1},k+1)] - V_\tau(\tilde{q}^k,\tilde{z}^k,k)\big]$$
$$+ k\delta\left[\frac{1}{k}\sum_{i=1}^{k}V_\tau\big(\tilde{q}^{k-1}_{\langle i\rangle},\tilde{z}^{k-1}_{\langle i\rangle},k-1\big) - V_\tau(\tilde{q}^k,\tilde{z}^k,k)\right]$$
$$+ \frac{\partial V_\tau(\tilde{q}^k,\tilde{z}^k,k)}{\partial\tilde{z}^k}\dot{\tilde{z}}^k\Bigg\}.$$

Conjecture the solution

$$(47) \qquad V_\tau(\tilde{q}^k,\tilde{z}^k,k) = \sum_{i=1}^{k}\frac{\tilde{z}_i\pi(\tilde{q}_i,\sigma)}{r+\delta-g(1-\sigma)} + kZ\Psi_\tau.$$

Insert this into (46),

$$rV_\tau(\tilde{q}^k,\tilde{z}^k,k) = \max_{\gamma\geq 0}\Bigg\{\sum_{i=1}^{k}\tilde{z}_i\pi(\tilde{q}_i,\sigma) - kwc(\gamma)$$
$$+ \sum_{i=1}^{k}\frac{\pi(\tilde{q}_i,\sigma)}{r+\delta-g(1-\sigma)}g(1-\sigma)\tilde{z}_i$$
$$+ k\gamma\left[\frac{Z\bar{\pi}_\tau(\sigma)}{r+\delta-g(1-\sigma)} + Z\Psi_\tau\right]$$
$$+ k\delta\Bigg[\frac{k-1}{k}\sum_{i=1}^{k}\frac{\tilde{z}_i\pi(\tilde{q}_i,\sigma)}{r+\delta-g(1-\sigma)}$$

$$+ \frac{(k-1)k}{k} Z\Psi_\tau - \sum_{i=1}^{k} \frac{\tilde{z}_i \pi(\tilde{q}_i, \sigma)}{r+\delta-g(1-\sigma)} - kZ\Psi_\tau \Bigg]\Bigg\}$$

$$= \max_{\gamma \ge 0} \Bigg\{ \sum_{i=1}^{k} \tilde{z}_i \pi(\tilde{q}_i, \sigma) - kwc(\gamma)$$

$$+ g(1-\sigma) \sum_{i=1}^{k} \frac{\tilde{z}_i \pi(\tilde{q}_i, \sigma)}{r+\delta-g(1-\sigma)}$$

$$+ \frac{k\gamma Z \bar{\pi}_\tau(\sigma)}{r+\delta-g(1-\sigma)} + k\gamma Z\Psi_\tau$$

$$- \delta \sum_{i=1}^{k} \frac{\tilde{z}_i \pi(\tilde{q}_i, \sigma)}{r+\delta-g(1-\sigma)} - k\delta Z\Psi_\tau \Bigg\},$$

which can be further manipulated to yield

$$(r+\delta-g(1-\sigma)) \sum_{i=1}^{k} \frac{\tilde{z}_i \pi(\tilde{q}_i, \sigma)}{r+\delta-g(1-\sigma)} + rkZ\Psi_\tau$$

$$= \max_{\gamma \ge 0} \Bigg\{ \sum_{i=1}^{k} \tilde{z}_i \pi(\tilde{q}_i, \sigma) - kwc(\gamma)$$

$$+ \frac{k\gamma Z \bar{\pi}_\tau(\sigma)}{r+\delta-g(1-\sigma)} + k\gamma Z\Psi_\tau - k\delta Z\Psi_\tau \Bigg\}$$

$$\Updownarrow$$

$$(r+\delta)\Psi_\tau = \max_{\gamma \ge 0} \left\{ \frac{\gamma \bar{\pi}_\tau(\sigma)}{r+\delta-g(1-\sigma)} - w\hat{c}(\gamma) + \gamma\Psi_\tau \right\}.$$

Hence, equation (47) is a solution to equation (46) for

$$\Psi_\tau = \max_{\gamma \ge 0} \left\{ \frac{\gamma \nu_\tau - w\hat{c}(\gamma)}{r+\delta} \right\},$$

where

$$\nu_\tau = \frac{\bar{\pi}_\tau(\sigma)}{r+\delta-g(1-\sigma)} + \Psi_\tau.$$

APPENDIX B: GROWTH DECOMPOSITION

The growth decomposition decomposes growth into three sources. Holding everything else equal, the entry/exit contribution is based on the counterfactual that $\gamma_\tau = 0$ for all τ. In this case, one obtains

$$g^{\text{entry}} = \sum_\tau \eta\phi_\tau \frac{E_\tau^{\text{entry}}\left(\dfrac{q'' m(q', \sigma)}{m(q'', \sigma)}\right)^{\sigma-1} - 1}{\sigma - 1},$$

where

$$\begin{aligned} &E_\tau^{\text{entry}}\left(\frac{q'' m(q', \sigma)}{m(q'', \sigma)}\right)^{\sigma-1} \\ &\quad = \int \left(\frac{q''}{m(q'', \sigma)}\right)^{\sigma-1} dF_\tau(q'') \sum_{\tau'} \phi_{\tau'} \int m(q', \sigma)^{\sigma-1}\, dF_{\tau'}(q'). \end{aligned}$$

The residual contribution is the contribution by incumbent firms given the counterfactual that they are represented by the type distribution at the birth of their cohort rather than the steady state product size distribution. Hence, holding everything else equal, the counterfactual is that $\eta = 0$ and $K_\tau = \phi_\tau$. This yields

$$g^{\text{residual}} = \sum_\tau \gamma_\tau \phi_\tau \frac{E_\tau^{\text{residual}}\left(\dfrac{q'' m(q', \sigma)}{m(q'', \sigma)}\right)^{\sigma-1} - 1}{\sigma - 1},$$

where

$$\begin{aligned} &E_\tau^{\text{residual}}\left(\frac{q'' m(q', \sigma)}{m(q'', \sigma)}\right)^{\sigma-1} \\ &\quad = \int \left(\frac{q''}{m(q'', \sigma)}\right)^{\sigma-1} dF_\tau(q'') \\ &\qquad \times \sum_{\tau'} \frac{\phi_{\tau'}\gamma_{\tau'}}{\sum_{\tau''} \phi_{\tau''}\gamma_{\tau''}} \int m(q', \sigma)^{\sigma-1}\, dF_{\tau'}(q'). \end{aligned}$$

The selection contribution is the difference between the total growth contribution from incumbent firms and the residual contribution. Hence,

$$g^{\text{selection}} = g - g^{\text{entry}} - g^{\text{residual}}.$$

For the case where $\sigma \leq 1$, this can also be written simply as

$$g = \sum_\tau \gamma_\tau \phi_\tau E\left[\frac{\tilde{q}_\tau^{\sigma-1}-1}{\sigma-1}\right] + \sum_\tau \gamma_\tau (K_\tau - \phi_\tau) E\left[\frac{\tilde{q}_\tau^{\sigma-1}-1}{\sigma-1}\right] + \eta \sum_\tau \phi_\tau E\left[\frac{\tilde{q}_\tau^{\sigma-1}-1}{\sigma-1}\right], \tag{48}$$

where the first term on the right hand side is the residual contribution, the second term is the selection contribution, and the third term is the entry/exit contribution.

APPENDIX C: STEADY STATE EQUILIBRIUM SOLUTION ALGORITHM

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

$$wc'(\eta/\mu) = \sum_\tau \phi_\tau v_\tau.$$

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

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

$$K_\tau = \frac{\eta\phi_\tau}{\delta - \gamma_\tau} > 0 \quad \forall\tau,$$

$$\delta = \eta + \sum_\tau K_\tau \gamma_\tau > 0,$$

$$g = \sum_\tau (K_\tau + \eta\phi_\tau)\gamma_\tau E[\ln(\tilde{q}_\tau)] \geq 0.$$

PROOF: The evolution of the distribution of products across the n firm types can be written as

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

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

$$\delta = \eta + \sum_{\tau=1}^{n} \gamma_\tau K_\tau \tag{49}$$

and

$$K_\tau = \frac{\eta\phi_\tau}{\delta - \gamma_\tau} = \frac{\eta\phi_\tau}{\eta + \sum_{i\neq\tau}\gamma_i K_i - (1-K_\tau)\gamma_\tau} \tag{50}$$

$$\Updownarrow$$

$$= \frac{\eta\phi_\tau + K_\tau(1-K_\tau)\gamma_\tau}{\eta + \sum_{i\neq\tau}\gamma_i K_i} \equiv \Gamma_\tau(K_\tau, K_{-\tau}), \quad \tau = 1, \dots, n. \tag{51}$$

$\Gamma_\tau(K_\tau, K_{-\tau})$ is a continuous, strictly concave function in $K_\tau \in [0,1]$. Since $\Gamma_\tau(0, K_{-\tau}) = \Gamma_\tau(1, K_{-\tau}) \in (0,1)$ there exists a unique fixed point $\hat{K}_\tau = \Gamma_\tau(\hat{K}_\tau, K_{-\tau})$ where $\hat{K}_\tau \in (0,1)$. Define the mapping $\Omega : [0,1]^n \to (0,1)^n$ such that $\Omega(K) = \Gamma(\Omega(K), K)$. $\Omega(K)$ is continuous in K and since $[0,1]^n$ is a compact and convex set, Brouwer's fixed point theorem can be applied to prove the existence of a fixed point $K^* = \Omega(K^*)$. It follows by $K^*_\tau \in (0,1)\ \forall\tau$, that any fixed point has the property that $\delta > \gamma_\tau\ \forall\tau$.

To prove uniqueness, suppose to the contrary that there exist two distinct fixed points $K^0 = \Omega(K^0)$ and $K^1 = \Omega(K^1)$, $K^0 \neq K^1$. Denote $\delta^i = \eta + \sum_\tau \gamma_\tau K^i_\tau$, $i = 1, 2$. Since K^i is a solution to (51), it must be that $\sum_\tau K^i_\tau = \sum_\tau(\eta\phi_\tau/(\delta^i - \gamma_\tau)) = 1$, $i = 1, 2$. Combined with $\delta^i > \gamma_\tau\ \forall\tau$, $i = 1, 2$, this implies that $\delta^0 = \delta^1$. But it then follows that $K^0_\tau = \eta\phi_\tau/(\delta^0 - \gamma_\tau) = \eta\phi_\tau/(\delta^1 - \gamma_\tau) = K^1_\tau\ \forall\tau$, contradicting the assumption of two distinct fixed points. *Q.E.D.*

The type conditional creation rate choice satisfies

$$w\hat{c}'(\gamma_\tau) = \nu_\tau, \tag{52}$$

where

$$\Psi_\tau = \max_{\gamma\geq 0} \frac{\gamma\nu_\tau - w\hat{c}(\gamma)}{r+\delta},$$

$$\nu_\tau = \frac{\bar{\pi}_\tau(\sigma)}{r+\delta-g(1-\sigma)} + \Psi_\tau.$$

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

$$r + \delta > g(1-\sigma),$$

$$\delta > 0,$$

$$\Psi_\tau \geq 0 \quad \forall\tau,$$

$$g \geq 0.$$

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

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

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

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

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

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

APPENDIX D: THE BHC DECOMPOSITION IN STOCHASTIC STEADY STATE

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

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

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

$$\begin{aligned}\Delta\Theta_t &= \sum_{i \in \widehat{E}_t} s_{it}\theta_{it} + \sum_{i \in \widehat{C}_t} s_{it}\theta_{it} - \sum_{i \in \widehat{C}_t} s_{it-1}\theta_{it-1} - \sum_{i \in \widehat{X}_{t-1}} s_{it-1}\theta_{it-1} \\ &= \sum_{i \in \widehat{C}_t} s_{it-1}\Delta\theta_{it} + \sum_{i \in \widehat{C}_t} \Delta s_{it}\theta_{it-1} + \sum_{i \in \widehat{C}_t} \Delta s_{it}\Delta\theta_{it} + \sum_{i \in \widehat{E}_t} s_{it}\theta_{it} \\ &\quad - \sum_{i \in \widehat{X}_{t-1}} s_{it-1}\theta_{it-1}.\end{aligned}$$

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

Now, state the summations in terms of groups of identical type firms j. Furthermore, define the resource share of the group of type j firms at time t by

$\bar{s}_{jt} = (1/|I_{jt}|)\sum_{i\in I_{jt}} s_{it}$. Similarly, define the average productivity of type j firms at time t by $\bar{\theta}_{jt} = (1/|I_{jt}|)\sum_{i\in I_{jt}} \theta_{it}$. In steady state it must be that $\bar{s}_{jt} = \bar{s}_{jt-1} = \bar{s}_j$, implying that $\Delta\bar{s}_{jt} = 0$. The change in productivity can be written as

$$\begin{aligned}
\Delta\Theta &= \sum_j \sum_{i\in\widehat{E}_t\cap I_{jt}} s_{it}\theta_{it} + \sum_j \sum_{i\in\widehat{C}_t\cap I_{jt}} s_{it}\theta_{it} - \sum_j \sum_{i\in\widehat{C}_t\cap I_{jt-1}} s_{it-1}\theta_{it-1} \\
&\quad - \sum_j \sum_{i\in\widehat{X}_{t-1}\cap I_{jt-1}} s_{it-1}\theta_{it-1} \\
&= \sum_j \sum_{i\in\widehat{C}_t\cap I_{jt}} s_{it}\theta_{it} - \sum_j \sum_{i\in\widehat{C}_t\cap I_{jt}} s_{it}\theta_{it-1} + \sum_j \sum_{i\in\widehat{C}_t\cap I_{jt}} s_{it}\theta_{it-1} \\
&\quad - \sum_j \sum_{i\in\widehat{C}_t\cap I_{jt-1}} s_{it-1}\theta_{it-1} + \sum_j \sum_{i\in\widehat{E}_t\cap I_{jt}} s_{it}\theta_{it} \\
&\quad - \sum_j \sum_{i\in\widehat{X}_{t-1}\cap I_{jt-1}} s_{it-1}\theta_{it-1} \\
&= \sum_j \sum_{i\in\widehat{C}_t\cap I_{jt}} s_{it}\Delta\theta_{it} + \sum_j \sum_{i\in\widehat{C}_t\cap I_{jt}} s_{it}\theta_{it-1} + \sum_j \sum_{i\in\widehat{E}_t\cap I_{jt}} s_{it}\theta_{it-1} \\
&\quad - \left[\sum_j \sum_{i\in\widehat{C}_t\cap I_{jt-1}} s_{it-1}\theta_{it-1} + \sum_j \sum_{i\in\widehat{X}_{t-1}\cap I_{jt-1}} s_{it-1}\theta_{it-1}\right] \\
&\quad + \sum_j \sum_{i\in\widehat{E}_t\cap I_{jt}} s_{it}\theta_{it} - \sum_j \sum_{i\in\widehat{E}_t\cap I_{jt}} s_{it}\theta_{it-1} \\
&= \sum_j \sum_{i\in\widehat{C}_t\cap I_{jt}} s_{it}\Delta\theta_{it} + \sum_j \sum_{i\in I_{jt}} s_{it}\theta_{it-1} - \sum_j \sum_{i\in I_{jt-1}} s_{it-1}\theta_{it-1} \\
&\quad + \sum_j \sum_{i\in\widehat{E}_t\cap I_{jt}} s_{it}\Delta\theta_{it} \\
&= \sum_j \Delta\bar{\theta}_{jt} \sum_{i\in I_{jt}} s_{it} + \sum_j \bar{\theta}_{jt-1} \sum_{i\in I_{jt}} s_{it} - \sum_j \bar{\theta}_{jt-1} \sum_{i\in I_{jt-1}} s_{it-1} \\
&= \sum_j \Delta\bar{\theta}_{jt}|I_{jt}|\bar{s}_{jt} + \sum_j \bar{\theta}_{jt-1}|I_{jt}|\bar{s}_{jt} - \sum_j \bar{\theta}_{jt-1}|I_{jt-1}|\bar{s}_{jt-1}.
\end{aligned}$$

In steady state it must be that

$$\Delta\Theta_t = \sum_j |I_j|\bar{s}_j\Delta\bar{\theta}_{jt} + \sum_j [|I_j|\bar{s}_j - |I_j|\bar{s}_j]\bar{\theta}_{jt-1} = \sum_j |I_j|\bar{s}_j\Delta\bar{\theta}_{jt}.$$

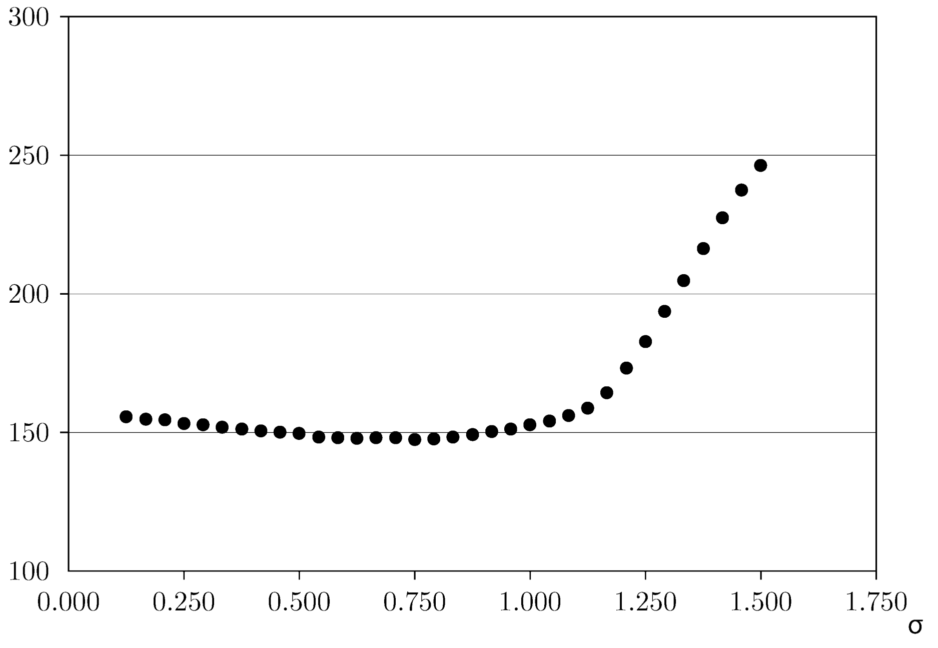

FIGURE 5.—Criterion function.

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

APPENDIX E: IDENTIFICATION OF σ

Figure 5 plots the output of a global search routine over σ. The points in the figure are minimized values of the criterion function for given values of σ. The criterion function takes its minimum for $\sigma = 0.75$.

APPENDIX F: ESTIMATION RESULTS BY INDUSTRY

Tables VII–XV present model estimates by industry. The estimations are done subject to the assumption of $\sigma = 1.0$. We show the results for the three largest industries. The smaller industries are too small to yield statistically interesting results.

TABLE VII

MANUFACTURING—PARAMETER ESTIMATES (STANDARD ERROR IN PARENTHESES)[a]

			$\tau = 1$	$\tau = 2$	$\tau = 3$
c_0/Z	283.3052	ϕ_τ	0.8637	0.0060	0.1303
	(26.7689)		(0.0135)	(0.0542)	(0.0559)
c_1	3.7112	ϕ^*_τ	0.7734	0.0099	0.2167
	(0.0509)		(0.0170)	(0.0907)	(0.0926)
κ	149.2475	ξ_τ	0.0000	0.6115	0.6329
	(2.7979)		(0.0000)	(0.2336)	(0.0637)
Z	20,738.9693	β_q	0.6654	0.6654	0.6654
	(567.5340)		(0.0419)	(0.0419)	(0.0419)
β_Z	0.8067	γ_τ	0.0000	0.0507	0.0510
	(0.0252)		(0.0000)	(0.0033)	(0.0024)
o_Z	3403.7779	K_τ	0.6104	0.0169	0.3727
	(488.6378)		(0.0207)	(0.1579)	(0.1595)
η	0.0478	ν_τ	0.0000	3.9740	4.0529
	(0.0022)		(0.0000)	(0.6156)	(0.4445)
σ^2_Y	0.0243	$\bar{\pi}_\tau(\sigma)$	0.0000	0.3090	0.3142
	(0.0032)		(0.0000)	(0.0431)	(0.0289)
σ^2_W	0.0177	$\text{Med}[\tilde{q}_\tau]$	1.0000	1.3525	1.3648
	(0.0000)		(0.0000)	(0.0802)	(0.1016)
δ	0.0677	$E[\tilde{k}_\tau]$	1.0000	2.1586	2.1798
	(0.0024)		(0.0000)	(0.1702)	(0.1743)
μ	1.6061				
	(0.0777)				
ℓ	56.0141				
	(1.4997)				
g	0.0124				
	(0.0008)				

[a]Equilibrium wage is estimated at $w = 190.660$. Estimation is performed for $\sigma = 1$.

TABLE VIII

MANUFACTURING—MODEL FIT (DATA IN TOP ROW, ESTIMATED MODEL IN BOTTOM ROW)

	1992	1997		1992	1996
Survivors	2051.000	1536.000	$\text{Cor}[\frac{Y}{N^*}, \frac{Y_{+1}}{N^*_{+1}}]$	0.650	0.728
	2051.000	1523.193		0.647	0.651
$E[Y]$	30,149.461	35,803.473	$\text{Cor}[\frac{Y}{N^*}, \Delta\frac{Y}{N^*}]$	−0.024	−0.195
	26,274.870	30,260.835		−0.394	−0.402
$\text{Med}[Y]$	15,117.552	18,858.445	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta Y}{Y}]$	−0.133	
	14,902.763	16,798.498		−0.127	
$\text{Std}[Y]$	56,081.995	69,574.991	$\text{Cor}[\frac{Y}{N^*}, \frac{\Delta N^*}{N^*}]$	0.145	
	33,705.646	39,438.547		0.152	
$E[W]$	15,047.636	17,318.195	$E[\frac{\Delta Y}{Y}]$	−0.035	
	13,382.645	15,164.299		0.002	
$\text{Med}[W]$	8031.273	9531.066	$\text{Std}[\frac{\Delta Y}{Y}]$	0.474	
	7975.387	8944.853		0.472	
$\text{Std}[W]$	24,667.884	27,159.439	$\text{Cor}[\frac{\Delta Y}{Y}, Y]$	−0.073	
	15,720.753	18,016.081		−0.041	
$E[\frac{Y}{N^*}]$	379.047	422.471	Within	0.863	
	374.724	407.562		0.838	
$\text{Med}[\frac{Y}{N^*}]$	347.100	375.300	Between	0.365	
	348.544	377.511		0.270	
$\text{Std}[\frac{Y}{N^*}]$	163.174	226.860	Cross	−0.297	
	140.524	154.961		−0.246	
$\text{Cor}[Y, W]$	0.889	0.855	Exit	0.068	
	0.933	0.933		0.137	
$\text{Cor}[\frac{Y}{N^*}, N^*]$	0.011	−0.003			
	−0.002	0.013			
$\text{Cor}[\frac{Y}{N^*}, Y]$	0.236	0.200			
	0.224	0.239			

TABLE IX

MANUFACTURING—GROWTH DECOMPOSITION (STANDARD ERRORS IN PARENTHESES)

	Point Estimate	Fraction of g
g	0.0124	1.0000
	(0.0008)	—
Entry/exit	0.0031	0.2472
	(0.0003)	(0.0187)
Residual	0.0033	0.2633
	(0.0002)	(0.0118)
Selection	0.0061	0.4894
	(0.0006)	(0.0296)

TABLE X

WHOLESALE AND RETAIL—PARAMETER ESTIMATES (STANDARD ERROR IN PARENTHESES)[a]

			$\tau=1$	$\tau=2$	$\tau=3$
c_0/Z	74.8290	ϕ_τ	0.9259	0.0008	0.0734
	(4.8147)		(0.0075)	(0.0075)	(0.0102)
c_1	2.9503	ϕ^*_τ	0.8705	0.0011	0.1284
	(0.0350)		(0.0096)	(0.0129)	(0.0157)
κ	178.2371	ξ_τ	0.0000	0.5095	2.1195
	(2.7383)		(0.0000)	(0.3410)	(0.1860)
Z	17,582.3421	β_q	0.8193	0.8193	0.8193
	(546.8752)		(0.0721)	(0.0721)	(0.0721)
β_Z	0.9366	γ_τ	0.0000	0.0369	0.0473
	(0.0384)		(0.0000)	(0.0045)	(0.0017)
o_Z	1410.3146	K_τ	0.7548	0.0016	0.2436
	(497.5269)		(0.0138)	(0.0245)	(0.0276)
η	0.0512	ν_τ	0.0000	3.2862	6.8521
	(0.0018)		(0.0000)	(1.0124)	(0.2999)
σ^2_Y	0.0174	$\bar{\pi}_\tau(\sigma)$	0.0000	0.2800	0.5303
	(0.0043)		(0.0000)	(0.0773)	(0.0216)
σ^2_W	0.0196	Med$[\tilde{q}_\tau]$	1.0000	1.3257	2.3551
	(0.3410)		(0.0000)	(0.1791)	(0.1608)
δ	0.0628	$E[\tilde{k}_\tau]$	1.0000	1.6098	2.1881
	(0.0019)		(0.0000)	(0.1515)	(0.1241)
μ	2.6149				
	(0.1153)				
ℓ	44.3084				
	(1.3828)				
g	0.0146				
	(0.0008)				

[a]Equilibrium wage is estimated at $w=187.720$. Estimation is performed for $\sigma=1$.

TABLE XI

WHOLESALE AND RETAIL—MODEL FIT (DATA IN TOP ROW, ESTIMATED MODEL IN BOTTOM ROW)

	1992	1997		1992	1996
Survivors	1584.000	1189.000	$\mathrm{Cor}[\frac{Y}{N^*}, \frac{Y_{+1}}{N^*_{+1}}]$	0.325	0.674
	1584.000	1183.114		0.736	0.743
$E[Y]$	22,952.920	28,386.719	$\mathrm{Cor}[\frac{Y}{N^*}, \Delta\frac{Y}{N^*}]$	−0.195	−0.259
	20,261.525	23,080.814		−0.326	−0.336
$\mathrm{Med}[Y]$	12,757.909	15,288.949	$\mathrm{Cor}[\frac{Y}{N^*}, \frac{\Delta Y}{Y}]$	−0.088	
	12,810.779	14,328.397		−0.097	
$\mathrm{Std}[Y]$	33,400.313	41,409.060	$\mathrm{Cor}[\frac{Y}{N^*}, \frac{\Delta N^*}{N^*}]$	0.189	
	23,880.447	27,936.033		0.162	
$E[W]$	10,696.683	12,712.898	$E[\frac{\Delta Y}{Y}]$	−0.042	
	9453.095	10,563.140		−0.003	
$\mathrm{Med}[W]$	6423.473	7650.564	$\mathrm{Std}[\frac{\Delta Y}{Y}]$	0.425	
	6237.066	6945.005		0.426	
$\mathrm{Std}[W]$	15,360.222	16,802.715	$\mathrm{Cor}[\frac{\Delta Y}{Y}, Y]$	−0.090	
	9909.854	11,174.611		−0.036	
$E[\frac{Y}{N^*}]$	410.234	466.591	Within	1.176	
	403.462	445.301		0.796	
$\mathrm{Med}[\frac{Y}{N^*}]$	373.928	408.244	Between	0.618	
	372.826	408.648		0.236	
$\mathrm{Std}[\frac{Y}{N^*}]$	171.661	278.495	Cross	−0.826	
	167.813	189.967		−0.163	
$\mathrm{Cor}[Y, W]$	0.922	0.914	Exit	0.032	
	0.903	0.899		0.131	
$\mathrm{Cor}[\frac{Y}{N^*}, N^*]$	−0.028	−0.039			
	−0.006	0.013			
$\mathrm{Cor}[\frac{Y}{N^*}, Y]$	0.252	0.188			
	0.285	0.310			

TABLE XII

WHOLESALE AND RETAIL—GROWTH DECOMPOSITION (STANDARD ERROR IN PARENTHESES)

	Point Estimate	Fraction of g
g	0.0146	1.0000
	(0.0008)	—
Entry/exit	0.0036	0.2460
	(0.0003)	(0.0226)
Residual	0.0033	0.2274
	(0.0002)	(0.0119)
Selection	0.0077	0.5267
	(0.0007)	(0.0343)

TABLE XIII

CONSTRUCTION—PARAMETER ESTIMATES (STANDARD ERROR IN PARENTHESES)[a]

			$\tau = 1$	$\tau = 2$	$\tau = 3$
c_0/Z	62.2679	ϕ_τ	0.4733	0.2350	0.2917
	(8.6072)		(0.0481)	(0.0214)	(0.0471)
c_1	3.6300	ϕ^*_τ	0.3627	0.2598	0.3775
	(0.0638)		(0.0408)	(0.0228)	(0.0471)
κ	80.1574	ξ_τ	0.0000	0.0751	0.2047
	(2.5200)		(0.0000)	(0.0150)	(0.0210)
Z	8414.3712	β_q	0.6727	0.6727	0.6727
	(193.1805)		(0.0309)	(0.0309)	(0.0309)
β_Z	1.2233	γ_τ	0.0000	0.0462	0.0583
	(0.1669)		(0.0000)	(0.0044)	(0.0041)
o_Z	4184.8523	K_τ	0.2468	0.2685	0.4847
	(1129.7484)		(0.0298)	(0.0281)	(0.0446)
η	0.0443	ν_τ	0.0000	0.7848	1.8282
	(0.0025)		(0.0000)	(0.1469)	(0.1532)
σ_Y^2	0.0357	$\bar{\pi}_\tau(\sigma)$	0.0000	0.0775	0.1632
	(0.0043)		(0.0000)	(0.0137)	(0.0110)
σ_W^2	0.0198	$\text{Med}[\tilde{q}_\tau]$	1.0000	1.0435	1.1187
	(0.0210)		(0.0000)	(0.0090)	(0.0130)
δ	0.0850	$E[\tilde{k}_\tau]$	1.0000	1.5185	1.8867
	(0.0045)		(0.0000)	(0.0669)	(0.1104)
μ	0.9832				
	(0.0785)				
ℓ	28.7271				
	(0.6870)				
g	0.0104				
	(0.0010)				

[a]Equilibrium wage is estimated at $w = 191.849$. Estimation is performed for $\sigma = 1$.

TABLE XIV

CONSTRUCTION—MODEL FIT (DATA IN TOP ROW; ESTIMATED MODEL IN BOTTOM ROW)

	1992	1997		1992	1996
Survivors	651.000	480.000	$\mathrm{Cor}[\frac{Y}{N^*}, \frac{Y_{+1}}{N^*_{+1}}]$	0.428	0.345
	651.000	476.879		0.368	0.361
$E[Y]$	15,191.354	16,869.551	$\mathrm{Cor}[\frac{Y}{N^*}, \Delta\frac{Y}{N^*}]$	−0.327	−0.560
	12,356.844	14,492.060		−0.548	−0.558
$\mathrm{Med}[Y]$	8688.501	10,711.648	$\mathrm{Cor}[\frac{Y}{N^*}, \frac{\Delta Y}{Y}]$	−0.187	
	8848.573	10,108.425		−0.239	
$\mathrm{Std}[Y]$	31287.564	22,454.655	$\mathrm{Cor}[\frac{Y}{N^*}, \frac{\Delta N^*}{N^*}]$	0.089	
	11,008.879	13,086.746		0.185	
$E[W]$	9973.166	10,594.737	$E[\frac{\Delta Y}{Y}]$	−0.025	
	8007.838	9332.966		0.009	
$\mathrm{Med}[W]$	5785.053	6838.405	$\mathrm{Std}[\frac{\Delta Y}{Y}]$	0.448	
	5905.672	6715.133		0.456	
$\mathrm{Std}[W]$	24,526.438	14,181.147	$\mathrm{Cor}[\frac{\Delta Y}{Y}, Y]$	−0.122	
	6738.723	8002.867		−0.081	
$E[\frac{Y}{N^*}]$	305.075	342.273	Within	0.986	
	303.289	324.261		1.037	
$\mathrm{Med}[\frac{Y}{N^*}]$	286.749	311.509	Between	0.635	
	287.309	307.529		0.383	
$\mathrm{Std}[\frac{Y}{N^*}]$	133.111	173.871	Cross	−0.870	
	94.018	99.614		−0.503	
$\mathrm{Cor}[Y, W]$	0.967	0.922	Exit	0.249	
	0.925	0.927		0.084	
$\mathrm{Cor}[\frac{Y}{N^*}, N^*]$	−0.040	−0.093			
	−0.091	−0.084			
$\mathrm{Cor}[\frac{Y}{N^*}, Y]$	0.131	0.174			
	0.190	0.192			

TABLE XV

CONSTRUCTION—GROWTH DECOMPOSITION
(STANDARD ERROR IN PARENTHESES)

	Point Estimate	Fraction of g
g	0.0104 (0.0010)	1.0000 —
Entry/exit	0.0035 (0.0004)	0.3411 (0.0282)
Residual	0.0044 (0.0006)	0.4249 (0.0225)
Selection	0.0024 (0.0005)	0.2340 (0.0438)

REFERENCES

AGHION, P., AND P. HOWITT (1992): "A Model of Growth Through Creative Destruction," *Econometrica*, 60, 323–351. [1317]

AW, B. Y., X. CHEN, AND M. J. ROBERTS (1997): "Firm-Level Evidence on Productivity Differentials, Turnover, and Exports in Taiwanese Manufacturing," Working Paper 6235, National Bureau of Economic Research. [1336]

BAILY, M., C. HULTEN, AND D. CAMPBELL (1992): "Productivity Dynamics in Manufacturing Plants," *Brookings Papers on Economic Activity, Microeconomics*, 187–249. [1335-1337,1356]

BAILY, M. N., E. J. BARTELSMAN, AND J. HALTIWANGER (1996): "Downsizing and Productivity Growth: Myth or Reality?" *Small Business Economics*, 8, 259–278. [1335,1336]

BALASUBRAMANIAN, N., AND J. SIVADASAN (2008): "What Happens When Firms Patent? New Evidence From US Economic Census Data," Working Paper Series 1090, Ross School of Business. [1318]

BARTELSMAN, E. J., AND P. J. DHRYMES (1994): "Productivity Dynamics: U.S. Manufacturing Plants, 1972–1986," in Board of Governors of the Federal Reserve System (U.S.), Finance and Economics Discussion Series, Vol. 94. [1336]

BARTELSMAN, E. J., AND M. DOMS (2000): "Understanding Productivity: Lessons From Longitudinal Microdata," *Journal of Economic Literature*, 38, 569–594. [1317,1321,1336]

BROWNING, M., M. EJRNÆS, AND J. ALVAREZ (2006): "Modelling Income Processes With Lots of Heterogeneity," Discussion Paper Series, Number 285, University of Oxford, Department of Economics. [1338]

FOSTER, L., J. HALTIWANGER, AND C. KRIZAN (2001): "Aggregate Productivity Growth: Lessons From Microeconomic Evidence," in *New Developments in Productivity Analysis*, ed. by C. R. Hulten, E. R. Dean, and M. J. Harper. Chicago: University of Chicago Press. [1335,1336,1353,1356,1363]

GIBRAT, R. (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. [1350]

GOURIEROUX, C., A. MONFORT, AND E. RENAULT (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8, S85–S118. [1338]

GRILICHES, Z., AND H. REGEV (1995): "Firm Productivity in Israeli Industry: 1979–1988," *Journal of Econometrics*, 65, 175–203. [1335,1336,1353]

GROSSMAN, G. M., AND E. HELPMAN (1991): *Innovation and Growth in the Global Economy*. Cambridge, MA: MIT Press. [1317,1322]

HALL, G., AND J. RUST (2003): "Simulated Minimum Distance Estimation of a Model of Optimal Commodity Price Speculation With Endogenously Sampled Prices," Unpublished Manuscript, Yale University. [1338]

HOROWITZ, J. L. (1998): "Bootstrap Methods for Covariance Structures," *Journal of Human Resources*, 33, 39–61. [1338]

KLETTE, T. J., AND S. KORTUM (2004): "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112, 986–1018. [1317,1318,1321,1325,1329,1347,1356]

LENTZ, R., AND D. T. MORTENSEN (2005): "Productivity Growth and Worker Reallocation," *International Economic Review*, 46, 731–751. [1317,1332]

——— (2008): "Supplement to 'An Empirical Model of Growth Through Product Innovation'," *Econometrica Supplemental Material*, 76, http://www.econometricsociety.org/ecta/Supmat/5997_data and programs.zip. [1317]

LIU, L., AND J. R. TYBOUT (1996): "Productivity Growth in Chile and Colombia: The Role of Entry, Exit, and Learning," in *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*, ed. by M. J. Roberts and J. R. Tybout. Oxford and New York: Oxford University Press for the World Bank, 73–103. [1336]

OLLEY, S. G., AND A. PAKES (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1297. [1336]

PETRIN, A., AND J. LEVINSOHN (2005): "Measuring Aggregate Productivity Growth Using Plant-Level Data," Working Paper 11887, NBER. [1337,1354]

ROSHOLM, M., AND M. SVARER (2000): "Wage, Training, and Job Turnover in a Search-Matching Model," Discussion Paper 223, IZA (Institute for the Study of Labor), Bonn, Germany. [1345]

ROSSI-HANSBERG, E., AND M. L. WRIGHT (2005): "Firm Size Dynamics in the Aggregate Economy," Working Paper Series 11261, NBER. [1350]

SCARPETTA, S., P. HEMMINGS, T. TRESSEL, AND J. WOO (2002): "The Role of Policy and Institutions for Productivity and Firm Dynamics: Evidence From Micro and Industry Data," Working Paper 329, OECD Economics Department. [1336]

SUTTON, J. (1997): "Gibrat's Legacy," *Journal of Economic Literature*, 35, 40–59. [1350]

TYBOUT, J. R. (1996): "Heterogeneity and Productivity Growth: Assessing the Evidence," in *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*, ed. by M. J. Roberts and J. R. Tybout. Oxford and New York: Oxford University Press for the World Bank, 43–72. [1336]

Dept. of Economics, University of Wisconsin–Madison, Social Science Building, 1180 Observatory Drive, Madison, WI 53706-1393, U.S.A. and Centre for Applied Microeconometrics, University of Copenhagen, Copenhagen, Denmark; rlentz@ssc.wisc.edu

and

Dept. of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, U.S.A. and NBER and Institute for the Study of Labor, Bonn, Germany; d-mortensen@northwestern.edu.

Manuscript received August, 2005; final revision received April, 2008.