Innovation, Reallocation and Growth[*]

Daron Acemoglu   Ufuk Akcigit   Harun Alp

Nicholas Bloom   William Kerr

November 14, 2017

Abstract

We build a model of firm-level innovation, productivity growth and reallocation featuring endogenous entry and exit. A new and central economic force is the selection between high- and low-type firms, which differ in terms of their innovative capacity. We estimate the parameters of the model using US Census micro data on firm-level output, R&D and patenting. The model provides a good fit to the dynamics of firm entry and exit, output and R&D. Taxing the continued operation of incumbents can lead to sizable gains (of the order of 1.4% improvement in welfare) by encouraging exit of less productive firms and freeing up skilled labor to be used for R&D by high-type incumbents. Subsidies to the R&D of incumbents do not achieve this objective because they encourage the survival and expansion of low-type firms.

JEL No. E2, L1, O31, O32 and O33

Keywords: entry, growth, industrial policy, innovation, R&D, reallocation, selection.

1 Introduction

Industrial policies that subsidize (often large) incumbent firms, either permanently or when they face distress, are pervasive. They have been the mainstay of government policies in China over the last two and a half decades as well as widely used in Europe (e.g., Owen, 1999; Lerner, 2009).[1] The majority of regional aid in Europe also ends up going to larger firms because they tend to be more effective at obtaining subsidies (Criscuolo et al., 2012). Despite the ubiquity of such policies, their effects are poorly understood. They may encourage incumbents to undertake greater investments, increase productivity and protect employment (e.g., Aghion et al., 2015). But they may also reduce economic growth by slowing down reallocation and even discouraging innovation by both continuing firms and new entrants.[2]

In this paper, we develop a model of endogenous reallocation and innovation with heterogeneous firms to investigate the implications of different types of industrial policies. Our model builds on the endogenous technological change literature (e.g., Romer, 1990; Aghion and Howitt, 1992; Grossman and Helpman, 1991) and in particular, on Klette and Kortum (2004)'s and Lentz and Mortensen (2008)'s analyses of firm-level innovation, but extends these models by incorporating endogenous exit and reallocation. These margins are critical for our investigation of different types of industrial policies as we explain below.

In our model, incumbents and entrants hire skilled labor to perform R&D. Successful innovation enables a firm to take over a leading-edge technology from its current holder, adding to the number of product lines the firm is operating. Because operating a product line entails a fixed cost (which is also in terms of skilled labor), firms may decide to exit some of the product lines in which they have the leading-edge technology when this technology has sufficiently low productivity relative to the equilibrium wage. Finally, firms have heterogeneous (high and low) types, which determine their "innovative capacity". We assume that firm type changes over time, and in particular, high-type firms can become low-type, which is important for accommodating the possibility that firms that have grown large over time may have ceased to be innovative.

The interplay of endogenous exit and innovation and exogenous transitions from high to low type introduces a *selection effect*, determining the composition of active product lines operated by high-type firms. There is positive selection as the fraction of active product lines operated by high-type firms expands over time because low-type firms innovate less and are more likely to

---

[1]The amount spent on bailouts and industrial policy by the European Union in 2010 was about 1.18 trillion euros, which amounts to 9.6% of EU GDP (European Commission, 2011, page 8).

[2]The impact of these policies on the reallocation of resources may be particularly important to take into account. Foster et al. (2001, 2006) report that reallocation, broadly defined to include entry and exit, accounts for around 50% of manufacturing and 90% of US retail productivity growth. These figures probably underestimate the full contribution of reallocation since entrants' prices tend to be below industry average leading to a downward bias in their estimated TFP (Foster et al., 2008). As a result the contribution of reallocation to aggregate productivity growth in the US across all sectors is probably substantially higher. Numerous papers looking at productivity growth in other countries also find a similarly important role for differences in reallocation in accounting for differences in aggregate productivity growth. For example, Hsieh and Klenow (2009, 2014), Bartelsman et al. (2013) and Syverson (2011) discuss how variations in reallocation across countries play a major role in explaining differences in productivity levels.

exit endogenously. Countering this there is also negative selection resulting from the fact that high-type firms transition to low type. The balance of these two forces will determine whether young (and small) firms are more innovative and contribute more to growth.

The key market failure in our model is related to skilled labor. Because of the quality ladder structure (whereby firms build on the quality level of existing leaders), R&D creates positive spillovers on other firms. This implies there will be underinvestment in R&D, and thus lower than socially optimal demand for the factor of production used in R&D, skilled labor. This implies that too high a fraction of skilled workers will be employed in operation activities, and thus all else equal, a welfare-maximizing social planner would like to reallocate skilled labor back to R&D, and especially away from the operations of low-type firms. However, our quantitative analysis will show that, despite the underinvestment in R&D and the emphasis on R&D subsidies in the previous literature, this objective cannot be successfully achieved by R&D subsidies to either incumbents or entrants, because such subsidies would go to both high- and low-type firms. Rather, taxing the continued operation of the incumbents (or alternatively subsidizing exit) is much more powerful in freeing up skilled labor, because such taxes fall disproportionately on low-type firms, which are more likely to be near the exit margin.

Our focus on the reallocation (and misallocation) of R&D inputs, which are critical for productivity growth, is different from that of much of the literature, which emphasizes the reallocation of production inputs. Though in practice there is not a hard line demarcating R&D and production inputs, our separation of these two sets of inputs enables us to highlight our main contribution in a more transparent manner, and emphasizes that misallocation may affect equilibrium growth as well.

Despite the various dimensions of firm-level decisions, heterogeneity, and selection effects, which will prove important in our estimation and quantitative exercises, we show that the model is tractable and that much of the equilibrium can be characterized in closed form (conditional on the wage rate, which does not admit a closed-form solution). This equilibrium characterization then enables the estimation of the model's parameters using simulated method of moments.

The data we use for estimation come from the Census Bureau's Longitudinal Business Database and Census of Manufacturers, the National Science Foundation's Survey of Industrial Research and Development, and the NBER Patent Database. We design our sample around innovative firms that are in operation during the 1987-1997 period. As discussed in greater detail below, the combination of these data sources and our sample design permits us to study the full distribution of innovative firms, which is important when considering reallocation of resources for innovation, and to match the model's focus on R&D-based firms. Our model closely links the growth dynamics of firms to their underlying innovation efforts and outcomes, and we quantify the reallocation of resources necessary for innovation. Our sample contains over 98% of the industrial R&D conducted in the US during this period.

We compute 18 moments capturing key features of firm-level R&D behavior, shipments growth, employment growth and exit, and how these moments vary by firm size and age. We use

these moments to estimate the 8 parameters of our model and 5 parameters are calibrated using conventional values. The model performs well and matches these 18 moments quite closely. In addition, we show that a variety of correlations implied by the model (not targeted in the estimation) are similar to the same correlations computed from the data, bolstering our confidence in the model and our subsequent policy analysis.

We then use our model to study the effects of various counterfactual policies and gain insights about whether substantial improvements in economic growth and welfare are possible. In addition to illustrating the aforementioned effects of different types of policies, our quantitative analysis enables us to compute the socially optimal allocation chosen by a planner who controls R&D investments, and entry and exit decisions of different types of firms. We find that such an allocation would achieve a 2.94% growth rate per annum (relative to 2.26% in equilibrium) and a 4.46% increase in welfare. The social planner achieves this by forcing low-type incumbents to exit at a substantial rate, reducing their R&D, and increasing the R&D of high-type incumbents. These policies induce a strong selection away from low-type firms where the productivity of skilled labor is less than in high-type firms. The socially optimal allocation is not achievable without type-specific taxes, however. Instead, with just (uniform) taxes on operations and subsidies to incumbent R&D, growth can be increased to about 2.54% and welfare can be increased by 1.4%. Optimal policies in this case involve a sizable tax, of about 70%, on the continued operation of incumbents alone, which leverages the selection effect (just like what the social planner was able to achieve directly).

Our baseline empirical analysis uses unweighted moments and focuses on continuously-innovative firms. We show that both our estimation results and quantitative policy conclusions are robust if we instead use employment-weighted moments or also include non-innovative firms in our sample (which however forces us to drop the R&D moments). The results are also not sensitive to excluding mergers and acquisitions related activities. We further document that our results are robust to various variations of the model, including modifying the technology of fixed costs so that it depends on both skilled and unskilled labor; including costs of factor reallocation; generalizing the model to more than two types of firms; and incorporating endogenous supply of skills.

Our paper is linked to a number of different literatures. First, it is most closely related to models of firm innovation and dynamics in general equilibrium pioneered by Klette and Kortum (2004) and Lentz and Mortensen (2008). As already mentioned, we extend these papers in a number of noteworthy dimensions. Most importantly, both papers assume unit elastic demands and no fixed costs of operations, and thus do not feature endogenous exit (obsolescence) of low-productivity products, which removes the issues related to our main focus in this paper—the impact of different types of policies on equilibrium reallocation and selection of firms. In addition, though Lentz and Mortensen allow for firm heterogeneity, this does not affect innovative capacity in their model, ruling out any misallocation of R&D inputs, which is central for our focus and policy analysis. Second, our paper is related to the growing literature on firm dynamics,

reallocation and misallocation[3], but is distinguished by our framework which marries the issue of reallocation to innovation, and by our focus on the reallocation and misallocation of R&D inputs (skilled labor). We are also not aware of any papers in these two literatures that investigate the equilibrium implications of different types of industrial policies, including R&D subsidies. On this last point, some of our emphasis on the distortions that are caused by R&D subsidies are related to Goolsbee (1998), Romer (2001) and Wilson (2009) who point out that R&D subsidies may primarily increase the wages of inelastic inputs (such as R&D workers) rather than spurring additional innovation, and to Akcigit et al. (2016a) who suggest that R&D subsidies may be ineffective when other complementary investments in basic science are not subsidized as well. None of these papers develop a comprehensive framework for studying the effects of different types of policies on selection, reallocation and innovation, nor do they obtain our main substantive conclusions on the ineffectiveness of R&D subsidies and the critical role of taxing incumbents for generating positive selection across firms and productivity growth.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 describes our data and quantitative framework. Section 4 presents our quantified parameter estimates, assesses the model's fit with the data, and provides validation tests. Section 5 examines the impact of counterfactual policy experiments on the economy's innovation and growth. Section 6 reports the results from a number of robustness exercises. The last section concludes, while Appendix A contains some of the proofs omitted from the text, and Appendix B, which is available online, contains additional results.

2 Model

In this section, we introduce our theoretical framework and characterize the stationary balanced growth equilibrium.

2.1 Preferences and Final Good Technology

Our economy is in continuous time and admits a representative household with the following CRRA preferences

$$U_0 = \int_0^\infty \exp(-\rho t) \frac{C(t)^{1-\vartheta} - 1}{1 - \vartheta} dt, \tag{1}$$

where $\rho > 0$ is the discount factor and $C(t)$ is a consumption aggregate given by

$$C(t) = \left(\int_{\mathcal{N}(t)} c_j(t)^{\frac{\varepsilon-1}{\varepsilon}} dj \right)^{\frac{\varepsilon}{\varepsilon-1}}, \tag{2}$$

---

[3]For example, Jovanovic (1982); Hopenhayn (1992, 2012); Hopenhayn and Rogerson (1993); Ericson and Pakes (1995); Davis et al. (2006); Restuccia and Rogerson (2008); Guner et al. (2008); Hsieh and Klenow (2009, 2014); Jones (2011); Peters (2016); Garcia-Macia et al. (2016); and Hsieh et al. (2013).

where $c_j(t)$ is the consumption of product j at time t, $\varepsilon > 1$ is the elasticity of substitution between products, and $\mathcal{N}(t) \subset [0,1]$ is the set of active product lines at time t. The reason why not all products are active at each point in time will be made clear below. Throughout we choose this consumption aggregate as the numeraire.

We assume that the economy is closed, and because R&D and production costs are in terms of labor, we have $c_j(t) = y_j(t)$, where $y_j(t)$ is the amount of product j produced at time t. This also implies that aggregate output (GDP) is equal to aggregate consumption,

$$Y(t) = C(t).$$

There are two types of labor in the economy, skilled and unskilled. Unskilled workers are used in the production of the active products (total labor demand denoted by L^P), while skilled workers perform R&D functions (total labor demand L^{RD}) and are also employed to cover the (fixed) costs of operations, such as management, back-office functions and other non-production work (total labor demand L^F). We assume that the operation of each product requires $\phi > 0$ units of skilled labor.

The representative household has a fixed skilled labor supply of measure L^S and an unskilled labor supply of measure 1, both supplied inelastically. The labor market-clearing condition then equates total labor demand to labor supply for each type of labor:

$$L^P = 1 \quad \text{and} \quad L^F + L^{RD} = L^S.$$

With this specification, the representative household maximizes its utility (1) subject to the flow budget constraint

$$\dot{A}(t) + C(t) \leq r(t) A(t) + w^u(t) + L^S w^s(t), \tag{3}$$

and the usual no-Ponzi condition, $\int_0^\infty \exp(-r(t)t)A(t)dt \geq 0$, where $A(t) = \int_{\mathcal{N}(t)} V_j(t)\,dj$ is the asset position of the representative household, $r(t)$ is the equilibrium interest rate on assets, and $w^s(t)$ and $w^u(t)$ denote skilled and unskilled wages, respectively. In what follows, we focus on stationary equilibria and drop the time subscripts when this causes no confusion.

For future reference, we also note that the representative household utility maximization problem delivers the standard Euler equation,

$$\frac{\dot{C}}{C} = \frac{r - \rho}{\vartheta}. \tag{4}$$

2.2 Intermediate Good Production

Intermediate good (product) j is produced by the monopolist who has the best (leading-edge) technology in that product line, though a single monopolist can own multiple product lines and

5

can produce multiple intermediate goods simultaneously.

At any given point in time, there are two different sets of firms: (i) a set of active firms $\mathcal{F}$ that own at least one product line; and (ii) a set of potential entrants of measure 1 that do not currently own any product line but invest in R&D for innovation.

Consider firm $f \in \mathcal{F}$ that has the leading-edge technology in product j. We assume that, once it hires ϕ units of skilled labor for operation, this firm has access to a linear technology in product line j of the form

$$y_{f,j} = q_{f,j} l_{f,j}, \tag{5}$$

where $q_{f,j}$ is the leading-edge technology of firm f in intermediate good j (which means that firm f has the best technology for this intermediate good), and $l_{f,j}$ is the number of workers it employs for producing this good.

Let us denote by $\mathcal{J}_f$ the set of *active* product lines where firm f has the leading-edge technology and chooses to produce, and by n_f the cardinality of this set, and also define

$$\mathcal{Q}_f \equiv \left\{ q_{f,j_1}, q_{f,j_2}, ..., q_{f,j_n} \right\}$$

as the set of productivities of firm f in product lines in the set $\mathcal{J}_f$. In what follows, we also drop the f subscript when this causes no confusion; for example, we refer to $q_{f,j}$ as q_j.

With this notation, equation (5) implies that the marginal cost of production in line j is simply w^u / q_j. Since all allocations will depend on productivity relative to the unskilled wage, we define the *relative productivity* of a product with productivity q as

$$\hat{q} \equiv \frac{q}{w^u}. \tag{6}$$

We also define the productivity index of the economy as

$$Q \equiv \left(\int_{\mathcal{N}} q_j^{\varepsilon-1} dj \right)^{\frac{1}{\varepsilon-1}}. \tag{7}$$

2.3 Firm Heterogeneity and Dynamics

Firms differ in terms of their innovative capacities. Upon successful entry into the economy, each firm draws its type $\theta \in \left\{ \theta^H, \theta^L \right\}$, corresponding to one of two possible types high (H) and low (L). We assume:

$$\Pr \left(\theta = \theta^H \right) = \alpha \text{ and } \Pr \left(\theta = \theta^L \right) = 1 - \alpha,$$

where $\alpha \in (0,1)$ and $\theta^H > \theta^L > 0$. Firm type impacts innovation as described below. We assume that while low-type is an absorbing state, high-type firms transition to low-type at the exogenous flow rate $\nu > 0$.

In addition to the transition from high to low type, each firm is also subject to an exogenous destructive shock at the rate φ. Once a firm is hit by this shock, its value declines to zero and it

exits the economy.

Innovation by incumbents is modeled as follows. When firm f with type θ_f hires h_f workers for developing a new product, it adds one more product into its portfolio at the flow rate

$$X_f = \theta_f^\gamma n_f^\gamma h_f^{1-\gamma}, \tag{8}$$

where $\gamma \in (0,1)$ and n_f is the number of product lines that firm f owns in total. Suppressing the f subscripts again, this implies the following cost function for R&D

$$C(x,n,\theta) = w^s n x^{\frac{1}{1-\gamma}} \theta^{-\frac{\gamma}{1-\gamma}} \equiv w^s n G(x,\theta), \tag{9}$$

where $x \equiv X/n$ is the "innovation intensity" (innovation effort per product) and $G(x,\theta) \equiv x^{\frac{1}{1-\gamma}} \theta^{-\frac{\gamma}{1-\gamma}}$, defined in (9), denotes the skilled labor requirement for a firm with innovative capacity θ to generate a per product innovation rate of x.

We assume that research is *undirected* across all product lines, meaning that firms do not know ex ante upon which particular product line they will innovate. This implies that their expected return to R&D is the expected value across all product lines $j \in [0,1]$.

When a firm innovates over a product line j, it increases the productivity of this product line j by $\lambda \bar{q}$, where $\lambda > 0$ and

$$\bar{q} \equiv \int_0^1 q_j dj$$

is the *average* quality over all product lines. That is,

$$q_j(t+) = q_j + \lambda \bar{q}, \tag{10}$$

where $t+$ refers to the instant after time t. Note also that equation (10) applies even if product line j is not currently active so that the dynamics of productivity at the product line level are independent of whether the product line in question is currently active or not.

What happens following innovation? The firm with the improved technology in product line j takes over this product line, but in principle, the firm that previously had the leading-edge technology might still compete if the current owner tried to set a very high price. To prevent this possibility, we follow Acemoglu et al. (2012) and assume that there is a two-stage pricing game between any firm that wishes to supply a product $j \in [0,1]$, whereby each firm first has to enter and pay a small cost $\epsilon > 0$, and then all firms that have entered simultaneously set prices. We take $\epsilon \to 0$ for simplicity. Since the price setting after entry forces Bertrand competition, the more productive firm will be able to make any sales and profits, and thus only this firm will pay the cost ϵ and enter. But then in equilibrium, the firm with the leading-edge technology can charge the monopoly price, regardless of the productivity gap between itself and the next best technology. This enables us to characterize prices in a simple fashion in the next subsection.

2.4 Equilibrium Prices and Profits

First note that from the utility function in (2), the inverse demand function for active product line $j \in \mathcal{N}$ is

$$p_j = C^{\frac{1}{\varepsilon}} c_j^{-\frac{1}{\varepsilon}}.$$

Given the market structure described in the previous subsection, the firm with the leading-edge technology can act as a monopolist and thus solves the following maximization problem,

$$\pi\left(\hat{q}_j\right) = \max_{c_j \geq 0} \left\{ \left(C^{\frac{1}{\varepsilon}} c_j^{-\frac{1}{\varepsilon}} - \hat{q}_j^{-1} \right) c_j \right\},$$

where we use $\pi\left(\hat{q}_j\right)$ to designate the firm's profit as a function of only its relative quality $\hat{q}_j$ after substituting for the unskilled wage, w^u, from (6). The price and consumption level of intermediate good j follow from this maximization as

$$p_j = \frac{\varepsilon}{\varepsilon - 1} \hat{q}_j^{-1} \text{ and } c_j = \left(\frac{\varepsilon - 1}{\varepsilon} \right)^{\varepsilon} C \hat{q}_j^{\varepsilon}, \tag{11}$$

and equilibrium profits can then be computed as

$$\pi\left(\hat{q}_j\right) = \frac{1}{\varepsilon - 1} \left(\frac{\varepsilon - 1}{\varepsilon} \right)^{\varepsilon} C \hat{q}_j^{\varepsilon - 1}.$$

Since the final good is the numeraire, (2) also implies

$$\left(\int_{\mathcal{N}} p_j^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}} = 1.$$

Substituting c_j from (11) into the production function (2) and integrating over $\mathcal{N}$, we obtain the unskilled wage rate as

$$w^u = \frac{\varepsilon - 1}{\varepsilon} Q, \tag{12}$$

where Q is given in (7).

2.5 Entry and Exit

There is a unit measure of potential entrants. Each entrant has access to an R&D technology $G\left(x^{entry}, \theta^E\right)$, where the function G was defined in (9) above and specifies the number of skilled workers necessary for generating an innovation rate of $x^{entry} > 0$. Thus an entrant wishing to achieve an innovation rate of x^{entry} would need to hire

$$h^{entry} = G\left(x^{entry}, \theta^E\right) \tag{13}$$

8

skilled workers. This specification implies that a potential entrant has access to the same R&D technology that an incumbent with innovative capacity θ^E and a single active product would have had.

Following a successful innovation, the entrant improves the productivity of a randomly chosen product line by $\lambda \bar{q}$, and at this point, the initial type of a firm, $\theta \in \{\theta^H, \theta^L\}$ is also realized. This description implies the following optimization problem for entrants:

$$\max_{x^{entry} \geq 0} \left\{ x^{entry} \mathbb{E} V^{entry} \left(\hat{q} + \lambda \bar{q}, \theta \right) - w^s G \left(x^{entry}, \theta^E \right) \right\}, \tag{14}$$

where $\mathbb{E} V^{entry} \left(\cdot \right)$ is the expected value of entry (and the expectation is over the relative productivity $\hat{q}$ of the single product the successful entrants will obtain and firm type $\theta \in \{\theta^H, \theta^L\}$). The maximization in (14) determines the R&D intensity of an entrant. Given that there is a unit measure of potential entrants, x^{entry} is also equal to the total entry flow rate.

Exit (of products and firms) has three causes:

1. There is an exogenous *destructive shock* at the rate $\varphi > 0$, which causes the firm to exit and shut down all its product lines.

2. There will be *creative destruction*, because of innovation by other firms replacing the leading-edge technology in a particular product line.

3. There will be *endogenous obsolescence*, meaning that firms will voluntarily shut down some product lines because they are no longer sufficiently profitable relative to the fixed cost of operation.

Due to the first and third factors, the measure of inactive product lines will be positive.

2.6 Value Functions

We normalize all the growing variables by $Q(t)$ to keep the stationary equilibrium values constant. Let us denote the normalized value of a generic variable X by $\tilde{X}$. Let τ denote the *average creative destruction rate* which is endogenously determined in equilibrium. Then the stationary equilibrium value function for a low-type firm can be written as

$$r\tilde{V}_l \left(\hat{\mathcal{Q}} \right) = \max \left\{ 0, \max_{x \geq 0} \left[\sum_{\hat{q} \in \hat{\mathcal{Q}}} \left[\begin{array}{c} \tilde{\pi} \left(\hat{q} \right) - \tilde{w}^s \phi + \\ \tau \left[\tilde{V}_l \left(\hat{\mathcal{Q}} \backslash \{\hat{q}\} \right) - \tilde{V}_l \left(\hat{\mathcal{Q}} \right) \right] + \frac{\partial \tilde{V}_l \left(\hat{\mathcal{Q}} \right)}{\partial \hat{q}} \frac{\partial \hat{q}}{\partial w^u} \frac{\partial w^u}{\partial t} \\ -n \tilde{w}^s G \left(x, \theta^L \right) \\ +nx \left[\mathbb{E} \tilde{V}_l \left(\hat{\mathcal{Q}} \cup \{\hat{q} + \lambda \bar{q}\} \right) - \tilde{V}_l \left(\hat{\mathcal{Q}} \right) \right] \\ +\varphi \left[0 - \tilde{V}_l \left(\hat{\mathcal{Q}} \right) \right] \end{array} \right] \right] \right\}, \tag{15}$$

where $\hat{\mathcal{Q}} \cup \{\hat{q}_{j'}\}$ denotes the new portfolio of the firm after successfully innovating in product line j'. Similarly $\hat{\mathcal{Q}} \backslash \{\hat{q}_j\}$ denotes the loss of a product with technology $\hat{q}_j$ from firm f's portfolio

$\hat{\mathcal{Q}}$ due to creative destruction.[4]

The value function (15) can be interpreted as follows. Given discounting at the rate r, the left-hand side is the flow value of a low-type firm with a set of product lines given by $\hat{\mathcal{Q}}$. The right-hand side includes the components that make up this flow value. The first line (inside the summation) includes the instantaneous operating profits, minus the fixed costs of operation, plus the change in firm value if any of its products gets replaced by another firm through creative destruction at the rate τ, plus the change in firm value due to the the increase in the economy-wide wage. This last term accounts for the fact that as the wage rate increases, the relative productivity of each of the products that the firm operates declines. The second line subtracts the R&D expenditure by firm f. The third line expresses the change in firm value when the low-type firm is successful with its R&D investment at the rate x. The last line shows the change in value when the firm has to exit due to an exogenous destructive shock at the rate φ.

Similarly, we can write the value function of a high-type firm as

$$r\tilde{V}_h\left(\hat{\mathcal{Q}}\right) = \max\left\{0, \max_{x \geq 0}\left[\begin{array}{c} \sum_{\hat{q} \in \hat{\mathcal{Q}}}\left[\begin{array}{c} \tilde{\pi}\left(\hat{q}\right) - \tilde{w}^s\phi+ \\ \tau\left[\tilde{V}_h\left(\hat{\mathcal{Q}}\backslash\{\hat{q}\}\right) - \tilde{V}_h\left(\hat{\mathcal{Q}}\right)\right] + \frac{\partial\tilde{V}_h(\hat{\mathcal{Q}})}{\partial\hat{q}}\frac{\partial\hat{q}}{\partial w^u}\frac{\partial w^u}{\partial t} \end{array}\right] \\ -n\tilde{w}^sG\left(x,\theta^H\right) \\ +nx\left[\mathbb{E}\tilde{V}_h\left(\hat{\mathcal{Q}}\cup\{\hat{q}+\lambda\bar{\hat{q}}\}\right) - \tilde{V}_h\left(\hat{\mathcal{Q}}\right)\right] \\ +\varphi\left[0 - \tilde{V}_h\left(\hat{\mathcal{Q}}\right)\right] \\ +\nu\left[\tilde{V}_l\left(\hat{\mathcal{Q}}\right) - \tilde{V}_h\left(\hat{\mathcal{Q}}\right)\right] \end{array}\right]\right\}. \tag{16}$$

The major difference from (15) is in the last line, where we incorporate the possibility of a transition to a low-type status at the rate ν. The remaining terms have the same interpretation as (15).

The next lemma shows that the value of each firm can be expressed as the sum of the *franchise values* of each of their product lines, defined as the net present discounted value of profits from a product line (as we will see these franchise values depend on the type of the firm).

Lemma 1 *The value function of a $k \in \{h, l\}$ type firm takes an additive form*

$$\tilde{V}_k\left(\hat{\mathcal{Q}}\right) = \sum_{\hat{q} \in \hat{\mathcal{Q}}} Y_k\left(\hat{q}\right),$$

where $Y_k\left(\hat{q}\right)$ is the franchise value of a product line of relative quality $\hat{q}$ to a firm of type k, and $Y_k\left(\hat{q}\right)$ is nondecreasing and increasing when it is greater than zero. Moreover, there exist thresholds $\hat{q}_{k,\min}$ such that a firm of type k shuts down a product line with relative quality $\hat{q} < \hat{q}_{k,\min}$ (and $Y_k\left(\hat{q}\right) > 0$ when $\hat{q} > \hat{q}_{k,\min}$).

Proof. See the Appendix A. ∎

---

[4]Note that in writing this expression, we have made use of the fact that there is a continuum of products, and thus even for a firm with a large number of product lines, the probability that it will innovate on one of its own products is zero. Consequently, τ is both the average creative destruction rate and the average innovation rate in the economy.

The next lemma characterizes the franchise value of a single product line as the solution to a simple differential equation and the type of the firm with the leading-edge best technology in this product line.

Lemma 2 *The franchise values of a product line of relative productivity $\hat{q}$ to low-type and high-type firms, respectively, are given by the following differential equations*

$$(r + \tau + \varphi)\,Y^l\,(\hat{q}) - \frac{\partial Y^l\,(\hat{q})}{\partial \hat{q}}\frac{\partial \hat{q}}{\partial w^u}\frac{\partial w^u}{\partial t} = \Pi \hat{q}^{\varepsilon-1} - \tilde{w}^s \phi + \Omega^l \text{ if } \hat{q} > \hat{q}_{l,\min} \tag{17}$$

$$Y^l\,(\hat{q}) = 0 \text{ otherwise}$$

and

$$(r + \tau + \varphi)\,Y^h\,(\hat{q}) - \frac{\partial Y^h\,(\hat{q})}{\partial \hat{q}}\frac{\partial \hat{q}}{\partial w^u}\frac{\partial w^u}{\partial t} = \left\{ \begin{array}{c} \Pi \hat{q}^{\varepsilon-1} - \tilde{w}^s \phi + \Omega^h + \\ \nu\left[Y^l\,(\hat{q}) - Y^h\,(\hat{q})\right] \end{array} \right\} \text{ if } \hat{q} > \hat{q}_{h,\min}$$

$$Y^h\,(\hat{q}) = 0 \text{ otherwise}$$

where $\Pi \equiv \frac{1}{\varepsilon-1}\left(\frac{\varepsilon-1}{\varepsilon}\right)^{\varepsilon}$, and

$$\Omega^k \equiv \max_{x \geq 0}\left\{ -\tilde{w}^s G\left(x, \theta^k\right) + x\mathbb{E}Y^k\left(\hat{q} + \lambda\bar{\hat{q}}\right) \right\}, \text{ for } k \in \{L, H\}$$

is the R&D value of a k-type firm. Moreover, the R&D policy function of a k-type firm is

$$x^k = \theta^k\left[\frac{(1-\gamma)\,\mathbb{E}Y^k\left(\hat{q} + \lambda\bar{\hat{q}}\right)}{\tilde{w}^s}\right]^{\frac{1-\gamma}{\gamma}} \text{ for } k \in \{L, H\}. \tag{18}$$

Finally, $\hat{q}_{k,\min}$ is given by

$$\hat{q}_{k,\min} = \left(\frac{\tilde{w}^s \phi - \Omega^k}{\Pi}\right)^{\frac{1}{\varepsilon-1}} \text{ for } k \in \{L, H\}. \tag{19}$$

Proof. This follows from the proof of Lemma 1. ∎

The expressions in this lemma are intuitive. So long as this product line remains active, the firm receives two returns: a flow of profits depending on $\hat{q}$, $\Pi \hat{q}^{\varepsilon-1}$, and an R&D value, denoted by Ω^k for a firm of type k. The R&D value accounts for the fact that the firm can undertake R&D building on the knowledge embedded in this active product line. While operating this product line, the firm also incurs the fixed cost of operation $\tilde{w}^s \phi$. The differential equation also takes into account that the relative productivity of this product line is declining proportionately at the growth rate of the economy, g, reducing profits at the rate $(\varepsilon - 1)\,g$, and that this product line is replaced by a higher productivity one at the rate τ and the firm exits for exogenous reasons at the rate φ, making the effective discount rate $r + \tau + \varphi$. If this product line is not replaced or the firm does not exit by the time its relative productivity reaches $\hat{q}_{k,\min}$ (for a firm of type k), at $\hat{q}_{k,\min}$ it will become "obsolete", providing a boundary condition for the differential equation.

Finally, for high-type firms there is an additional term incorporating the possibility of switching to low-type.

The differential equations in Lemma 2 can be solved explicitly, and in the next proposition, we provide the solution for low-type firms, which is simpler. We present the solution for high-type firms in Appendix A.

Proposition 1 *Let g and $\tilde{w}^s$ be the stationary equilibrium growth rate of the economy and the normalized skilled wage rate, respectively. Moreover, let*

$$F_k(x) \equiv 1 - \left(\frac{\hat{q}_{k,\min}}{\hat{q}} \right)^x.$$

Then, the franchise value of a product line with relative productivity $\hat{q}$ for a low-type firm is

$$Y^l(\hat{q}) = \frac{\Pi \hat{q}^{\varepsilon-1}}{r + \tau + \varphi + (\varepsilon-1)g} F_l \left(\frac{r + \tau + \varphi + (\varepsilon-1)g}{g} \right) + \frac{\Omega^l - \tilde{w}^s \phi}{r + \tau + \varphi} F_l \left(\frac{r + \tau + \varphi}{g} \right),$$

where $\Pi \equiv \frac{1}{\varepsilon-1} \left(\frac{\varepsilon-1}{\varepsilon} \right)^{\varepsilon}$.

Proof. See Appendix A. ∎

Intuitively, the franchise value of a product line can be obtained in closed-form because it is given by a combination of two forces: a proportional decline in the value of a product line as the unskilled wage rate increases (and the relative quality of the product line declines), accounting for the term $(\varepsilon-1)g$, and effective discounting coming from the interest rate, creative destruction and exogenous firm exit, accounting for the term $r + \tau + \varphi$.

2.7 Labor Market and Stationary Equilibrium Distributions

The relative productivity distribution for type-k firms has a stationary equilibrium distribution function, $F_k(\hat{q})$ on $[\hat{q}_{k,\min}, \infty)$. Let the shares of product lines that belong to two different types of firms and inactive product lines be denoted by Φ^h, Φ^l and Φ^{np}, respectively. Naturally,

$$\Phi^h + \Phi^l + \Phi^{np} = 1.$$

Then the labor market-clearing condition for unskilled workers is

$$\int_{\mathcal{N}} l(\hat{q}_j) \, dj = \left(\frac{\varepsilon-1}{\varepsilon} \right)^{\varepsilon} (w^u)^{-\varepsilon} C \int_{\mathcal{N}} q_j^{\varepsilon-1} dj = 1. \tag{20}$$

Using (7), (11) and (12), the previous labor market condition gives

$$Y = C = Q. \tag{21}$$

The labor market-clearing for skilled workers, on the other hand, sets the total demand, made

12

up of demand from entrants (first term) and demand from incumbents (second term), equal to the total supply, L^S:

$$G\left(x^{entry}, \theta^E\right) + \sum_{k \in \{h,l\}} \Phi^k \left[h_k\left(w^s\right) + \phi\right] = L^S. \tag{22}$$

To solve for the labor market-clearing condition, we need to characterize the measures of active product lines Φ^k and the stationary equilibrium productivity distributions conditional on firm type k. These detailed derivations are provided in Lemma 3 in Appendix A.

2.8 Aggregate Growth

Equation (21) shows that aggregate output is equal to the productivity index, Q. Thus the growth rate of aggregate output is given by $g = \dot{Q}/Q$. The following proposition characterizes the growth rate.

Proposition 2 *The growth rate of the economy is equal to*

$$g = \lambda \tau. \tag{23}$$

Proof. See Appendix A. ∎

The intuition for the growth rate in (23) is as in standard quality ladder models, linking growth to the frequency and size of innovations.

Finally we summarize the equilibrium of this economy.

Definition 1 (Stationary Equilibrium) *A stationary equilibrium of this economy is a tuple*

$$\{y_j, p_j, l_j, \tilde{V}_l, \tilde{V}_h, \hat{q}_{h,\min}, \hat{q}_{l,\min}, x^h, x^l, x^{entry}, h^h, h^l, h^{entry}, \Phi^h, \Phi^l, \Phi^{np}, F_l\left(\hat{q}\right), F_h\left(\hat{q}\right), w^s, w^u, g, r\}$$

such that [**i**] *y_j and p_j maximize profits as in (11) and the labor demand l_j satisfies (5); [**ii**] $\tilde{V}_l$ and $\tilde{V}_h$ are given by the low-type and high-type value functions in (15) and (16); [**iii**] $(\hat{q}_{h,\min}, \hat{q}_{l,\min})$ satisfy the threshold rule in (19); [**iv**] x^h and x^l are given by the R&D policy functions in (18) and x^{entry} solves the entrants' problem in (14); [**v**] the skilled worker demands h^h, h^l and h^{entry} satisfy (8) and (13); [**vi**] the stationary equilibrium productivity distributions $\left(\tilde{F}_l\left(\hat{q}\right), \tilde{F}_h\left(\hat{q}\right)\right)$ and the product line shares $\left(\Phi^h, \Phi^l, \Phi^{np}\right)$ satisfy Lemma 3; [**vii**] the growth rate is given by (23); [**viii**] the interest rate satisfies the Euler equation (4); and [**ix**] w^s and w^u are consistent with labor market-clearing for unskilled and skilled workers as given by (20) and (22).*

Though the stationary equilibrium in this model is a relatively complex object, the values for different types of firms can be computed in closed form given the equilibrium wage as shown in Proposition 1. There are no closed-form solutions for the equilibrium wage rate and stationary distributions, but these can be computed numerically. We will also use this computation for the simulated method of moments estimation as outlined in Section 3.2.

13

2.9 Welfare and Distortions

Recall that output and consumption are equal to the productivity index Q, so that the initial level of consumption satisfies $C_0 = Q_0$, where

$$Q_0 = \left(\int_{\mathcal{N}_0} q_{j0}^{\varepsilon-1} dj \right)^{\frac{1}{\varepsilon-1}}.$$

We normalize the initial productivity level of all active product lines to 1, i.e., $q_{j0} = 1$ for all $j \in \mathcal{N}_0$, which implies, $C_0 = Q_0 = \Phi^{\frac{1}{\varepsilon-1}}$, where $\Phi_0 = \Phi_0^h + \Phi_0^l$ is the endogenous measure of active product lines at date $t = 0$. Then welfare can be obtained as

$$U_0\left(C_0, g\right) = \int_0^\infty \exp\left(-\rho t\right) \frac{C_t^{1-\vartheta} - 1}{1 - \vartheta} dt = \frac{1}{1 - \vartheta} \left[\frac{\Phi_0^{\frac{1-\vartheta}{\varepsilon-1}}}{\rho - (1 - \vartheta) g} - \frac{1}{\rho} \right], \tag{24}$$

where the first equality simply repeats the definition of discounted utility from (1), the second equality imposes the assumption that we are in stationary equilibrium (thus implying that we are not evaluating welfare implications of transitioning from one stationary equilibrium to another), and solves the integral using $C_t = C_0 e^{gt}$ and $C_0 = \Phi_0^{\frac{1}{\varepsilon-1}}$.

In comparing welfare in two economies, say with subsidy policies s_1 and s_2, and resulting growth rates $g(s_1)$ and $g(s_2)$ and initial consumption levels $C_0(s_1)$ and $C_0(s_2)$, we compute consumption-equivalent changes in welfare by considering the fraction of initial consumption ξ that will ensure the same discounted utility with the new growth rate as with the initial allocation. More formally, the consumption-equivalent change ξ is given such that

$$U_0\left(\xi C_0\left(s_2\right), g\left(s_2\right)\right) = U_0\left(C_0\left(s_1\right), g\left(s_1\right)\right).$$

It is also useful at this point to note that the decentralized equilibrium is typically inefficient. As in models of endogenous technological change, there is insufficient R&D because firms do not appropriate the full value of new innovations (see, e.g., Acemoglu, 2008, for a discussion). In our model, this lack of appropriation results because future innovations build on the current knowledge stock, as captured by equation (10), and thus current innovations create a positive spillover to future innovators. The resulting underinvestment takes the form of too little employment of skilled workers in R&D, and thus too much employment in operations (covering the fixed costs of active firms).[5] However, this underinvestment does not apply to the two types of

---

[5]Counteracting this lack of full appropriation are two other effects. First, as in other quality ladder models such as Aghion and Howitt (1992), there is a business stealing effect, encouraging firms to undertake R&D in order to capture monopoly profits. Second, the love-for-variety resulting from the imperfect substitution of different varieties means that consumers benefit from having more active products. Nevertheless, these two effects are typically dominated by the lack of full appropriation, which leads to underinvestment in R&D.

We should also note that even though there are monopoly markups in this model, these do not directly distort the allocation, since there is no elastic supply of production inputs.

firms equally. The social value of one more active product is greater in the hands of a high-type firm, because such a firm is more productive in R&D, and thus is more likely to undertake a socially valuable (and under-provided) innovation. Consequently, the social planner would like to allocate more skilled labor to R&D, and to be able to do this, she would need to free up this labor from operations, especially from the operations of low-type firms. We will see below how different policies achieve this objective.

3   Estimation and Quantitative Analysis

To perform the policy experiments described in the Introduction, we first estimate the parameters of our model using simulated method of moments (SMM). In this section, we describe our data set and estimation procedures, and the next two sections provide our results and policy counterfactual experiments.

3.1   Data

We employ the Longitudinal Business Database (LBD), the Census of Manufacturers (CMF), the NSF Survey of Industrial Research and Development (RAD), and the NBER Patent Database (PAT). The LBD and CMF are the backbone for our study. The LBD is a business registry that contains annual employment levels for every private-sector establishment in the United States with payroll from 1976 onward. The CMF is conducted every five years and provides detailed records on manufacturing plant and firm operations (e.g., output). Sourced from US tax records and Census Bureau surveys, these micro-records document the universe of establishments and firms, enabling us to study reallocation, entry/exit, and related firm dynamics.

The Survey of Industrial Research and Development (RAD) is the US government's primary instrument for surveying the R&D expenditures and innovative efforts of US firms. This is an annual or biannual survey conducted jointly by the Census Bureau and NSF. The survey includes with certainty all public and private firms, as well as foreign-owned firms, undertaking over one million dollars of R&D within the US. The survey frame also subsamples firms conducting less than the certainty expenditure threshold. The certainty threshold was raised after 1996 to five million dollars of R&D for future years (before subsequently being lowered after our sample frame). RAD surveys are linked to the LBD's and CMF's operating data through Census Bureau identifiers. These micro-records begin in 1972 and provide the most detailed statistics available on firm-level R&D efforts. In 1997, 3,741 firms reported positive R&D expenditures that sum to $158 billion (Foster and Grim (2010) provide additional details on the data). To complement the RAD, we also match patent data into the Census Bureau data. We employ the individual records of all patents granted by the United States Patent and Trademark Office (USPTO) from January 1975 to May 2009. Each patent record provides information about the invention and the inventors submitting the application. Hall et al. (2001) provide extensive details about these data, and

Griliches (1990) surveys the use of patents as economic indicators of technology advancement. We only employ patents (i) filed by inventors living in the US at the time of the patent application; and (ii) assigned to industrial firms. In 1997, this group comprised about 77 thousand patents. We match these patent data to the LBD using firm name and location matching algorithms.[6]

Our main sample focuses on "continuously-innovative" firms (though we later consider the broader manufacturing sample). We define a firm as "innovative" if it is conducting R&D or patenting within the US. Our operating data come from the years 1987, 1992, and 1997 when the CMF is conducted, and the data are specific to those years. We develop our measures of innovation using five-year windows surrounding these CMF years (e.g., 1985-1989 for the 1987 CMF). These local averages assist with RAD's biannual reporting when it occurs, and they ensure that we include two RAD surveys with the lower certainty threshold for the 1997 CMF group. The local averages also provide a more consistent measure of patent filings, which can be lumpy for firms with few patents. We describe the use of patents in further detail shortly.

The "continuous" part of our sample selection is important and is structured as follows. We only include a firm in our sample if it conducts R&D or patents during the five-year window surrounding each CMF year in which it is operating (i.e., has positive employment and sales in the CMF). Thus, a firm that is in operation in 1987 and 1992 is included in our sample if it is also conducting R&D or patents during 1985-1989 and 1990-1994. Similarly, a firm that is in operation in 1992 and 1997 is included in our sample if it is also conducting R&D or patents during 1990-1994 and 1995-1999. The firm does not need to conduct R&D or patent in every year of the five-year window, but must do one of the two activities at least once.

This selection process has several features to point out. First, the entrants in our sample (i.e., firms first appearing in the 1992 or 1997 CMF) will be innovative throughout their lifecycle until the 1995-1999 period. Second, we do not consider switches into innovation among already existing firms. For example, we exclude firms that are present in the 1987 and 1992 CMF, patent or conduct R&D in the 1990-1994 period, but do not patent or conduct R&D during 1985-1989 (the probability that an existing, non-innovative firm commences R&D or patenting over the ensuing five years, conditional on survival, is only about 1%). Third, and on a similar note, we do not include in our sample firms that cease to be innovative but continue in operation. Exits in our economy are thus defined over firms that patent or conduct R&D until they cease to operate.

Finally, our sample does not condition on innovative activity before 1985-1989. Thus, the incumbents in our sample who were in operation prior to the 1987 CMF may have had some point in their past when they did not conduct R&D or patent. We only require that incumbents be innovative in every period when they are in operation during our sample. This choice allows us to construct a full distribution of innovative firms in the economy, which is important when considering the reallocation of resources for innovation. Of course, this choice is also partly due to necessity as we do not observe the full history of older incumbents. We discuss further below

---

[6] Akcigit and Kerr (2017) discuss the R&D and patent data in greater detail. The patent matching builds upon the prior work of Balasubramanian and Sivadasan (2011) and Kerr and Fu (2008). See also Kogan et al. (2017).

the aggregate implications for reallocation and growth measurement of this design.[7]

We now describe the use of the patenting data. In accordance with our model, the moments below focus on R&D intensities (i.e., inputs into the innovation production function) as well as employment, sales and exit dynamics. We face the challenge that the RAD subsamples firms conducting less than one million dollars in R&D. By contrast, the patent data are universally observed. To provide a more complete distribution, we use patents to impute R&D values for firms that are less than the certainty threshold and not sub-sampled. Thus, our moments combine the R&D and patent data into a single measure that accords with the model. As the R&D expenditures in these sub-sampled cases are very low (by definition), this imputation choice versus treating unsurveyed R&D expenditures as zero expenditures conditional on patenting is not very important.

Overall, our compiled dataset includes innovative manufacturing firms from the years 1987, 1992, and 1997 when the CMF is conducted. A record in our dataset is a firm-year observation that aggregates over the firm's manufacturing establishments. We have 17,055 observations from 9,835 firms. By abstracting from the extensive margin of entry or exit into innovation for continuing firms, all of our moments are consistently defined and well measured in the data. At the same time, our selection procedures provide as complete a distribution of innovative firms as possible, which is important when considering reallocation. Our sample accounts for 98% of industrial R&D conducted during the period. When compared to a single cross-section of data, our sample is slightly more skewed towards larger firms. Specifically, in the average year during our sample period, 22% of the firms conducting R&D or patenting have more than 500 employees. In our sample, 32% of observations have more than 500 employees.

Our main sample thus focuses on the reallocation of resources for innovation and thus excludes firms that do not report R&D or patents, which we define as "non-innovative firms". It is important to place our sample within the overall distribution of economic activity. Our sample of continuously-innovative firms accounts for 2% of firms, 50% of employment, and 64% of sales within manufacturing. The greater share of employment and sales activity than firm counts is because the great majority of small firms are non-innovative. In a similar manner and due to the link of innovation to growth, our sample accounts for a substantial portion of reallocation occurring. Many small firms are not oriented for growth (e.g., Haltiwanger et al., 2013) and thus play a limited role in reallocation. As one statistic, our sample includes 58% and 65% of employment and sales reallocation, respectively, among continuing manufacturing firms between 1987 and 1997. As a second statistic, among firms that were either very small (fewer than 20 employees) or did not exist in 1987, we capture 94% of those that then grew to 10,000+ employees by 1997. We likewise capture 80% of small firms or new entrants that grow to one billion dollars in sales by 1997.

---

[7]Note that it would have been impossible to build a consistent sample for "ever innovative" firms rather than for continuously innovative firms. To see why, consider keeping all of the past records for firms that conduct R&D in 1997. In both 1987 and 1992, this approach would induce a mismeasurement of exit propensities and growth dynamics because a portion of the sample will include firms conditioned on survival until 1997.

Our central moments are firm entry/exit rates, the age and size distribution of firms, transitions across the firm size distribution over time, firm growth rates by age and size, firm innovation intensity by age and size, and entrants' share of employment in the economy. Large firms are defined to be those with more than 200 employees, which is roughly the median firm size in our sample. The age distribution is similarly separated into whether a firm is 0-9 years or 10+ years old. We calculate firm age as the count of years since the firm was first observed in the LBD with positive employment, and we later consider robustness checks that exclude inorganic entry and exit (e.g., spinouts and acquisitions). We define moments related to entry/exit, growth, and age-size distribution transitions as changes between CMF years expressed in per annum terms.[8] Shipments are deflated using the 2009 NBER Productivity Database.[9]

3.2 Computational Algorithm

The model can be solved computationally as a fixed point of the following vector of six aggregate equilibrium variables

$$\left\{ \tilde{w}^s, \Phi^h, \Phi^l, \bar{\hat{q}}, \mathbb{E}[Y^h(\hat{q} + \lambda \bar{\hat{q}})], \mathbb{E}[Y^l(\hat{q} + \lambda \bar{\hat{q}})] \right\}. \tag{25}$$

Our characterization above shows that equilibrium innovation decisions can be determined given these aggregate variables. While the skilled wage $\tilde{w}^s$ directly gives the cost of innovation, the rest of the variables in (25) determine the expected return to innovation. We can solve for the stationary equilibrium by first posing a conjecture for (25), then solving for the individual innovation decisions and then verifying the initial conjecture. Specifically, using the guess for these variables:

1. we compute the innovation rates (x^h, x^l, x^{entry}), R&D values (Ω^h, Ω^l), and growth rate g;

2. using the innovation intensities, we calculate the stationary equilibrium distribution over active/inactive product lines and over values of $\hat{q}$ by using Lemma 3;

3. we check the labor market-clearing conditions using the innovation intensities and the above distributions and compute the equilibrium wage rates from (20) and (22), updating $\tilde{w}^s$;

4. we update the values for $\bar{\hat{q}}$, $\mathbb{E}[Y^h(\hat{q} + \lambda \bar{\hat{q}})]$ and $\mathbb{E}[Y^l(\hat{q} + \lambda \bar{\hat{q}})]$ by using the productivity distribution and Lemma 2.

This procedure gives us (25) as a fixed point and also generates the stationary equilibrium distributions of relative productivities. Note that all these variables are determined at the product-line level. We compute firm-level moments by simulating the evolution of a panel of 2^{17} firms until they reach approximate stationary equilibrium after $10,000$ periods. Each period corre-

---

[8]We measure growth rates relative to base years over the five-year period to allow a direct decomposition to per annum terms. These growth rates are winsorized at their 0.5% and 99.5% values. The patterns are similar when expressing growth relative to the average of base and end years. We then calculate geometric averages over these firm-level growth rates. We similarly winsorize R&D intensities to be conservative.

[9]Though prices in industries related to computers and semiconductors behave differently from those in other parts of manufacturing, we find very similar moments when excluding these industries from our moment calculations.

sponds to 0.02 of a year, and hence the total simulation time comes out to 200 years. At each iteration, firms gain and lose products according to the flow probabilities specified in the model.

3.3 Estimation

We set the discount rate equal to $\rho = 2\%$, which roughly corresponds to an annual discount factor of 97%, and the inverse of the intertemporal elasticity of substitution to $\vartheta = 2$. We choose $L^S = 0.166$ to match the share of managers, scientists and engineers in the workforce in 1990, which is 14.2% ($= 0.166/1.166$). Following Broda and Weinstein (2006), we take the elasticity of substitution between different products to be $\varepsilon = 2.9$.

TABLE 1: CALIBRATED PARAMETERS

#	Parameter	Description	Value
1.	ε	CES	2.900
2.	L^S	Measure of high-skilled workers	0.166
3.	γ	Innovation elasticity	0.500
4.	ϑ	Inverse of the intertemporal elasticity of substitution	2.000
5.	ρ	Discount rate	0.020

Following the microeconometric innovation literature, we choose the elasticity of successful innovation with respect to R&D γ as 0.5. In particular, using count data models, Blundell et al. (2002) report an elasticity of $\gamma = 0.5$, while Hall and Ziedonis (2001) find similar results in a study of the semiconductor industry. Estimates exploiting variations in tax credits also yield similar elasticities. Both studies exploiting over-time variation in the US tax code (e.g., Hall, 1993) and those relying on cross-state variation in R&D tax credits (e.g., Bloom et al., 2002; Wilson, 2009) typically estimate a tax elasticity of R&D around unity. These tax elasticities are equivalent to the R&D elasticity with respect to the scientist wage, w^s, since this is the only cost of R&D in our model. Because $\frac{\%\Delta R\&D}{\%\Delta w^s} = \frac{\gamma-1}{\gamma}$, a unit tax elasticity also implies $\gamma = 0.5$ in our setup.[10]

The remaining 8 parameters, which are listed in Table 2, are estimated with SMM.[11] We compute the model-implied moments from the simulation strategy described above and compare them to the data-generated moments to minimize

$$\min \sum_{i=1}^{18} \frac{|\text{model}(i) - \text{data}(i)|}{\frac{1}{2}|\text{model}(i)| + \frac{1}{2}|\text{data}(i)|},$$

where we index each moment by i. SMM iteratively searches repeatedly across sets of parameter values in the model until the model's moments are as close as possible to the empirical moments.

Our SMM procedure targets the 18 moments outlined in Table 3. These moments center on

---

[10]To see this, substitute the equilibrium innovation choice (18) into R&D cost function (9) to obtain $R\&D = n\theta^k (w^s)^{\frac{\gamma-1}{\gamma}} \left[(1-\gamma) \mathbb{E} Y^k (\hat{q} + \lambda\bar{q})\right]^{\frac{1}{\gamma}}$.

[11]See McFadden (1989) and Pakes and Pollard (1989) for the statistical properties of the SMM estimator.

firm entry (measured through employment shares), exit rates, size transition rates, employment and sales growth rates, and innovation intensities, selected in each case because of their economic importance for the mechanisms of the model. We have a single aggregate moment, the growth of output per worker in our sample of firms, and we give this moment 5 times the weight of the micro-moments to ensure that we are in the ballpark of matching the aggregate growth.

We compute the standard errors of the data moments by bootstrap. Specifically, we draw samples of equal size to our original sample from either the Census Bureau data or from the Census of Populations. We use 1000 iterations in each case. For the firm data, we stratify the sample draws by firm age, size, year and industry. The sample draws are conducted at the firm-year level and retain the firm-specific information like whether the firm is an entrant in that year and its forward growth rates for sales and employment. We recalculate our aggregate moments like entrant shares of employment and overall growth rate for each bootstrap sample. The resulting standard errors are quite similar across a range of techniques, such as removing the firm selection stratification or sampling whole firm histories (i.e., retaining all years of a sampled firm).

Standard errors of the parameter estimates are also computed by bootstrap. We estimate the model parameters 1000 times by targeting the empirical moments that are randomly generated based on the bootstrapped distribution of the data moments, and then derive their standard errors from their distribution across these 1000 estimations.[12]

4 Results

In this section, we present our estimation results and evaluate the fit of our model to various targeted and non-targeted moments in the data.

4.1 Parameter Estimates

Table 2 reports the parameter estimates from our SMM procedure and the bootstrapped standard errors, which are uniformly very small reflecting the size of our micro data.

The estimate of the fixed cost of operation indicates that the ratio of fixed workers to variable production workers is around 13.3%. Our estimates also show that high-type firms are about 26% more innovative than low-type firms ($\theta^H/\theta^L \approx 1.26$). Entrants have a 93% chance of being a high-type firm ($\alpha = 0.93$), and high-type firms face an annual 21% probability of transitioning to low-type ($\nu = 0.206$). This pattern implies a very high degree of negative selection—firms are much more likely to be high-type when young than later in their life cycle. The parameter λ is estimated as 0.132, which implies that an innovation leads to 13.2% increase in quality for an average product line. We also estimate a small exogenous destruction rate, $\varphi = 0.037$. Recall,

---

[12]Due to disclosure restrictions, we cannot use the bootstrapped distribution of the data moments directly. Instead, we generate the data moments from a multivariate normal distribution with mean and covariance matrix that are calculated from bootstrapped data moments.

TABLE 2: PARAMETER ESTIMATES

#	Parameter	Description	Value	St Errors
1.	ϕ	Fixed cost of operation	0.216	(0.012)
2.	θ^H	Innovative capacity of high-type firms	1.751	(0.020)
3.	θ^L	Innovative capacity of low-type firms	1.391	(0.017)
4.	θ^E	Innovative capacity of entrants	0.024	(0.001)
5.	α	Probability of being high-type entrant	0.926	(0.023)
6.	ν	Transition rate from high-type to low-type	0.206	(0.005)
7.	λ	Innovation step size	0.132	(0.010)
8.	φ	Exogenous destruction rate	0.037	(0.001)

however, that the overall rate of firm exit will be higher than this because of endogenous exit due to creative destruction and obsolescence, as we show below.

4.2 Goodness of Fit

Table 3 reports the empirical moments that we target (together with their standard errors) and the predicted values from our model. The solid bars in Figures 1(a)-(e) for the model-implied moments provides a graphical depiction.

TABLE 3: MODEL AND DATA MOMENTS

#	Moments	Model	Data	St Error	#	Moments	Model	Data	St Error
1.	Firm exit (small-young)	0.097	0.107	(0.002)	10.	Sales growth (small-young)	0.101	0.107	(0.006)
2.	Firm exit (small-old)	0.092	0.077	(0.002)	11.	Sales growth (small-old)	0.040	0.024	(0.004)
3.	Firm exit (large-old)	0.036	0.036	(0.001)	12.	Sales growth (large-old)	-0.005	-0.003	(0.002)
4.	Trans. from large to small	0.021	0.010	(0.001)	13.	R&D to sales (small-young)	0.086	0.064	(0.004)
5.	Trans. from small to large	0.038	0.014	(0.001)	14.	R&D to sales (small-old)	0.066	0.059	(0.004)
6.	Prob. of small (cond on entry)	0.848	0.753	(0.005)	15.	R&D to sales (large-old)	0.059	0.037	(0.001)
7.	Emp. growth (small-young)	0.101	0.106	(0.004)	16.	5-year entrant share	0.336	0.393	(0.003)
8.	Emp. growth (small-old)	0.040	0.035	(0.003)	17.	Fixed cost-R&D labor ratio	4.175	5.035	(0.015)
9.	Emp. growth (large-old)	-0.005	-0.005	(0.002)	18.	Aggregate growth	0.023	0.022	(0.007)

Both Table 3 and Figure 1 show a relatively good fit between our model-implied moments and data. Our model replicates salient characteristics of data, including lower exit rates for larger and older firms, similar transition rates across firm sizes and entry/exit, quite similar growth rates for sales and employment by firm size and age bins, and similar R&D/sales intensities by type of firm. The last three economy-wide moments are also well aligned. On the whole, despite the overidentification of matching 18 moments with 8 parameters, the fit is quite good.[13]

---

[13]We do not report tests of the overidentifying restrictions for the usual reason that, given our sample size, standard errors are tiny, and even the most minor deviation from these 18 moments would constitute a rejection of the overidentifying restrictions. At the bottom of this, of course, is the fact that standard errors are computed without allowing for "model misspecification".

FIGURE 1: DATA AND SIMULATED MOMENTS
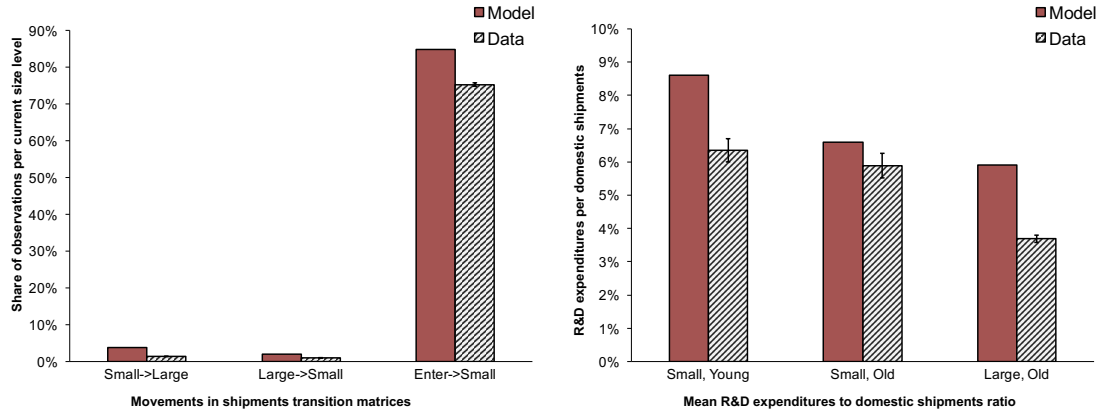
(a) Transition Rates

(b) R&D Intensity
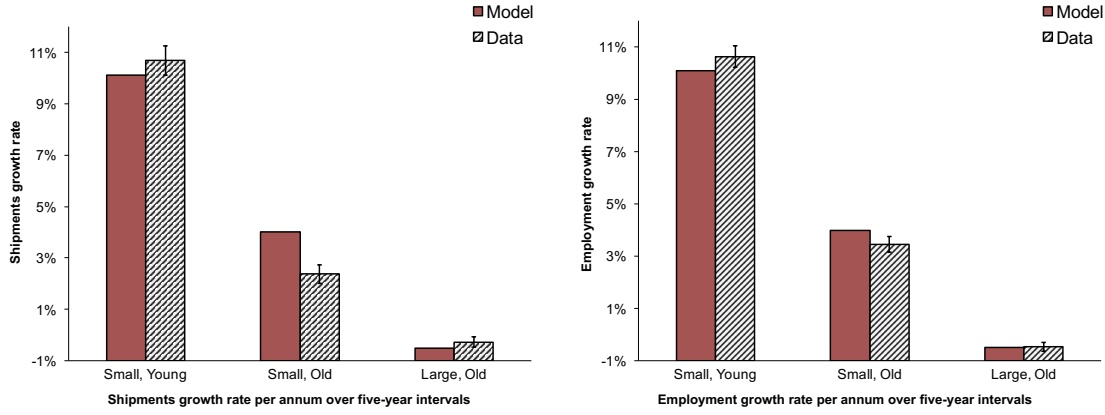
(c) Sales Growth

(d) Employment Growth
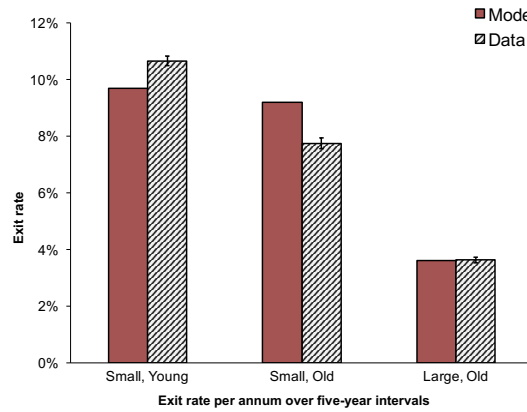
(e) Exit Rates

Table 4 shows the important equilibrium implications in our baseline economy (all numbers in this and subsequent tables, except welfare, are in percentage points). These moments will be used extensively for comparison in our policy analysis in the next section. The equilibrium growth rate is $g = 2.26\%$. This is driven by entry as well as R&D investments by high- and low-type firms. The table shows that the per product innovation rate of high-type firms is about 50% greater than that of low-type firms, which reflects their greater innovative capacity ($x^h = 38.1\%$ vs. $x^l = 25.9\%$). As explained above, the distribution of product lines across high- and low-type firms is determined by different rates of innovation for these two types of firms, different obsolescence rates, and negative selection due to transitions to low-type. Our model finds 6.3% of product lines are held by high-type firms (Φ^h), 55% by low-types firms (Φ^l), and 38.7% are inactive. Together with the 0.51% flow rate of innovations by entrants, these innovation efforts lead to the employment of about 19.9% of all skilled workers in R&D ($L^{R\&D}/L^S$) and an average creative destruction rate of $\tau = 17.2\%$. We also normalize baseline welfare to 100 for ease of comparison in our subsequent policy analysis.

TABLE 4: BASELINE ECONOMY

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
0.51	25.90	38.13	55.04	6.28	147.26	130.33	19.86	17.16	2.26	100

Figure 2 shows the productivity distribution across product lines among high- and low-type firms. An important point to note is that the threshold productivity for high-type firms is lower because of their greater R&D value of operating a product line ($\hat{q}_{h,\min} = 1.30$ vs. $\hat{q}_{l,\min} = 1.47$).

FIGURE 2: PRODUCTIVITY DISTRIBUTION AND SELECTION

(a) Productivity Distribution and Obsolescence

(b) Impact of Incumbent Tax
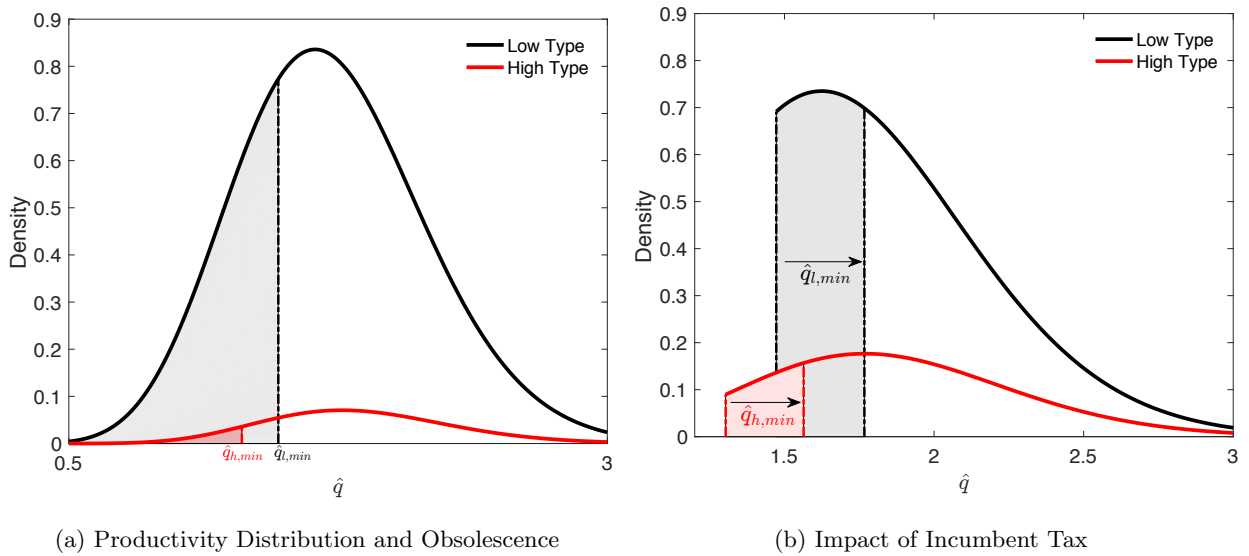
4.3 Non-Targeted Moments

We assess the performance of our model by comparing its implications for a range of non-targeted moments, which capture important economic quantities, but have played no role in our estimation. This strategy thus provides an out-of-sample test of the structure imposed by our model and the values of the parameters we have estimated. Reassuringly, we will see that our model performs fairly well, raising our confidence in the model's ability to provide a good approximation to data and the conclusions that will follow from the policy experiments.

First, Panel A of Table 5 considers persistence in growth rates among firms that survive over the whole sample period. Table 4 shows that about 10% of active product lines are operated by high-type firms in our model, and so we look at persistence in the second period of differences between the top 10% of firms and the remaining 90% in terms of first-period employment growth. For both employment growth and the R&D-to-sales ratio, the model generates patterns consistent with the data, though the differences in the data are somewhat larger in our model. For example, the future employment growth of bottom 90% and top 10% of firms in the data are 0.011 and 0.016, respectively, while they are 0.011 and 0.037 in our model.

TABLE 5: NON-TARGETED MOMENTS

Moments	Model	Data Mean	St Error
Panel A			
Employment growth of bottom 90%	0.011	0.011	(0.004)
Employment growth of top 10%	0.016	0.037	(0.016)
R&D to sales of bottom 90%	0.061	0.038	(0.004)
R&D to sales of top 10%	0.071	0.052	(0.018)
Panel B			
R&D per employee ratio (high to low)	1.737	1.461	(*n/a*)
Patent per employee (high to low)	1.578	1.838	(*n/a*)
Panel C			
Productivity distribution - 75/25: small/young	1.328	1.344	(0.006)
Productivity distribution - 75/25: small/old	1.231	1.311	(0.006)
Productivity distribution - 75/25: large/young	1.150	1.388	(0.017)
Productivity distribution - 75/25: large/old	1.087	1.294	(0.004)
Panel D			
Add a product (or more)	2.9%	8.0%	(0.003)
Drop a product (or more)	2.3%	8.3%	(0.003)
Do both	91.7%	76.8%	(0.005)
Do neither	3.1%	7.0%	(0.003)

Panel B uses the recent Management and Organizational Practices Survey (MOPS) conducted by the U.S. Census Bureau. Bloom et al. (2017) summarize some initial findings from MOPS 2010 survey, which we compare with the implications of our model. These authors group firms into deciles of management scores. Since high-type firms contain 10% of the active product lines in our simulations, we compare the innovation performances our high-type firms to the top-decile of the Bloom et al. (2017) sample. The ratio of the R&D per employee of the top 10% of firms to the bottom 90% is 1.5 in the data, and our model predicts a similar rate, 1.7. Likewise, the patent per employee ratio is 1.8 in the data versus 1.6 in our model. All in all, the model performs fairly well with respect to these non-targeted comparisons, which is reassuring.

Panel C presents the ratio of productivities at the 75*th* and 25*th* percentiles by size and age group. In the data, these are calculated with 5% fuzzy bands around each percentile point to allow for disclosure. Both the model and data exhibit similar productivity distributions within each size and age category, even though these distributions were not targeted in our estimation.

Panel D reports the fractions of firms that gained at least one product without losing any in a year, lost at least one product without gaining any, both gained and lost at least one product and neither gained nor lost any product. The model and data exhibit similar patterns for these numbers as well.

Finally we compare the product line distribution that is generated from our model to its empirical counterpart in Figure 3.

FIGURE 3: PRODUCT LINE DISTRIBUTION
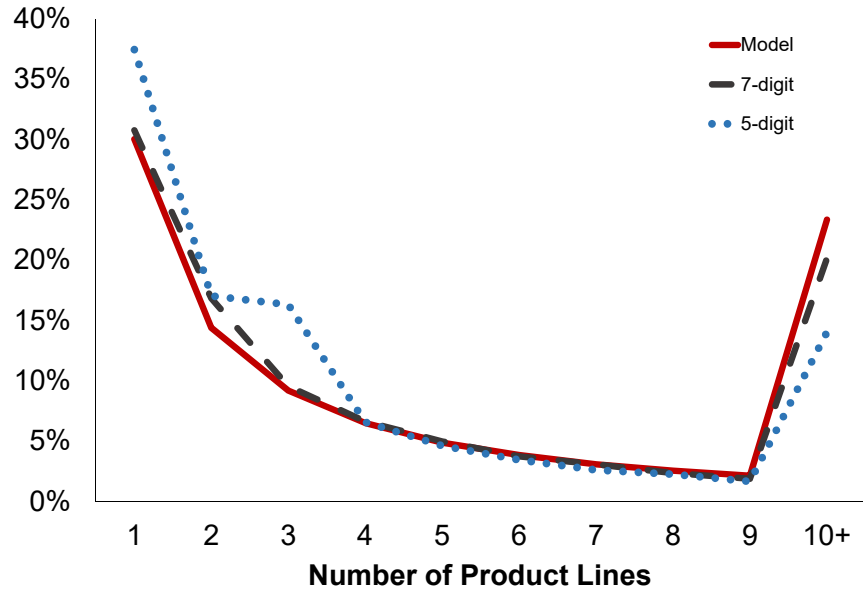
Product information for firms is taken from the Product Trailers to the Census of Manufacturers. Our model generates a product line distribution that is almost identical to the 7-digit

product distribution in the data. In addition, we plot 5-digit product distribution which is not too different from our model either. Panel D of Table 5 also shows that the unweighted rate at which firms add and drop products over five year periods in the model and data are reasonably aligned and in accordance with Bernard et al. (2010). This comparability for product count distributions and firm-level adjustments is encouraging since information on these product line distributions are not used in our estimation.

We next follow Foster et al. (2001), Bartelsman and Doms (2000), and Lentz and Mortensen (2008) and perform a simple growth decomposition according to the following identity,

$$\Delta\Theta_t = \underbrace{\sum_{i\in\mathbb{C}_t} s_{it-1}\Delta\hat{Y}_{it}}_{within} + \underbrace{\sum_{i\in\mathbb{C}_t} \left(\hat{Y}_{it-1}-\Theta_{t-1}\right)\Delta s_{it}}_{between} + \underbrace{\sum_{i\in\mathbb{C}_t} \Delta s_{it}\Delta\hat{Y}_{it}}_{cross} + \underbrace{\sum_{i\in\mathbb{E}_t} s_{it}\left(\hat{Y}_{it}-\Theta_{t-1}\right)}_{entry} - \underbrace{\sum_{i\in\mathbb{X}_t} s_{it-1}\left(\hat{Y}_{it-1}-\Theta_{t-1}\right)}_{exit},$$

where Y_{it} is value added for firm i at time t, N_{it} is the number of employees of the firm, and $\mathbb{C}_t$, $\mathbb{E}_t$ and $\mathbb{X}_t$, respectively, denote the set of continuing, entering and exiting firms. In addition, we have $\Theta_t = \sum_i s_{it}\hat{Y}_{it}$, $\hat{Y}_{it} = Y_{it}/N_{it}$, and $s_{it} = N_{it}/\sum_i N_{it}$. We report each of the five components in this decomposition both in the data and from our model in Table 6.

TABLE 6: DECOMPOSITION

	Model	**Data**	**St Error**
Within Share	0.607	0.999	(0.176)
Between Share	-0.024	-0.049	(0.057)
Cross Share	0.239	-0.305	(0.176)
Entry Share	0.175	0.192	(0.062)
Exit Share	0.003	0.164	(0.033)
Net Entry Share	0.178	0.356	(0.074)
10-year Cumulative Growth	0.254	0.261	(0.051)

The model and data both show the largest component of growth coming from the within-firm labor productivity growth term, and the signs and magnitudes mostly line up over the components too. The one exception is in the cross term, where the model finds growing share of employment connected to growing labor productivity, whereas the data finds a negative correlation. This discrepancy is not a very robust feature, however; the cross term becomes positive, for example, when we look at the 1987-1992 subsample.

5   Policy Experiments and Efficiency

In this section, we perform counterfactual policy analysis to gain insight on both the implications of different types of industrial policies and the form of optimal policy in this economy. Before

turning to our analysis of optimal policy, we first show how incumbent R&D subsidies, fixed cost subsidies and entry subsidies impact the equilibrium.[14]

5.1 Incumbent R&D Subsidy

The results from subsidizing the R&D of incumbents are shown in Table 7. As in other policy experiments, we choose the subsidy rate to be equivalent to 1% of GDP, and also show the key equilibrium objects from our baseline economy (from Table 4) in Panel A for comparison.

TABLE 7: INCUMBENT R&D SUBSIDY

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
				Panel A. Baseline						
0.51	25.90	38.13	55.04	6.28	147.26	130.33	19.86	17.16	2.26	100.00
				Panel B. 1% of GDP ($s_i = 14\%$)						
0.46	27.38	40.71	53.03	6.85	150.95	133.76	21.74	17.77	2.34	100.63

A subsidy equivalent to 1% of GDP translates into a 14% subsidy on R&D spending of continuing firms. Unsurprisingly, this leads to higher R&D by these incumbents. Low-type incumbents increase their innovation rate from 25.9% to 27.4%, while high-type incumbents go from 38.1% to 40.7%—both of these are about 6% higher than the baseline. However, the overall impact on innovation and growth is much less than this direct effect. The average rate of creative destruction, τ, increases only by 3%, for instance. This is for two reasons. First, at a given skilled wage, greater incumbent R&D would increase creative destruction and thus discourage entry. Second, and more important, the greater demand for skilled workers from incumbent R&D increases the skilled wage. This reduces R&D by entrants from 0.51% to 0.46%, and also modestly reduces the amount of skilled labor allocated to operations and thus raises the ratio of skilled labor employed in R&D from 19.9% to 21.7%. The overall result is a modest increase in growth 2.26% to 2.34%, and aggregate welfare goes up by 0.6% (in consumption equivalent terms).

5.2 Subsidy to Operating Costs

We next consider an industrial policy subsidizing the continued operation of incumbents by subsidizing their fixed costs of operations $w^s\phi$, which approximates policies that support large firms that are in economic trouble.[15] A subsidy equivalent to 1% of GDP in this case corresponds

---

[14]To focus on the key economic implications of our model in the clearest fashion, we abstract from the costs of raising taxes. In any case, we will see below that optimal policies typically involve taxes on the operation of continuing firms, thus raising rather than reducing revenues to tax authorities.

[15]Or equivalently, their exit is taxed or some combination thereof. We consider subsidies or taxes on the fixed cost of operations rather than on all costs or on accounting profits, because these alternative policies would also affect markups, partly confounding the main effect we are interested in. All the same, such subsidies or taxes have broadly similar impacts.

to a 4% subsidy on the fixed costs of operation of continuing firms.

Panel B of Table 8 shows that this subsidy discourages exit, increasing the fraction of active product lines (Panel A again gives the baseline for comparison). It also leads to modest declines in the innovation rates of entrants, low-type incumbents and high-type incumbents. In particular, because now more firms are operating, the demand for skilled labor increases, the skill wage goes up and fewer skilled workers perform R&D (the fraction of skilled workers allocated to R&D goes down modestly, from 19.9% to 19.4%). Because low-type firms are overrepresented among those at the margin of obsolescence (recall Figure 2), this policy also induces further negative selection: the share of product lines operated by low-type firms in the economy increases from 55.0% to 55.6%, while the share operated by high-type firms declines from 6.3% to 6.1%. As a consequence of all of these negative effects, the growth rate of the economy declines from 2.26% to 2.24%, and aggregate welfare declines by 0.2%.

TABLE 8: OPERATION SUBSIDY

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
				Panel A. Baseline						
0.51	25.90	38.13	55.04	6.28	147.26	130.33	19.86	17.16	2.26	100.00
				Panel B. 1% of GDP ($s_o = 4\%$)						
0.50	25.57	37.58	55.63	6.05	146.27	129.30	19.39	17.00	2.24	99.80

In sum, a subsidy to the operating costs of incumbents reduces growth and welfare because it causes a negative selection effect, increasing the share of product lines controlled by low-type firms, as low-type firms tend to benefit more from this subsidy which is directed to low-productivity product lines.

5.3 Entry Subsidy

Finally, for comparison, we also consider the implications of an entry subsidy equivalent to 1% of GDP. The results are reported in Table 9.

TABLE 9: ENTRY SUBSIDY

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
				Panel A. Baseline						
0.51	25.90	38.13	55.04	6.28	147.26	130.33	19.86	17.16	2.26	100.00
				Panel B. 1% of GDP ($s_e = 65\%$)						
1.35	24.59	35.08	49.50	10.16	151.30	138.46	22.02	17.09	2.25	98.36

The direct effect of the subsidy is to increase entry. In Panel B we see that the innovation effort of entrants increases from 0.51% to 1.35%, but now there is a decline in the innovation rates of

continuing firms. The total effect is a modest reduction in the average creative destruction rate of the economy from 17.2% to 17.1%. This in turn leads to slightly lower growth and aggregate welfare.

5.4 Social Planner

The results of the previous subsection show only small effects from subsidies to incumbent R&D, entrant R&D and operations. We will see now, however, that the social planner can significantly increase welfare. Since we are not interested in monopoly distortions per se, we restrict the social planner to the same production and pricing decisions as the equilibrium, and only allow her to control the entry, exit and R&D margins of different firms. It is straightforward to see that the social planner will choose the same per product R&D for all high-type firms and also the same R&D for all low-type product lines. Then, we can represent the problem of the planner as choosing $\{\hat{q}_{h,\min}, \hat{q}_{l,\min}, x^h, x^l\}$ to maximize representative household welfare (24) subject to the skilled labor market-clearing condition (22). Table 10 summarizes the allocation implied by social planner's choices.

TABLE 10: SOCIAL PLANNER

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
					Panel A. Baseline					
0.51	25.90	38.13	55.04	6.28	147.26	130.33	19.86	17.16	2.26	100.00
					Panel B. Social Planner					
0.61	27.13	45.32	6.07	44.25	238.26	21.76	34.23	22.30	2.94	104.46

The social planner improves growth and welfare quite significantly. Growth increases from 2.26% to 2.94%. Welfare increases by 4.46%, underscoring that the equilibrium was far from optimal in the baseline model, and the limited consequences of the subsidy policies considered so far stemmed from the fact that each was by itself ineffective in triggering a reallocation of resources towards R&D by high-type firms. How the social planner is achieving such a reallocation can also be seen from Table 10, which illustrates the form of the optimal allocation. Most notably, the exit threshold for low-type firms, $\hat{q}_{l,\min}$, increases substantially (from 1.47 to 2.38) whereas the threshold for high-type firms, $\hat{q}_{h,\min}$, actually decreases (from 1.30 to 0.22). The social planner also differentially increases R&D by firm type: high-type incumbents increase from $x^h = 0.38$ to 0.45, while R&D for low-type firms remains essentially unchanged (there is also a modest increase in the entry rate). The combined effect of the large increase in the exit threshold for low-type firms and increased R&D for high-type firms is a significant change in the selection effect—the ratio of high- to low-type firms (Φ^h/Φ^l) increases from 0.11 to 7.29.

Table 11 further dissects how the social planner is improving welfare relative to the baseline economy. Row 3 shows that if the social planner can only change the entry and innovation rates

(keeping the exit thresholds at their baseline equilibrium values, $\hat{q}_{l,\min}$ and $\hat{q}_{h,\min}$), there is essentially no effect on welfare. On the other hand, when she only controls the exit thresholds (keeping the innovation and entry rates at their baseline equilibrium values), she achieves most of the selection gains and can increase welfare by 1.58% in consumption-equivalent terms. Naturally, when the two margins are combined, she can achieve much greater growth and welfare gains as we have seen in Table 10.

TABLE 11: RESTRICTED SOCIAL PLANNER

	x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	g	Wel
1. Baseline	0.51	25.90	38.13	55.04	6.28	147.26	130.33	2.26	100.00
2. Social Planner (SP)	0.61	27.13	45.32	6.07	44.25	238.26	21.76	2.94	104.46
3. SP choosing innovation	0.51	25.65	38.72	54.49	6.87	147.26	130.33	2.26	100.00
4. SP choosing $\hat{q}_{\min}$	0.94	25.90	38.13	39.74	18.92	161.16	29.87	2.43	101.58

5.5 Uniform Optimal Policy

The social planner's allocation discussed in the previous subsection relied on choosing the exit thresholds and R&D rates of different types of firms. In practice, policies cannot be directly conditioned on type (at least not without also specifying relevant incentive compatibility constraints).[16] Motivated by this restriction, in this subsection we study how much of the gap between the baseline equilibrium allocation and the social planner's allocation characterized in the previous subsection can be closed with uniform policies. In Table 12, we start by looking at the optimal choice of each one of the three policies already introduced previously.

Panel A again depicts the baseline equilibrium for comparison. In Panel B, we show that the optimal rate of incumbent R&D subsidy (by itself) would be $s_i = 39\%$, which is higher than what we considered in Table 7 above, but has similar implications. In Panel C, we turn to taxes/subsidies on operations. Here, we see that the optimal policy is a rather large tax (instead of the subsidy considered in Table 8 above). With this optimal tax rate of $s_o = -69\%$,[17] we can obtain a significant increase in growth, achieving $g = 2.54\%$. As with the social planner's allocation, this is made possible by increasing the exit thresholds and generating a significant selection effect—the fraction of product lines operated by high-type firms increases from 10% to 18%. Finally, Panel D shows that entry subsidies have a very small effect.

In sum, the results of single uniform tax/subsidy policies Panels A-D in Table 12 suggest that taxes on the operations (or the fixed costs of operations) may be the most potent policies.

We next analyze the optimal combination of these uniform policies, with the results presented

---

[16]See Scotchmer (2004), Hopenhayn et al. (2006), and Akcigit et al. (2016b) on the design of policies to encourage innovation under asymmetric information.

[17]Recall that this is a tax on the fixed costs of operation, $w^s\phi$, not on all costs or revenues of firms. The 69% tax on the fixed costs of operation of incumbents is equivalent to an average tax of 8% on the revenues of incumbents.

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
					Panel A. Baseline					
0.51	25.90	38.13	55.04	6.28	147.26	130.33	19.86	17.16	2.26	100.00
					Panel B. Incumbent R&D ($s_i = 39\%$)					
0.38	30.74	46.54	47.67	8.65	160.07	142.83	26.40	19.06	2.51	101.22
					Panel C. Operation ($s_o = -69\%$)					
0.61	30.78	46.04	45.95	9.84	161.61	145.83	27.08	19.29	2.54	101.42
					Panel D. Entry ($s_e = 18\%$)					
0.62	25.74	37.69	54.26	6.95	147.58	131.35	20.00	17.20	2.27	100.04
					Panel E. Incumbent R&D and Operation ($s_i = -3\%, s_o = -74\%$)					
0.63	30.74	45.94	45.90	9.90	161.50	145.81	27.07	19.29	2.54	101.42

in Panel E of Table 12. Panel D already showed that entry subsidies are not very effective, and it turns out that conditional on using incumbent R&D subsidies and operation taxes, there is no further gain from using entry subsidies. So, Panel E focuses on the optimal combination of incumbent R&D subsidies and taxes on the fixed costs for continuing firms. The optimal combination of these uniform policies involves a large tax on fixed costs ($s_o = -74\%$) and perhaps surprisingly also a small tax on incumbent R&D ($s_i = -3\%$). The resulting allocation increases the growth rate of the economy to 2.54% and secures a 1.42% consumption-equivalent welfare gain. This gain is achieved by substantially increasing the exit threshold for low-type firms, which then increases the ratio of product lines operated by high-type firms to those operated by low-type firms from 11% in the baseline to 22%. With the skilled labor freed from operations, overall R&D investments also increase, though because these are uniform policies, R&D investments by both types of firms increase in tandem.

6 Robustness

The broad pattern of estimation results and policy analyses reported so far is quite robust. In this section, we illustrate this by considering a number of variations on our sample and model. In each case, the implied parameter estimates and the match between model and data moments are depicted in online Appendix B, while in the text we report the baseline equilibrium moments, the social planner's allocation and the allocation that results from the optimal choice of uniform incumbent R&D subsidies and taxes on operations.

6.1 Employment-weighted Sample

Our baseline estimation targets unweighted moments. Our first variation shows that targeting moments weighted by beginning of period employment (which means that we are using such

weighted moments both in the model and the data) makes little difference. The results are shown in Table 13, where we see similar values for most key equilibrium objects. The social planner's allocation reported in Panel B is also very similar to the baseline, though the increase in the growth rate is a little more modest—from 2.22% to 2.54%, with a corresponding 1.25% consumption-equivalent welfare gain. The implications of the optimal uniform policies are also similar (and these policies again involve a large tax on operations and in this case, no tax or subsidy on incumbent R&D), increasing the growth rate to 2.39%, with a consumption-equivalent welfare gain of 0.56%. The main reason for the smaller gains from both the social planner's allocation and the optimal uniform policies is that the ratio of product lines operated by high-type firms to low-type firms is not as low in this case as in our baseline estimation, thus limiting the extent of the selection effects that optimal policies leverage.

TABLE 13: EMPLOYMENT-WEIGHTED ESTIMATION

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
					Panel A. Baseline					
0.52	25.26	47.76	63.37	11.62	126.53	89.02	23.86	22.08	2.22	100.00
					Panel B. Social Planner					
0.58	24.43	52.98	38.24	28.80	152.33	30.79	31.93	25.18	2.54	101.25
				Panel C. Incumbent R&D and Operation ($s_i = 0\%, s_o = -47\%$)						
0.57	27.28	51.53	53.96	16.41	138.78	109.77	28.54	23.74	2.39	100.56

6.2 Organic Sample that Excludes M&A Activities

Our baseline sample includes "inorganic" entry and exit, taking the form of mergers and acquisitions (M&A) and spinouts (where part of an existing firm becomes a new legal entity). We next reestimate the model after removing all observations we determined to be potentially influenced by inorganic activity on these margins.[18] The results from this exercise are reported in Table 14. The broad patterns of various policy implications remain very similar to the baseline—for example, the social planner is now able to increase growth from 2.24% to 2.81%, with a 4.04% consumption-equivalent welfare gain, and the optimal policies once again involve a substantial tax on operations of continuing firms and no tax or subsidy on incumbent R&D.

---

[18]We identify these cases following the procedures of Haltiwanger et al. (2013). We use the establishment identifiers, which are distinct from firm identifiers, to identify cases where an establishment exists before or after the associated firm id. We flag as being a potentially inorganic birth the cases where more than 10% of the firm's initial employment appears to come from a pre-existing establishment owned by another firm in the prior year; similarly, a potential inorganic exit is flagged when more than 10% of the exiting firm's employment is in a plant that transfers to a new firm in the following year. This 10% bar is aggressive, but also serves well to test the issues. About 19% of births, 30% of exits, and 41% of firms overall show some measure of inorganic activity in our innovative firm sample. Excluding these firms leaves a sample size of 9,854 firm-period observations.

TABLE 14: EXCLUDING M&A ACTIVITIES

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
					Panel A. Baseline					
0.46	27.24	33.97	55.48	2.41	154.57	146.64	18.33	16.39	2.24	100.00
					Panel B. Social Planner					
0.62	27.94	43.56	10.09	39.16	224.99	88.54	30.52	20.50	2.81	104.04
				Panel C. Incumbent R&D and Operation ($s_i = 0\%, s_o = -76\%$)						
0.60	33.98	42.99	48.80	3.49	168.44	160.26	26.22	18.69	2.56	101.82

6.3 Manufacturing Sample

Because of our reliance on R&D moments, our baseline sample includes continuously-innovative firms as explained in Section 3.1, and thus excludes most manufacturing firms. We believe that the same dynamics should apply to many firms that do not report R&D but still engage in innovation-type activities to take over product lines currently operated by competitors.[19] To investigate this issue, we first reestimated our model dropping all R&D moments and calculating the remaining 15 moments using the universe of manufacturing firms (982,559 firm-period observations). We weight each firm such that the firm size distribution matches that of our core sample using 16 size bins. The results of this estimation are reported in Panel A of Table 15.[20]

TABLE 15: FULL MANUFACTURING (NON-INNOVATING FIRMS INCLUDED)

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
					Panel A. Baseline					
1.71	5.08	29.69	22.92	4.63	215.50	118.64	25.47	4.25	1.92	100.00
					Panel B. Social Planner					
1.95	5.29	35.03	16.93	6.63	256.76	53.90	36.29	5.17	2.34	101.88
				Panel C. Incumbent R&D and Operation ($s_i = 7\%, s_o = -41\%$)						
1.80	5.72	33.80	20.34	5.10	233.44	149.89	31.17	4.69	2.12	100.75

The overall patterns are similar to our baseline, though with lower innovation rates and aggregate growth, likely reflecting the inclusion of less innovative firms in the sample. Panel B shows that the social planner's allocation can again increase growth significantly (from 1.92% to 2.34%), and achieves this once again by leveraging the selection effects. The implications of

---

[19]See National Research Council (2004) and Corrado et al. (2005) on the range of innovation activities not recorded in R&D surveys.

[20]We also verified that dropping the R&D moments in our baseline sample leads to similar estimation results and policy conclusions. These results are reported in Appendix B-3.

optimal uniform policies in Panel C are also similar, though now there is a small subsidy to incumbent innovation too.

6.4 Model with Unskilled Overhead Labor

In this subsection, we return to our initial sample but modify our baseline model to allow for the fixed operations cost to consist of both skilled and unskilled labor. Namely, we assume that a β fraction of the overhead labor ϕ has to be skilled, and the remaining $1 - \beta$ fraction is from unskilled labor. This leads to a simple generalization of our setup, with the Bellman equation for a $k-$type firm now taking the form

$$r\tilde{V}_k\left(\hat{Q}\right) - \dot{\tilde{V}}_k\left(\hat{Q}\right) = \max_{x_k \geq 0} \left\{ \begin{array}{c} \Sigma_{\hat{q} \in \hat{Q}} \left[\begin{array}{c} \tilde{\pi}(\hat{q}) - \phi \left[\beta \tilde{w}_s + (1 - \beta) \tilde{w}_u \right] \\ + \tau \left[\tilde{V}_k\left(\hat{Q} \setminus \{\hat{q}\}\right) - \tilde{V}_k\left(\hat{Q}\right) \right] \end{array} \right] \\ -n\tilde{w}^s G(x, \theta^k) + n x_k \left[\mathbb{E}\tilde{V}_k\left(\hat{Q} \cup \{\hat{q} + \lambda \bar{\hat{q}}\}\right) - \tilde{V}_k\left(\hat{Q}\right) \right] \\ + \varphi \left[0 - \tilde{V}_k\left(\hat{Q}\right) \right] + \mathbb{I}_{k=h} \nu \left[\tilde{V}_l\left(\hat{Q}\right) - \tilde{V}_h\left(\hat{Q}\right) \right] \end{array} \right\}, \text{ for } k \in \{l, h\}.$$

The labor-market clearing conditions are then modified to accommodate the use of both skilled and unskilled labor in operations as follows

$$L^S = L^{RD} + \Phi \beta \phi \text{ and } 1 = L^P + \Phi(1 - \beta)\phi.$$

We also set the parameter β to match the fraction of managers who have a college degree or above, which is 45.7%. The results are reported in Table 16.

Table 16: Model with Unskilled Overhead Labor

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
					Panel A. Baseline					
0.56	21.48	36.46	54.08	10.26	134.58	104.20	28.70	15.91	2.23	100.00
					Panel B. Social Planner					
0.59	19.77	39.08	38.48	22.23	151.12	30.07	32.71	16.89	2.37	100.56
					Panel C. Incumbent R&D and Operation ($s_i = -4\%, s_o = -9\%$)					
0.59	21.76	36.92	52.64	10.93	136.93	107.82	29.54	16.08	2.26	100.02

The baseline estimation leads to very similar results. The implications of the social planner's allocation and optimal uniform policies are also similar, but generate smaller gains relative to the baseline—in large part because the ability of these policies to free up skilled labor from operations is now more limited. All the same, the qualitative patterns are similar, and both the social planner's direct intervention and the optimal uniform policies again leverage the selection effect.[21]

---

[21]Perhaps the most important difference is that the tax on the operations of continuing firms is now smaller, 9%, as

6.5 Model with Reallocation Cost

Our baseline model does not incorporate any costs for reallocating labor from the original firm operating a product line to a new one taking it over. In practice, there may be several types of reallocation costs, both because some workers might go through unemployment and also because some employees may need to be retrained to work for their new employers or with new technologies. Here, we investigate the implications of allowing for these types of reallocation costs by introducing them in a reduced-form manner. Namely, we assume that hiring new workers entails training costs, and training each type of worker requires v workers of the same type for training. As a result, when a new firm hires l new unskilled workers and ϕ skilled workers for operations, it incurs an additional cost of $v\left[\tilde{w}^u l + \tilde{w}^s \phi\right]$ (the reallocation of R&D inputs is assumed to be costless). This modification leads to a small modification in the Bellman equations, which now take the form

$$
r\tilde{V}_k\left(\hat{Q}\right) - \dot{\tilde{V}}_k\left(\hat{Q}\right) = \max_{x_k \geq 0} \left\{ \begin{array}{c} \sum_{\hat{q} \in \hat{Q}} \left[\tilde{\pi}(\hat{q}) - \tilde{w}^s \phi + \tau\left[\tilde{V}_k\left(\hat{Q} \backslash \{\hat{q}\}\right) - \tilde{V}_k\left(\hat{Q}\right)\right]\right] \\ -n\tilde{w}^s G(x, \theta^k) \\ +nx_k \left[\begin{array}{c} \mathbb{E}\tilde{V}_k\left(\hat{Q} \cup \{\hat{q} + \lambda\bar{\hat{q}}\}\right) - \tilde{V}_k\left(\hat{Q}\right) \\ -v\tilde{w}^u \mathbb{E}\left[l(\hat{q} + \lambda\bar{\hat{q}})\right] - v\tilde{w}^s \phi \end{array} \right] \\ +\varphi\left[0 - \tilde{V}_k\left(\hat{Q}\right)\right] + \mathbb{I}_{k=h}\nu\left[\tilde{V}_l\left(\hat{Q}\right) - \tilde{V}_h\left(\hat{Q}\right)\right] \end{array} \right\}, \quad \text{for } k \in \{l, h\},
$$

where we have imposed that the reallocation costs are paid when the firm expands by taking over a product line from another incumbent. Because in equilibrium reallocation costs are incurred at the rate of average creative destruction τ, the labor-market clearing conditions become

$$
L^S_{\text{supply}} = L^S_{\text{demand}} + v\tau\phi \quad \text{and} \quad L^P_{\text{supply}} = L^P_{\text{demand}} + \tau v L^P_{\text{demand}}.
$$

We identify the new cost parameter v using estimates from Bloom et al. (2014) on the costs of training as equivalent to one month of a worker's time, which translates into $v = 1/12$. The resulting baseline equilibrium values and policy experiments are reported in Table 17, which shows very similar results to the baseline.

TABLE 17: MODEL WITH REALLOCATION COST

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
				Panel A. Baseline						
0.41	23.49	43.60	60.26	4.89	137.43	106.61	19.57	16.70	2.25	100.00
				Panel B. Social Planner						
0.49	24.53	52.38	31.57	23.58	173.87	30.47	31.65	20.58	2.77	102.55
			Panel C. Incumbent R&D and Operation ($s_i = 7\%, s_o = -73\%$)							
0.46	26.55	52.72	48.15	10.81	154.34	127.68	27.05	18.94	2.55	101.44

opposed to the taxes that were around 70% in our other samples and variations.

For example, the social planner's allocation increases the growth rate from 2.25% to 2.77%, with a consumption-equivalent welfare gain of 2.55%. Optimal uniform policies again impose a substantial tax operations and achieve a 1.44% consumption-equivalent welfare gain.

6.6   Model with Three Types of Firms

We next verify that our results are not unduly sensitive to assuming two types of firms by extending the model to three types of firms. The estimation results reported in Appendix B show that the innovative capacities of high-type and middle-type firms are estimated to be similar. Unsurprisingly in view of this, we find Table 18 that the policy implications also remain similar. For example, the social planner's allocation increases the growth rate from 2.20% to 2.94%, with a consumption-equivalent welfare gain of 5.35%. Optimal uniform policies again substantially tax the fixed cost of operations for continuing firms and achieve 1.81% consumption-equivalent welfare gain.

TABLE 18: MODEL WITH THREE TYPES $\left(\theta^H, \theta^M, \theta^L\right)$

x^{entry}	x^l	x^m	x^h	Φ^l	Φ^m	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{m,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
						Panel A. Baseline							
0.51	25.83	38.19	40.85	52.32	5.07	1.12	153.14	136.78	132.04	19.11	16.41	2.20	100.00
						Panel B. Social Planner							
0.59	29.36	45.88	38.19	2.49	44.32	0.69	276.16	25.91	93.28	34.32	21.92	2.94	105.35
				Panel C. Incumbent R&D and Operation ($s_i = 6\%, s_o = -69\%$)									
0.59	31.35	46.90	50.83	42.30	6.50	3.78	169.22	154.47	149.30	27.29	18.82	2.52	101.81

6.7   Model with Endogenous Supply of Skills

Finally, we extend our model to endogenize the supply of skilled workers. Specifically, we adapt our framework to an overlapping generations setup where each individual faces a constant death rate of ζ, and a measure ζ of new agents arrive at each instant, so that total population remains constant. In addition, each agent has a type indexed by κ. Upon entry into the economy, agents have a decision to acquire skills. Each agent can supply one unit of unskilled labor without any investment, and can also supply one unit of skilled labor if they acquire education, which is assumed to last a^* years for everybody. Education requires some of the skilled workers to be allocated to teaching, and we assume that an agent with type κ requires the services of $1/\kappa$ teachers during his education. Thus, the costs of education are higher for agents with low κ, and because these agents will have to bear this cost of education, they are less likely to become skilled.

We take the distribution of κ to be truncated Pareto,

$$\kappa \sim A\kappa^{\chi-1},$$

for convenience, where $\chi < 1$, $\kappa \in [0, \bar{\kappa}]$ and $A = \chi\bar{\kappa}^{-\chi}$.

Education decisions are some of the most heavily subsidized activities in practice. In our model too the social planner will face a strong incentive to subsidize education because skilled workers create positive externalities when they perform R&D. If we rule out such subsidies, then other optimal policies would try to mimic them, potentially distorting the results of our policy analysis. For this reason, we also introduce an education subsidy at the rate $s_{edu} \in [0, 1]$ that reduces the cost of education faced by the agents. Incorporating this subsidy, we can see that an agent of type κ will acquire education if

$$\underbrace{\frac{e^{-(r-g+\zeta)a^*}w_S}{r-g+\zeta}}_{\text{present value of skilled worker}} - \underbrace{(1-s_{edu})\frac{1}{\kappa}w_S\int_0^{a^*}e^{-(r-g+\zeta)t}dt}_{\text{present value of education cost}} > \underbrace{\frac{w_U}{r-g+\zeta}}_{\text{present value of unskilled worker}}.$$

The right-hand side of this expression is the present discounted value of working as an unskilled worker, taking into account that the unskilled wage at the moment, w_U, will grow at the rate g, and that the agent has an effective discount rate of $r + \zeta$. The first term on the left-hand side is the present discounted value of working as a skilled labor, which recognizes that skilled workers will have no earnings during the first a^* years of their lives. Finally, the second term on the left-and side is the subsidized cost of education for a worker of type κ. This comparison gives a threshold for κ,

$$\kappa^* = (1-s_{edu})\left[1 - e^{-(r-g+\zeta)a^*}\right]\left(e^{-(r-g+\zeta)a^*} - \frac{w_U}{w_S}\right)^{-1},$$

such that only those with $\kappa > \kappa^*$ will become skilled.

We denote the total population by L, which comprises unskilled labor (L^P), skilled R&D labor (L^{RD}), skilled labor working in operations (L^F), skilled teachers (L^T), and students still in the education process (L^E). Given the exponential age structure (due to the constant death rate), the fraction of workers becoming skilled who are still below the age of a^* is $1 - e^{-\zeta a^*}$, which implies that in the stationary equilibrium, the masses of teachers and students are, respectively,

$$L^T = L\left(1 - e^{-\zeta a^*}\right)\int_{\kappa^*}^{\bar{\kappa}}\frac{1}{\kappa}dF(\kappa) \quad \text{and} \quad L^E = L\left(1 - e^{-\zeta a^*}\right)(1 - F(\kappa^*)).$$

Incorporating the employment of skilled workers as teachers, the labor market-clearing conditions become

$$L^{RD} + L^F = L\left[e^{-\zeta a^*}\left(1 - \frac{A}{\chi}(\kappa^*)^\chi\right) - \left(1 - e^{-\zeta a^*}\right)\frac{A}{\chi-1}\left(\bar{\kappa}^{\chi-1} - (\kappa^*)^{\chi-1}\right)\right] \quad \text{and} \quad L^P = L\frac{A}{\chi}(\kappa^*)^\chi.$$

37

To estimate this extended model with endogenous supply of skills, we choose the parameter ζ as 35 years to approximate the working life of skilled workers, and set $a^* = 4$ as the length of post-secondary education. We then choose $\chi = 0.035$, $\bar{\kappa} = 95.55$, and $L = 1.193$ so that this extended model replicates the supply of skilled and unskilled labor in our benchmark economy $(L^{RD} + L^F = 0.166$ and $L^U = 1)$ and 0.6% of total employment $(= L^T/(L^{RD} + L^F + L^P + L^T))$ being devoted to post-secondary teaching as in the US economy. By construction, the estimates for the remaining parameters are identical to our baseline estimates reported in Table 2 (because $L^{RD} + L^F = 0.166$ and $L^U = 1$ as before).

Table 19 reports the results of our policy analysis in this case. The baseline allocation without the education subsidy is identical to our benchmark results by construction and is reported in Panel A for comparison. Panel B shows that introducing an optimal education subsidy, at the rate $s_{edu} = 0.8$, increases the growth rate from 2.26 to 2.69, and secures a 10.9% improvement in welfare. This sizable welfare effect reflects the severe underprovision of skilled labor in the benchmark allocation. Panel C provides the social planner's optimal allocation, which exploits the same selection effect as in our baseline model and increases the growth rate further by another 0.59 percentage points to 3.28 and welfare by an additional 3.48% relative to the allocation with optimal education subsidy. The additional welfare and growth gains from the social planner's allocation over the one with just education subsidies are similar to the gains from the social planner's allocation in our benchmark economy.

TABLE 19: MODEL WITH ENDOGENOUS SUPPLY OF SKILLS

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	L^S	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
					Panel A. Baseline						
0.51	25.90	38.13	55.04	6.28	147.26	130.33	16.55	19.86	17.16	2.26	100.00
					Panel B. Baseline with Optimal Education Subsidy $(s_{edu} = 80\%)$						
0.55	27.46	40.98	60.02	8.26	133.97	114.56	18.94	22.03	20.41	2.69	110.94
					Panel C. Social Planner						
0.63	26.81	47.73	16.13	41.68	188.84	28.80	18.93	33.93	24.84	3.28	114.42
Panel D. Incumbent R&D, Operation Cost and Education Policies $(s_i = -4\%, s_o = -64\%, s_{edu} = 93\%)$											
0.66	31.31	47.19	50.56	12.39	147.02	129.82	18.94	28.10	22.33	2.95	112.08

Panel D shows that the same mix of uniform policies as before—incumbent R&D subsidy and tax on operation costs—but now combined with education subsidies lead to somewhat smaller gains than the social planner, but again achieve this by leveraging the selection effect. In particular, in addition to a higher education subsidy, $s_{edu} = 0.93$, we have a tax on operations, which has a very similar magnitude to our baseline results $(s_o = -0.64)$, and a small tax on incumbent R&D $(s_i = -0.04)$. These policies again increase the exit thresholds, and especially for low-type firms, and increase the growth rates to 2.95% and lead to 12.08% improvement in consumption-equivalent welfare. Thus overall, we conclude that our policy conclusions are robust to endogenizing the supply of skilled labor.

7 Conclusions

In this paper we build a micro-founded model of firm innovation and growth. The model enables us to examine the forces jointly driving innovation, productivity growth and reallocation. We estimate the parameters of the model using simulated method of moments on detailed US Census Bureau micro data on employment, output, R&D and patenting. Our model fits the key moments from microdata reasonably well, and also performs well on non-targeted moments and is in line with the range of micro estimates in the literature.

We use the model to investigate the implications of several types of industrial policies on long-run growth and welfare. We find that industrial policies (subsidies to incumbent R&D or to their operating costs) reduce growth and welfare, and entry subsidies are also ineffective. These small effects are not because the equilibrium of our model is approximately optimal. On the contrary, a social planner limited to affecting only R&D, entry and exit decisions can increase growth from 2.26% to 2.94%, and increase welfare by 4.46%. The social planner achieves this by strongly leveraging the selection effect. She forces low-type incumbents to exit at a very high rate, reduces their R&D, and increases the R&D of high-type incumbents.

Our general equilibrium model, which incorporates both reallocation and selection effects, also highlights the potential pitfalls of industrial policies supporting incumbents. Though there is substantial underinvestment in R&D, the optimal policy is not to subsidize R&D-type activities, because such subsidies increase R&D investments by both low-type and high-type firms. Instead, optimal policy should free up resources from the operations of low-type firms to be used for R&D by high-type firms, and this can be achieved by encouraging the exit of low-productivity firms, for example by taxing the operations of all firms.

Several further topics of inquiry are left for future research. First, it would be interesting to extend our analysis to incorporate an endogenous selection between non-innovation and innovation, and also to incorporate reallocation of other resources (unskilled labor and capital). Second, our analysis has been confined to comparisons of stationary equilibria (balanced growth paths), thus ignoring transition costs, which could be nontrivial. Third, and related, our baseline model did not incorporate any reallocation costs, though we allowed for such costs in a reduced-form manner in our extensions. A more systematic investigation of such costs would necessitate an micro-founded model of closely allocation of resources, for example via search (see Lentz and Mortensen, 2010) for a complementary approach on this question). Fourth, an interesting possible extension of our framework would be to model the joint dynamics of innovation, reallocation and unemployment, which can enrich the analysis of the effects of various policies, and also enable us to incorporate some of the potential unemployment benefits of supporting incumbent producers. Fifth, we have also abstracted from political constraints. It would be important to consider the political economy of different types of industrial policies, which have often been politically difficult to manage and prone to capture. Sixth, our model can also be used to study mergers between high- and low-type firms which might be able to make more efficient use of

the existing knowledge stock of low-type firms in certain circumstances. Finally, supplementing our approach with more direct estimation of the costs and benefits of different types of policies targeted at R&D by incumbents is a major area for future research.

References

Acemoglu, D.: 2008, *Introduction to Modern Economic Growth*. Princeton University Press.

Acemoglu, D., G. Gancia, and F. Zilibotti: 2012, 'Competing Engines of Growth: Innovation and Standardization'. *Journal of Economic Theory* **147**(2), 570–601.

Aghion, P., J. Cai, M. Dewatripont, L. Du, A. Harrison, and P. Legros: 2015, 'Industrial Policy and Competition'. *American Economic Journal: Macroeconomics* **7**(4), 1–32.

Aghion, P. and P. Howitt: 1992, 'A Model of Growth through Creative Destruction'. *Econometrica* **60**(2), 323–351.

Akcigit, U., D. Hanley, and N. Serrano-Velarde: 2016a, 'Back to Basics: Basic Research Spillovers, Innovation Policy and Growth'. CEPR Discussion Paper 11707.

Akcigit, U., D. Hanley, and S. Stantcheva: 2016b, 'Optimal Taxation and R&D Policies'. National Bureau of Economic Research Working Paper #22908.

Akcigit, U. and W. Kerr: 2017, 'Growth through Heterogeneous Innovations'. *Journal of Political Economy*. forthcoming.

Balasubramanian, N. and J. Sivadasan: 2011, 'What Happens When Firms Patent? New Evidence from US Economic Census Data'. *Review of Economics and Statistics* **93**(1), 126–146.

Bartelsman, E. J. and M. Doms: 2000, 'Understanding Productivity: Lessons from Longitudinal Microdata'. *Journal of Economic Literature* **38**(3), 569–594.

Bartelsman, E. J., J. C. Haltiwanger, and S. Scarpetta: 2013, 'Cross-Country Differences in Productivity: The Role of Allocative Efficiency'. *American Economic Review* **103**(1), 305–334.

Bernard, A. B., S. J. Redding, and P. K. Schott: 2010, 'Multiple-product Firms and Product Switching'. *American Economic Review* **100**(1), 70–97.

Bloom, N., E. Brynjolfsson, L. Foster, R. S. Jarmin, M. Patnaik, I. Saporta-Eksten, and J. Van Reenen: 2017, 'What Drives Differences in Management?'. National Bureau of Economic Research WP 23300.

Bloom, N., R. Griffith, and J. Van Reenen: 2002, 'Do R&D Tax Credits Work? Evidence from a Panel of Countries 1979–1997'. *Journal of Public Economics* **85**(1), 1–31.

Bloom, N., J. Liang, J. Roberts, and Z. J. Ying: 2014, 'Does Working from Home Work? Evidence from a Chinese Experiment'. *Quarterly Journal of Economics* **130**(1), 165–218.

Blundell, R., R. Griffith, and F. Windmeijer: 2002, 'Individual Effects and Dynamics in Count Data Models'. *Journal of Econometrics* **108**(1), 113–131.

Broda, C. and D. E. Weinstein: 2006, 'Globalization and the Gains from Variety'. *Quarterly Journal of Economics* **121**(2), 541–585.

Corrado, C., C. Hulten, and D. Sichel: 2005, 'Measuring Capital and Technology: An Expanded Framework'. In: *Measuring Capital in the New Economy*. University of Chicago Press, pp. 11–46.

Criscuolo, C., R. Martin, H. Overman, and J. Van Reenen: 2012, 'The Causal Effects of an Industrial Policy'. National Bureau of Economic Research WP 17842.

Davis, S. J., J. Haltiwanger, R. Jarmin, J. Miranda, C. Foote, and E. Nagypal: 2006, 'Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms [with Comments and Discussion]'. *NBER Macroeconomics Annual* **21**, 107–179.

Ericson, R. and A. Pakes: 1995, 'Markov-perfect Industry Dynamics: A Framework for Empirical Work'. *Review of Economic Studies* **62**(1), 53–82.

European Commission: 2011, 'State Aid Scorecard'. http://ec.europa.eu/competition/state_aid/studies_reports/archive/2011_autumn_working_paper_en.pdf. [Online; accessed 20-July-2017].

Foster, J., J. C. Haltiwanger, and C. Syverson: 2008, 'Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?'. *American Economic Review* **98**(1), 394–425.

Foster, L. and C. Grim: 2010, 'Characteristics of the Top R&D Performing Firms in the US: Evidence from the Survey of Industrial R&D'. CES Working Paper 10-33.

Foster, L., J. Haltiwanger, and C. J. Krizan: 2006, 'Market Selection, Reallocation, and Restructuring in the US Retail Trade Sector in the 1990s'. *Review of Economics and Statistics* **88**(4), 748–758.

Foster, L., J. C. Haltiwanger, and C. J. Krizan: 2001, 'Aggregate Productivity Growth: Lessons from Microeconomic Evidence'. In: *New Developments in Productivity Analysis*. University of Chicago Press, pp. 303–372.

Garcia-Macia, D., C.-T. Hsieh, and P. Klenow: 2016, 'How Destructive is Innovation?'. Stanford University mimeo.

Goolsbee, A.: 1998, 'Does Government R&D Policy Mainly Benefit Scientists and Engineers?'. *American Economic Review, Papers and Proceedings* **88**(2), 298–302.

Griliches, Z.: 1990, 'Patent Statistics as Economic Indicators: A Survey'. *Journal of Economic Literature* **28**(4), 1661–1707.

Grossman, G. and E. Helpman: 1991, 'Quality Ladders in the Theory of Growth'. *Review of Economic Studies* **58**(1), 43–61.

Guner, N., G. Ventura, and Y. Xu: 2008, 'Macroeconomic Implications of Size-dependent Policies'. *Review of Economic Dynamics* **11**(4), 721–744.

Hall, B. H.: 1993, 'R&D Tax Policy during the 1980s: Success or Failure?'. In: *Tax Policy and the Economy*, Vol. 7. National Bureau of Economic Research and The MIT Press, pp. 1–35.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg: 2001, 'The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools'. National Bureau of Economic Research WP 8498.

Hall, B. H. and R. H. Ziedonis: 2001, 'The Patent Paradox Revisited: An Empirical Study of Patenting in the US Semiconductor Industry, 1979-1995'. *RAND Journal of Economics* pp. 101–128.

Haltiwanger, J., R. S. Jarmin, and J. Miranda: 2013, 'Who Creates Jobs? Small versus Large versus Young'. *Review of Economics and Statistics* **95**(2), 347–361.

Hopenhayn, H.: 1992, 'Entry, Exit and Firm Dynamics in Long-Run Equilibrium'. *Econometrica* **60**(5), 1127–1150.

Hopenhayn, H.: 2012, 'On the Measure of Distortions'. *Working Paper*.

Hopenhayn, H., G. Llobet, and M. Mitchell: 2006, 'Rewarding Sequential Innovators: Prizes, Patents, and Buyouts'. *Journal of Political Economy* **114**(6), 1041–1068.

Hopenhayn, H. and R. Rogerson: 1993, 'Job Turnover and Policy Evaluation: A General Equilibrium Analysis'. *Journal of Political Economy* **101**(5), 915–938.

Hsieh, C.-T., E. Hurst, P. Klenow, and C. Jones: 2013, 'The Allocation of Talent and U.S. Economic Growth'. National Bureau of Economic Research Working Paper #18693.

Hsieh, C.-T. and P. Klenow: 2009, 'Misallocation and Manufacturing TFP in China and India'. *Quarterly Journal of Economics* **124**(2), 771–807.

Hsieh, C.-T. and P. J. Klenow: 2014, 'The Life Cycle of Plants in India and Mexico'. *Quarterly Journal of Economics* **129**(3), 1035–1084.

Jones, C. I.: 2011, 'Misallocation, Economic Growth, and Input-output Economics'. National bureau of economic research WP 16742.

Jovanovic, B.: 1982, 'Selection and the Evolution of Industry'. *Econometrica* **50**(3), 649–670.

Kerr, W. R. and S. Fu: 2008, 'The Survey of Industrial R&D-patent Database Link Project'. *Journal of Technology Transfer* **33**(2), 173–186.

Klette, T. J. and S. Kortum: 2004, 'Innovating Firms and Aggregate Innovation'. *Journal of Political Economy* **112**(5), 986–1018.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman: 2017, 'Technological Innovation, Resource Allocation, and Growth'. *Quarterly Journal of Economics* **132**(2), 665–712.

Lentz, R. and D. T. Mortensen: 2008, 'An Empirical Model of Growth through Product Innovation'. *Econometrica* **76**(6), 1317–1373.

Lentz, R. and D. T. Mortensen: 2010, 'Labor Market Friction, Firm Heterogeneity, and Aggregate Employment and Productivity'. *University of Wisconsin Discussion Paper*.

Lerner, J.: 2009, *Boulevard of Broken Dreams: Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed – And What To Do About it*. Princeton University Press.

McFadden, D.: 1989, 'A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration'. *Econometrica* **57**(5), 995–1026.

National Research Council: 2004, *Measuring Research and Development Expenditures in the US Economy*. National Academies Press.

Owen, G.: 1999, *From Empire to Europe: The Decline and Revival of British Industry since the Second World War*. Harper Collins.

Pakes, A. and D. Pollard: 1989, 'Simulation and the Asymptotics of Optimization Estimators'. *Econometrica* pp. 1027–1057.

Peters, M.: 2016, 'Heterogeneous Mark-Ups, Growth and Endogenous Misallocation'. Working Paper.

Restuccia, D. and R. Rogerson: 2008, 'Policy Distortions and Aggregate Productivity with Heterogenous Establishments'. *Review of Economic Dynamics* **11**(4), 707–720.

Romer, P.: 2001, 'Should the Government Subsidize Supply or Demand in the Market for Scientists and Engineers?'. In: *Innovation Policy and the Economy*, Vol. 1. MIT Press, pp. 221–252.

Romer, P. M.: 1990, 'Endogenous Technological Change'. *Journal of Political Economy* **98**(5), 71–102.

Scotchmer, S.: 2004, *Innovation and Incentives*. MIT Press.

Syverson, C.: 2011, 'What Determines Productivity?'. *Journal of Economic Literature* **49**(2), 326–365.

Wilson, D.: 2009, 'Beggar thy Neighbor? The In-State, Out-of-State and Aggregate Effects of R&D Tax Credits'. *Review of Economics and Statistics* pp. 431–436.

Appendix A

Proof of Lemma 1. Consider the low-type firms and conjecture $\tilde{V}_l\left(\hat{\mathcal{Q}}\right) = \sum_{\hat{q} \in \hat{\mathcal{Q}}} Y^l(\hat{q})$:

$$r \sum_{\hat{q} \in \hat{\mathcal{Q}}} Y^l(\hat{q}) = \sum_{\hat{q} \in \hat{\mathcal{Q}}} \max \left\{ 0, \max_{x \geq 0} \left[\begin{array}{c} \tilde{\pi}\left(\hat{q}_j\right) - \tilde{w}^s \phi^s - \tilde{w}^s G\left(x, \theta^L\right) + \frac{\partial Y^l(\hat{q})}{\partial \hat{q}} \frac{\partial \hat{q}}{\partial w^u} \frac{\partial w^u}{\partial t} \\ + x \mathbb{E} Y^l\left(\hat{q} + \lambda \bar{\hat{q}}\right) - (\tau + \varphi) Y^l(\hat{q}) \end{array} \right] \right\},$$

which implies

$$r Y^l(\hat{q}) = \max \left\{ 0, \left\{ \begin{array}{c} \tilde{\pi}(\hat{q}) - \tilde{w}^s \phi^s + \frac{\partial Y^l(\hat{q})}{\partial \hat{q}} \frac{\partial \hat{q}}{\partial w^u} \frac{\partial w^u}{\partial t} - (\tau + \varphi) Y^l(\hat{q}) \\ + \max_{x \geq 0} \left[x \mathbb{E} Y^l\left(\hat{q} + \lambda \bar{\hat{q}}\right) - \tilde{w}^s G\left(x, \theta^L\right) \right] \end{array} \right\} \right\},$$

where we also use the fact that a firm can choose not to operate an individual product line.

Next consider the high-type firms and conjecture $\tilde{V}_h\left(\hat{\mathcal{Q}}\right) = \sum_{\hat{q} \in \hat{\mathcal{Q}}} Y^h(\hat{q})$:

$$r \sum_{\hat{q} \in \hat{\mathcal{Q}}} Y^h(\hat{q}) = \sum_{\hat{q} \in \hat{\mathcal{Q}}} \max \left\{ 0, \max_{x \geq 0} \left[\begin{array}{c} \tilde{\pi}(\hat{q}) - \tilde{w}^s \phi^s - \tilde{w}^s G\left(x, \theta^H\right) + \frac{\partial Y^h(\hat{q})}{\partial \hat{q}} \frac{\partial \hat{q}}{\partial w^u} \frac{\partial w^u}{\partial t} \\ + x \mathbb{E} Y^h\left(\hat{q} + \lambda \bar{\hat{q}}\right) \\ - (\tau + \varphi) Y^h(\hat{q}) + \nu \left[\mathbb{I}_{\hat{q} > \hat{q}_{l,\min}} \cdot Y^l(\hat{q}) - Y^h(\hat{q}) \right] \end{array} \right] \right\},$$

which similarly implies

$$r Y^h(\hat{q}) = \max \left\{ 0, \max_{x \geq 0} \left[\begin{array}{c} \tilde{\pi}(\hat{q}) - \tilde{w}^s \phi^s - \tilde{w}^s G\left(x, \theta^H\right) + \frac{\partial Y^h(\hat{q})}{\partial \hat{q}} \frac{\partial \hat{q}}{\partial w^u} \frac{\partial w^u}{\partial t} + \\ x \mathbb{E} Y^h\left(\hat{q} + \lambda \bar{\hat{q}}\right) \\ - (\tau + \varphi) Y^h(\hat{q}) + \nu \left[\mathbb{I}_{\hat{q} > \hat{q}_{l,\min}} \cdot Y^l(\hat{q}) - Y^h(\hat{q}) \right] \end{array} \right] \right\}.$$

Monotonicity follows from the fact that the per-period return function is increasing in $\hat{q}$. ∎

Proof of Proposition 1. First note that $\tilde{\pi}(q) = \left(\frac{\varepsilon-1}{\varepsilon}\right)^{\varepsilon} \frac{1}{\varepsilon-1} \hat{q}^{\varepsilon-1} = \Pi \hat{q}^{\varepsilon-1}$. Then, defining $\Psi \equiv r + \tau + \varphi$, equation (17) can be written as the following linear differential equation

$$\Psi Y^l(\hat{q}) + g \hat{q} \frac{\partial Y^l(\hat{q})}{\partial \hat{q}} = \Pi \hat{q}^{\varepsilon-1} + \Omega^l - \tilde{w}^s \phi \text{ if } \hat{q} > \hat{q}_{l,\min}$$

or

$$\xi_1 \hat{q}^{-1} Y^l(\hat{q}) + \frac{\partial Y^l(\hat{q})}{\partial \hat{q}} = \xi_2 \hat{q}^{\varepsilon-2} - \xi_3 \hat{q}^{-1}, \tag{A-1}$$

where $\xi_1 \equiv \frac{\Psi}{g}$, $\xi_2 \equiv \frac{\Pi}{g}$ and $\xi_3 \equiv \frac{\tilde{w}^s \phi - \Omega^l}{g}$. Then the solution to (A-1) can be written as

$$Y^l(\hat{q}) = \hat{q}^{-\xi_1} \left(\int \left[\xi_2 t^{\xi_1 + \varepsilon - 2} - \xi_3 t^{\xi_1 - 1} \right] dt + D \right) = \frac{\xi_2 \hat{q}^{\varepsilon-1}}{\xi_1 + \varepsilon - 1} - \frac{\xi_3}{\xi_1} + D \hat{q}^{-\xi_1}.$$

44

Imposing the boundary condition $Y^l(\hat{q}_{l,\min}) = 0$, we can solve out for the constant of integration D, obtaining

$$
\begin{aligned}
Y^l(\hat{q}) &= \frac{\xi_2 \hat{q}^{\varepsilon-1}}{\xi_1 + \varepsilon - 1} - \frac{\xi_3}{\xi_1} + \left(\frac{\xi_3 \hat{q}_{l,\min}^{\xi_1}}{\xi_1} - \frac{\xi_2 \hat{q}_{l,\min}^{\xi_1+\varepsilon-1}}{\xi_1 + \varepsilon - 1} \right) \hat{q}^{-\xi_1} \\
&= \frac{\Pi \hat{q}^{\varepsilon-1}}{\Psi + (\varepsilon - 1)g} \left(1 - \left(\frac{\hat{q}_{l,\min}}{\hat{q}} \right)^{\frac{\Psi}{g}+\varepsilon-1} \right) + \frac{\Omega^l - \tilde{w}^s \phi}{\Psi} \left(1 - \left(\frac{\hat{q}_{l,\min}}{\hat{q}} \right)^{\frac{\Psi}{g}} \right).
\end{aligned}
\tag{A-2}
$$

We next provide the derivation of the value for a high-type product line. Let us rewrite the expression in (A-2) as

$$
Y^l(\hat{q}) = \xi_4 \hat{q}^{\varepsilon-1} + \xi_5 \hat{q}^{-\frac{\Psi}{g}} - \xi_6,
$$

where

$$
\xi_4 \equiv \frac{\Pi}{\Psi + (\varepsilon - 1)g}, \quad \xi_5 \equiv \frac{(\tilde{w}^s \phi - \Omega^l) \hat{q}_{l,\min}^{\frac{\Psi}{g}}}{\Psi} - \frac{\Pi \hat{q}_{l,\min}^{\frac{\Psi}{g}+\varepsilon-1}}{\Psi + g(\varepsilon - 1)}, \quad \text{and } \xi_6 \equiv \frac{\tilde{w}^s \phi - \Omega^l}{\Psi}.
$$

Recall the value of a product line of a high-type firm

$$
(\Psi + \nu) Y^h(\hat{q}) + \frac{\partial Y^h(\hat{q})}{\partial \hat{q}} g\hat{q} = \Pi \hat{q}^{\varepsilon-1} + \Omega^h - \tilde{w}^s \phi + \nu \left(\xi_4 \hat{q}^{\varepsilon-1} + \xi_5 \hat{q}^{-\frac{\Psi}{g}} - \xi_6 \right) \text{ for } \hat{q} \geq \hat{q}_{l,\min}
$$

$$
(\Psi + \nu) Y^h(\hat{q}) + \frac{\partial Y^h(\hat{q})}{\partial \hat{q}} g\hat{q} = \Pi \hat{q}^{\varepsilon-1} + \Omega^h - \tilde{w}^s \phi \text{ for } \hat{q}_{l,\min} > \hat{q} \geq \hat{q}_{h,\min},
$$

which can be rewritten as

$$
K_1 Y^h(\hat{q}) \hat{q}^{-1} + \frac{\partial Y^h(\hat{q})}{\partial \hat{q}} = K_2 \hat{q}^{\varepsilon-2} + K_3 \hat{q}^{-\frac{\Psi+g}{g}} - K_4 \hat{q}^{-1},
$$

where

$$
K_1 \equiv \frac{\Psi + \nu}{g}, \quad K_2 \equiv \frac{\Pi + \nu \xi_4}{g}, \quad K_3 \equiv \frac{\nu \xi_5}{g} \text{ and } K_4 \equiv \frac{\nu \xi_6 + \tilde{w}^s \phi - \Omega^h}{g} \text{ for } \hat{q} \geq \hat{q}_{l,\min} \tag{A-3}
$$

$$
K_1 \equiv \frac{\Psi + \nu}{g}, \quad K_2 \equiv \frac{\Pi}{g}, \quad K_3 \equiv 0 \text{ and } K_4 \equiv \frac{\tilde{w}^s \phi - \Omega^h}{g} \text{ for } \hat{q}_{l,\min} > \hat{q} \geq \hat{q}_{h,\min}. \tag{A-4}
$$

Then we can express the general solution for the high-type value function as

$$
\begin{aligned}
Y^h(\hat{q}) &= \hat{q}^{-K_1} \left(\int \left[K_2 \hat{q}^{K_1+\varepsilon-2} + K_3 \hat{q}^{K_1 - \frac{\Psi+g}{g}} - K_4 \hat{q}^{K_1-1} \right] d\hat{q} + D \right) \\
&= \frac{K_2 \hat{q}^{\varepsilon-1}}{K_1 + \varepsilon - 1} + \frac{K_3 \hat{q}^{1-\frac{\Psi+g}{g}}}{K_1 + 1 - \frac{\Psi+g}{g}} - \frac{K_4}{K_1} + D\hat{q}^{-K_1}.
\end{aligned}
\tag{A-5}
$$

To find the constant of integration D, we use $Y^h(\hat{q}_{h,\min}) = 0$, which yields

$$D = -\frac{K_2 \hat{q}_{h,\min}^{K_1+\varepsilon-1}}{K_1+\varepsilon-1} - \frac{K_3 \hat{q}_{h,\min}^{K_1+1-\frac{\Psi+g}{g}}}{K_1+1-\frac{\Psi+g}{g}} + \frac{K_4 \hat{q}_{h,\min}^{K_1}}{K_1} \text{ for } \hat{q} \in [\hat{q}_{h,\min}, \hat{q}_{l,\min}].$$

Then we can express the value function as

$$Y^h(\hat{q}) = \left\{ \begin{array}{c} \frac{K_2 \hat{q}^{\varepsilon-1}}{K_1+\varepsilon-1} + \frac{K_3 \hat{q}^{1-\frac{\Psi+g}{g}}}{K_1+1-\frac{\Psi+g}{g}} - \frac{K_4}{K_1} \\ + \left[-\frac{K_2 \hat{q}_{h,\min}^{\varepsilon-1}}{K_1+\varepsilon-1} - \frac{K_3 \hat{q}_{h,\min}^{1-\frac{\Psi+g}{g}}}{K_1+1-\frac{\Psi+g}{g}} + \frac{K_4}{K_1} \right] \left(\frac{\hat{q}_{h,\min}}{\hat{q}} \right)^{K_1} \end{array} \right\} = \begin{array}{c} \frac{K_2 \hat{q}^{\varepsilon-1}}{K_1+\varepsilon-1} \left[1 - \left(\frac{\hat{q}_{h,\min}}{\hat{q}} \right)^{K_1+\varepsilon-1} \right] \\ + \frac{K_3 \hat{q}^{-\frac{\Psi}{g}}}{K_1-\frac{\Psi}{g}} \left[1 - \left(\frac{\hat{q}_{h,\min}}{\hat{q}} \right)^{K_1-\frac{\Psi}{g}} \right] \\ - \frac{K_4}{K_1} \left[1 - \left(\frac{\hat{q}_{h,\min}}{\hat{q}} \right)^{K_1} \right] \end{array}.$$

Then from (A-4), we have that for $\hat{q} \in [\hat{q}_{h,\min}, \hat{q}_{l,\min}]$,

$$Y^h(\hat{q}) = \frac{\Pi \hat{q}^{\varepsilon-1}}{\Psi+\nu+(\varepsilon-1)g} \left(1 - \left(\frac{\hat{q}_{h,\min}}{\hat{q}} \right)^{\frac{\Psi+\nu+(\varepsilon-1)g}{g}} \right) + \frac{\Omega^h - \tilde{w}^s \phi}{\Psi+\nu} \left(1 - \left(\frac{\hat{q}_{h,\min}}{\hat{q}} \right)^{\frac{\Psi+\nu}{g}} \right).$$

Intuitively, because product lines with relative quality $\hat{q} \in [\hat{q}_{h,\min}, \hat{q}_{l,\min}]$ immediately become obsolete when operated by low-type firms, but not by high-type firms, the flow rate of transitioning from high-type to low-type, ν, becomes part of the effective discount rate in this range.

For $\hat{q} \geq \hat{q}_{l,\min}$, the appropriate values for K's from (A-3) delivers (A-5) as

$$Y^h(\hat{q}) = \frac{\Pi \hat{q}^{\varepsilon-1}}{\Psi+(\varepsilon-1)g} \left(1 - \left(\frac{\hat{q}_{l,\min}}{\hat{q}} \right)^{\frac{\Psi+(\varepsilon-1)g}{g}} \right) + \frac{\Omega^l - \tilde{w}^s \phi}{\Psi} \left(1 - \left(\frac{\hat{q}_{l,\min}}{\hat{q}} \right)^{\frac{\Psi}{g}} \right) + \frac{\Omega^h - \Omega^l}{\Psi+\nu} + D \hat{q}^{-\frac{\Psi+\nu}{g}}$$

We also have the boundary condition

$$Y^h(\hat{q}_{l,\min}) = \frac{\Pi \hat{q}_{l,\min}^{\varepsilon-1}}{\Psi+\nu+(\varepsilon-1)g} \left(1 - \left(\frac{\hat{q}_{h,\min}}{\hat{q}_{l,\min}} \right)^{\frac{\Psi+\nu+(\varepsilon-1)g}{g}} \right) + \frac{\Omega^h - \tilde{w}^s \phi}{\Psi+\nu} \left(1 - \left(\frac{\hat{q}_{h,\min}}{\hat{q}_{l,\min}} \right)^{\frac{\Psi+\nu}{g}} \right).$$
(A-6)

Hence, the constant of integration for $\hat{q} \geq \hat{q}_{l,\min}$ must satisfy (A-6). Next using (A-3) and (A-5), $Y^h(\hat{q}_{l,\min})$ for $\hat{q} \geq \hat{q}_{l,\min}$ can be computed as

$$
\begin{aligned}
Y^h(\hat{q}_{l,\min}) &= \frac{K_2 \hat{q}_{l,\min}^{\varepsilon-1}}{K_1+\varepsilon-1} + \frac{K_3 \hat{q}_{l,\min}^{1-\frac{\Psi+g}{g}}}{K_1+1-\frac{\Psi+g}{g}} - \frac{K_4}{K_1} + D \hat{q}_{l,\min}^{-K_1} \\
&= \frac{(\Pi+\nu\xi_4) \hat{q}_{l,\min}^{\varepsilon-1}}{\Psi+\nu+g(\varepsilon-1)} + \xi_5 \hat{q}_{l,\min}^{-\frac{\Psi}{g}} - \frac{\nu\xi_6 + \tilde{w}^s \phi - \Omega^h}{\Psi+\nu} + D \hat{q}_{l,\min}^{-\frac{\Psi+\nu}{g}},
\end{aligned}
$$
(A-7)

which must be equal to $(A\text{-}6)$. Equating $(A\text{-}6)$ to $(A\text{-}7)$, we get

$$D = \left\{ \begin{array}{c} -\frac{\Pi}{\Psi+\nu+(\varepsilon-1)g}\hat{q}_{h,\min}^{\frac{\Psi+\nu+(\varepsilon-1)g}{g}} + \frac{\tilde{w}^s\phi-\Omega^h}{\Psi+\nu}\hat{q}_{h,\min}^{\frac{\Psi+\nu}{g}} \\ -\frac{\nu\xi_4}{\Psi+\nu+g(\varepsilon-1)}\hat{q}_{l,\min}^{\frac{\Psi+\nu+(\varepsilon-1)g}{g}} - \xi_5\hat{q}_{l,\min}^{\frac{\nu}{g}} + \frac{\nu\xi_6}{\Psi+\nu}\hat{q}_{l,\min}^{\frac{\Psi+\nu}{g}} \end{array} \right\}.$$

Hence

$$\hat{q}^{-\frac{\Psi+\nu}{g}}D = \left\{ \begin{array}{c} \frac{\Pi\hat{q}^{\varepsilon-1}}{\Psi+\nu+g(\varepsilon-1)}\left(1-\left(\frac{\hat{q}_{h,\min}}{\hat{q}}\right)^{\frac{\Psi+\nu+(\varepsilon-1)g}{g}}\right) - \frac{\Pi\hat{q}^{\varepsilon-1}}{\Psi+\nu+g(\varepsilon-1)}\left(1-\left(\frac{\hat{q}_{l,\min}}{\hat{q}}\right)^{\frac{\Psi+\nu+(\varepsilon-1)g}{g}}\right) \\ +\frac{\Omega^l-\Omega^h}{\Psi+\nu} \\ -\frac{\tilde{w}^s\phi-\Omega^h}{\Psi+\nu}\left(1-\left(\frac{\hat{q}_{h,\min}}{\hat{q}}\right)^{\frac{\Psi+\nu}{g}}\right) + \frac{\tilde{w}^s\phi-\Omega^l}{\Psi+\nu}\left(1-\left(\frac{\hat{q}_{l,\min}}{\hat{q}}\right)^{\frac{\Psi+\nu}{g}}\right) \end{array} \right\}.$$

Therefore, for $\hat{q} \geq \hat{q}_{l,\min}$ we have

$$Y^h(\hat{q}) = \left\{ \begin{array}{c} \frac{\Pi\hat{q}^{\varepsilon-1}}{\Psi+\nu+g(\varepsilon-1)}\left(1-\left(\frac{\hat{q}_{h,\min}}{\hat{q}}\right)^{\frac{\Psi+\nu+(\varepsilon-1)g}{g}}\right) + \frac{\Omega^h-\tilde{w}^s\phi}{\Psi+\nu}\left(1-\left(\frac{\hat{q}_{h,\min}}{\hat{q}}\right)^{\frac{\Psi+\nu}{g}}\right) \\ \frac{\Pi\hat{q}^{\varepsilon-1}}{\Psi+(\varepsilon-1)g}\left(1-\left(\frac{\hat{q}_{l,\min}}{\hat{q}}\right)^{\frac{\Psi+(\varepsilon-1)g}{g}}\right) + \frac{\Omega^l-\tilde{w}^s\phi}{\Psi}\left(1-\left(\frac{\hat{q}_{l,\min}}{\hat{q}}\right)^{\frac{\Psi}{g}}\right) \\ -\left(\frac{\Pi\hat{q}^{\varepsilon-1}}{\Psi+\nu+g(\varepsilon-1)}\left(1-\left(\frac{\hat{q}_{l,\min}}{\hat{q}}\right)^{\frac{\Psi+\nu+(\varepsilon-1)g}{g}}\right) + \frac{\Omega^l-\tilde{w}^s\phi}{\Psi+\nu}\left(1-\left(\frac{\hat{q}_{l,\min}}{\hat{q}}\right)^{\frac{\Psi+\nu}{g}}\right)\right) \end{array} \right\}.$$

Finally, we need to determine the values for the exit thresholds $\hat{q}_{l,\min}$ and $\hat{q}_{h,\min}$. Using the above differential equations we get

$$\left.\frac{\partial Y^l(\hat{q})}{\partial t}\right|_{\hat{q}=\hat{q}_{l,\min}} = \frac{1}{g}\left(\Pi\hat{q}_{l,\min}^{\varepsilon-2} + \frac{\Omega^l-\tilde{w}^s\phi}{\hat{q}_{l,\min}}\right).$$

From the smooth-pasting condition we get

$$\left.\frac{\partial Y^l(\hat{q})}{\partial\hat{q}}\right|_{\hat{q}=\hat{q}_{l,\min}} = 0 \implies \hat{q}_{l,\min} = \left(\frac{\tilde{w}^s\phi-\Omega^l}{\Pi}\right)^{\frac{1}{\varepsilon-1}}.$$

Similarly, we also have

$$\left.\frac{\partial Y^h(\hat{q})}{\partial\hat{q}}\right|_{\hat{q}=\hat{q}_{h,\min}} = \frac{\Pi}{\Psi+\nu+(\varepsilon-1)g}\left((\varepsilon-1)\hat{q}^{\varepsilon-2} + \frac{\Psi+\nu}{g}\hat{q}_{h,\min}^{\frac{\Psi+\nu+(\varepsilon-1)g}{g}}\hat{q}^{-\frac{\Psi+\nu}{g}-1}\right) - \frac{\tilde{w}^s\phi-\Omega^h}{g}\hat{q}_{h,\min}^{\frac{\Psi+\nu}{g}}\hat{q}^{-\frac{\Psi+\nu}{g}-1}$$

and $\left.\frac{\partial Y^h(\hat{q})}{\partial\hat{q}}\right|_{\hat{q}=\hat{q}_{h,\min}} = 0$ implies

$$\hat{q}_{h,\min} = \left(\frac{\tilde{w}^s\phi-\Omega^h}{\Pi}\right)^{\frac{1}{\varepsilon-1}}.$$

∎

Lemma 3 *Let F denote the overall relative productivity distribution, including both active and inactive product lines. In stationary equilibrium, it satisfies the following differential equation:*

$$g\hat{q}f(\hat{q}) = \tau \left[F(\hat{q}) - F(\hat{q} - \lambda\bar{\hat{q}}) \right],$$

where $\tau = \Phi^h x^h + \Phi^l x^l + x^{entry}$ and $\bar{\hat{q}} = \int_0^\infty \hat{q}f(\hat{q}) \, d\hat{q}$. Moreover let $\tilde{F}_k$ denote the (unnormalized) distribution of relative productivities of active product lines, owned by type $k \in \{h,l\}$. In stationary equilibrium, they satisfy

$$g\hat{q}\tilde{f}_h(\hat{q}) = g\hat{q}_{h,\min}\tilde{f}_h(\hat{q}_{h,\min}) + \left(\tau^l + \varphi + \nu\right)\tilde{F}_h(\hat{q}) - \tau^h \left[F(\hat{q} - \lambda\bar{\hat{q}}) - F(\hat{q}_{h,\min} - \lambda\bar{\hat{q}}) - \tilde{F}_h(\hat{q}) \right]$$

$$g\hat{q}\tilde{f}_l(\hat{q}) = g\hat{q}_{l,\min}\tilde{f}_l(\hat{q}_{l,\min}) + \left(\tau^h + \varphi\right)\tilde{F}_l(\hat{q}) - \tau^l \left[F(\hat{q} - \lambda\bar{\hat{q}}) - F(\hat{q}_{l,min} - \lambda\bar{\hat{q}}) - \tilde{F}_l(\hat{q}) \right] - \nu \left[\tilde{F}_h(\hat{q}) - \tilde{F}_h(\hat{q}_{l,min}) \right],$$

where $\tau^l = \Phi^l x^l + (1 - \alpha) x^{entry}$ and $\tau^h = \Phi^h x^h + \alpha x^{entry}$. The measure of active product lines are given by

$$\Phi^k = \tilde{F}_k(\infty), \ k \in \{h,l\}.$$

Proof of Lemma 3. In a stationary equilibrium inflows and outflows into different parts of the distributions have to be equal. First consider overall productivity distribution F. Given a time interval of Δt, this implies that $F_t(\hat{q}) = F_{t+\Delta t}(\hat{q})$,

$$F_t(\hat{q}) = F_t(\hat{q}(1 + g\Delta t)) - \tau\Delta t \left[F_t(\hat{q}) - F_t(\hat{q} - \lambda\bar{\hat{q}}) \right]$$

Next, subtract $F_t(\hat{q}(1 + g\Delta t))$ from both sides, multiply both sides by -1, divide again sides by Δt, and take the limit as $\Delta t \to 0$, so that

$$\lim_{\Delta t \to 0} \frac{F(\hat{q}(1 + g\Delta t)) - F(\hat{q})}{\Delta t} = g\hat{q}f(\hat{q}).$$

Using this last expression delivers

$$g\hat{q}f(\hat{q}) = \tau \left[F(\hat{q}) - F(\hat{q} - \lambda\bar{\hat{q}}) \right].$$

Similarly, for active product line distributions $\tilde{F}_k$, we can write

$$\tilde{F}_{h,t}(\hat{q}) = \tilde{F}_{h,t}(\hat{q}(1 + g\Delta t)) - \tilde{F}_{h,t}(\hat{q}_{h,\min}(1 + g\Delta t)) + \tau^h\Delta t \left[F_t(\hat{q} - \lambda\bar{\hat{q}}) - \tilde{F}_{h,t}(\hat{q}) - F_t(\hat{q}_{h,min} - \lambda\bar{\hat{q}}) \right]$$
$$- \left(\tau^l + \varphi + \nu\right)\Delta t\tilde{F}_{h,t}(\hat{q})$$

$$\tilde{F}_{l,t}(\hat{q}) = \tilde{F}_{l,t}(\hat{q}(1 + g\Delta t)) - \tilde{F}_{l,t}(\hat{q}_{l,\min}(1 + g\Delta t)) + \tau^l\Delta t \left[F_t(\hat{q} - \lambda\bar{\hat{q}}) - \tilde{F}_{l,t}(\hat{q}) - F_t(\hat{q}_{l,min} - \lambda\bar{\hat{q}}) \right]$$
$$- \left(\tau^h + \varphi\right)\Delta t\tilde{F}_{l,t}(\hat{q}) + \nu\Delta t \left[\tilde{F}_{h,t}(\hat{q}) - \tilde{F}_{h,t}(\hat{q}_{l,min}) \right].$$

Again, by subtracting $\tilde{F}_{k,t}\left(\hat{q}\left(1+g\Delta t\right)\right)-\tilde{F}_{k,t}(\hat{q}_{k,\min}(1+g\Delta t))$ from both sides, dividing by $-\Delta t$, and taking the limit as $\Delta t \to 0$, we get the desired equations for $k \in \{h,l\}$ in Lemma 3. ∎

Proof of Proposition 2. As shown in Lemma 3, overall productivity distribution satisfies

$$\hat{q}f(\hat{q}) = \frac{\tau}{g}\left[F(\hat{q}) - F\left(\hat{q} - \lambda\bar{\hat{q}}\right)\right]$$

By integrating both sides over the domain, we get

$$\mathbb{E}(\hat{q}) \equiv \int_0^\infty \hat{q}f(\hat{q})d\hat{q} = \frac{\tau}{g}\int_0^\infty \left[F(\hat{q}) - F\left(\hat{q} - \lambda\bar{\hat{q}}\right)\right]d\hat{q}$$

We can write above equation as follows

$$\mathbb{E}(\hat{q}) = \frac{\frac{\tau}{g}}{1 + \frac{\tau}{g}}\int_0^\infty \left[1 - F\left(\hat{q} - \lambda\bar{\hat{q}}\right)\right]d\hat{q}.$$

as $\int_0^\infty \left[1 - F(\hat{q})\right]d\hat{q} = \mathbb{E}(\hat{q})$.
By changing of variable as $x = \hat{q} - \lambda\bar{\hat{q}}$, which implies $dx = d\hat{q}$, we have

$$\mathbb{E}(\hat{q}) = \frac{\frac{\tau}{g}}{1 + \frac{\tau}{g}}\int_{-\lambda\bar{\hat{q}}}^\infty \left[1 - F\left(x\right)\right]dx = \frac{\tau}{g}\lambda\bar{\hat{q}}$$

Last equality follows from the fact that $F(x) = 0$ for $x \leq 0$. In equilibrium we have, $\bar{\hat{q}} = \mathbb{E}(\hat{q})$. Therefore

$$g = \tau\lambda.$$

∎

Appendix B: Online Appendix for "Innovation, Reallocation and Growth"

– Not for Publication unless Requested –

Daron Acemoglu   Ufuk Akcigit   Harun Alp

Nicholas Bloom   William Kerr

In this Appendix, we provide additional details from the robustness exercises reported in the text.

Estimation Results from the Robustness Exercises

B-1   Employment Weighted Sample

TABLE B-1: ESTIMATED PARAMETERS

#	Parameter	Description	Value
1.	ϕ	Fixed cost of operation	0.168
2.	θ^H	Innovative capacity of high-type firms	2.209
3.	θ^L	Innovative capacity of low-type firms	1.532
4.	θ^E	Innovative capacity of entrants	0.025
5.	α	Probability of being high-type entrant	0.917
6.	ν	Transition rate from high-type to low-type	0.258
7.	λ	Innovation step size	0.101
8.	φ	Exogenous destruction rate	0.039

TABLE B-2: MODEL AND DATA MOMENTS

#	Moments	Model	Data	#	Moments	Model	Data
1.	Firm exit (small-young)	0.104	0.107	10.	Sales growth (small-young)	0.079	0.079
2.	Firm exit (small-old)	0.102	0.077	11.	Sales growth (small-old)	0.020	0.019
3.	Firm exit (large-old)	0.039	0.036	12.	Sales growth (large-old)	-0.023	-0.022
4.	Trans. from large to small	0.025	0.010	13.	R&D to sales (small-young)	0.100	0.075
5.	Trans. from small to large	0.036	0.014	14.	R&D to sales (small-old)	0.066	0.048
6.	Prob. of small (cond on entry)	0.795	0.753	15.	R&D to sales (large-old)	0.066	0.055
7.	Emp. growth (small-young)	0.078	0.073	16.	5-year Entrant Share	0.361	0.393
8.	Emp. growth (small-old)	0.020	0.028	17.	Fixed cost-R&D labor ratio	3.284	5.035
9.	Emp. growth (large-old)	-0.023	-0.033	18.	Aggregate growth	0.022	0.022

B-2 Organic Sample that Excludes M&A Activities

TABLE B-3: PARAMETER ESTIMATES

#	Parameter	Description	Value
1.	ϕ	Fixed cost of operation	0.233
2.	θ^H	Innovative capacity of high-type firms	1.708
3.	θ^L	Innovative capacity of low-type firms	1.480
4.	θ^E	Innovative capacity of entrants	0.023
5.	α	Probability of being high-type entrant	0.806
6.	ν	Transition rate from high-type to low-type	0.213
7.	λ	Innovation step size	0.137
8.	φ	Exogenous destruction rate	0.030

TABLE B-4: MODEL AND DATA MOMENTS

#	Moments	Model	Data	#	Moments	Model	Data
1.	Firm exit (small-young)	0.097	0.112	10.	Sales growth (small-young)	0.086	0.102
2.	Firm exit (small-old)	0.091	0.067	11.	Sales growth (small-old)	0.049	0.014
3.	Firm exit (large-old)	0.030	0.022	12.	Sales growth (large-old)	-0.003	-0.003
4.	Trans. from large to small	0.021	0.009	13.	R&D to sales (small-young)	0.070	0.048
5.	Trans. from small to large	0.037	0.010	14.	R&D to sales (small-old)	0.065	0.061
6.	Prob. of small (cond on entry)	0.873	0.899	15.	R&D to sales (large-old)	0.057	0.035
7.	Emp. growth (small-young)	0.088	0.106	16.	5-year Entrant Share	0.319	0.381
8.	Emp. growth (small-old)	0.049	0.028	17.	Fixed cost-R&D labor ratio	4.592	5.035
9.	Emp. growth (large-old)	-0.002	-0.002	18.	Aggregate growth	0.022	0.022

B-3 Baseline Estimation without R&D Moments

TABLE B-5: ESTIMATED PARAMETERS

#	Parameter	Description	Value
1.	ϕ	Fixed cost of operation	0.215
2.	θ^H	Innovative capacity of high-type firms	1.711
3.	θ^L	Innovative capacity of low-type firms	1.407
4.	θ^E	Innovative capacity of entrants	0.030
5.	α	Probability of being high-type entrant	0.894
6.	ν	Transition rate from high-type to low-type	0.207
7.	λ	Innovation step size	0.130
8.	φ	Exogenous destruction rate	0.035

TABLE B-6: MODEL AND DATA MOMENTS

#	Moments	Model	Data	#	Moments	Model	Data
1.	Firm exit (small-young)	0.098	0.107	10.	Sales growth (small-young)	0.095	0.107
2.	Firm exit (small-old)	0.092	0.077	11.	Sales growth (small-old)	0.046	0.024
3.	Firm exit (large-old)	0.035	0.036	12.	Sales growth (large-old)	-0.005	-0.003
4.	Trans. from large to small	0.021	0.010	13.	R&D to sales (small-young)	-	-
5.	Trans. from small to large	0.037	0.014	14.	R&D to sales (small-old)	-	-
6.	Prob. of small (cond on entry)	0.853	0.753	15.	R&D to sales (large-old)	-	-
7.	Emp. growth (small-young)	0.096	0.106	16.	5-year Entrant Share	0.333	0.393
8.	Emp. growth (small-old)	0.046	0.035	17.	Fixed cost-R&D labor ratio	4.263	5.035
9.	Emp. growth (large-old)	-0.005	-0.005	18.	Aggregate growth	0.022	0.022

TABLE B-7: EXCLUDING R&D MOMENTS

x^{entry}	x^l	x^h	Φ^l	Φ^h	$\hat{q}_{l,\min}$	$\hat{q}_{h,\min}$	$\frac{L^{R\&D}}{L^S}$	τ	g	Wel
				Panel A. Baseline						
0.62	26.04	35.71	56.67	5.15	146.54	133.84	19.62	17.24	2.23	100.00
				Panel B. Social Planner						
0.75	26.52	44.29	10.46	41.24	217.08	29.67	32.88	21.79	2.82	103.63
			Panel C. Incumbent R&D and Operation ($s_i = -2\%, s_o = -69\%$)							
0.76	31.10	43.34	49.08	7.52	159.59	147.21	26.57	19.29	2.50	101.38

B-4 Manufacturing Sample

TABLE B-8: PARAMETER ESTIMATES

#	Parameter	Description	Value
1.	ϕ	Fixed cost of operation	0.448
2.	θ^H	Innovative capacity of high-type firms	0.277
3.	θ^L	Innovative capacity of low-type firms	0.058
4.	θ^E	Innovative capacity of entrants	0.017
5.	α	Probability of being high-type entrant	0.699
6.	ν	Transition rate from high-type to low-type	0.460
7.	λ	Innovation step size	0.452
8.	φ	Exogenous destruction rate	0.044

TABLE B-9: MODEL AND DATA MOMENTS

#	Moments	Model	Data	#	Moments	Model	Data
1.	Firm exit (small-young)	0.096	0.081	10.	Sales growth (small-young)	0.018	0.018
2.	Firm exit (small-old)	0.103	0.059	11.	Sales growth (small-old)	0.001	0.001
3.	Firm exit (large-old)	0.053	0.037	12.	Sales growth (large-old)	-0.059	0.008
4.	Trans. from large to small	0.037	0.011	13.	R&D to sales (small-young)	-	-
5.	Trans. from small to large	0.020	0.009	14.	R&D to sales (small-old)	-	-
6.	Prob. of small (cond on entry)	0.530	0.669	15.	R&D to sales (large-old)	-	-
7.	Emp. growth (small-young)	0.018	0.020	16.	5-year Entrant Share	0.390	0.425
8.	Emp. growth (small-old)	0.008	-0.003	17.	Fixed cost-R&D labor ratio	4.955	5.035
9.	Emp. growth (large-old)	-0.055	-0.008	18.	Aggregate growth	0.019	0.019

B-5 Model with Unskilled Overhead Labor

TABLE B-10: PARAMETER ESTIMATES

#	Parameter	Description	Value
1.	ϕ	Fixed cost of operation	0.219
2.	θ^H	Innovative capacity of high-type firms	1.925
3.	θ^L	Innovative capacity of low-type firms	1.404
4.	θ^E	Innovative capacity of entrants	0.030
5.	α	Probability of being high-type entrant	0.883
6.	ν	Transition rate from high-type to low-type	0.196
7.	λ	Innovation step size	0.140
8.	φ	Exogenous destruction rate	0.049
9.	β	Fraction of managers with a college degree or above	0.457

TABLE B-11: MODEL AND DATA MOMENTS

#	Moments	Model	Data	#	Moments	Model	Data
1.	Firm exit (small-young)	0.099	0.107	10.	Sales growth (small-young)	0.107	0.107
2.	Firm exit (small-old)	0.098	0.077	11.	Sales growth (small-old)	0.048	0.024
3.	Firm exit (large-old)	0.046	0.036	12.	Sales growth (large-old)	-0.006	-0.003
4.	Trans. from large to small	0.020	0.010	13.	R&D to sales (small-young)	0.108	0.064
5.	Trans. from small to large	0.039	0.014	14.	R&D to sales (small-old)	0.076	0.059
6.	Prob. of small (cond on entry)	0.807	0.753	15.	R&D to sales (large-old)	0.065	0.037
7.	Emp. growth (small-young)	0.104	0.106	16.	5-year Entrant Share	0.369	0.393
8.	Emp. growth (small-old)	0.047	0.035	17.	Fixed cost-R&D labor ratio	5.656	5.035
9.	Emp. growth (large-old)	-0.005	-0.005	18.	Aggregate growth	0.022	0.022

B-6 Model with Reallocation Cost

TABLE B-12: PARAMETER ESTIMATES

#	Parameter	Description	Value
1.	ϕ	Fixed cost of operation	0.201
2.	θ^H	Innovative capacity of high-type firms	1.840
3.	θ^L	Innovative capacity of low-type firms	1.287
4.	θ^E	Innovative capacity of entrants	0.017
5.	α	Probability of being high-type entrant	0.960
6.	ν	Transition rate from high-type to low-type	0.300
7.	λ	Innovation step size	0.134
8.	φ	Exogenous destruction rate	0.038

TABLE B-13: MODEL AND DATA MOMENTS

#	Moments	Model	Data	#	Moments	Model	Data
1.	Firm exit (small-young)	0.093	0.107	10.	Sales growth (small-young)	0.103	0.107
2.	Firm exit (small-old)	0.088	0.077	11.	Sales growth (small-old)	0.033	0.024
3.	Firm exit (large-old)	0.036	0.036	12.	Sales growth (large-old)	-0.005	-0.003
4.	Trans. from large to small	0.020	0.010	13.	R&D to sales (small-young)	0.090	0.064
5.	Trans. from small to large	0.037	0.014	14.	R&D to sales (small-old)	0.058	0.059
6.	Prob. of small (cond on entry)	0.841	0.753	15.	R&D to sales (large-old)	0.052	0.037
7.	Emp. growth (small-young)	0.099	0.106	16.	5-year Entrant Share	0.321	0.393
8.	Emp. growth (small-old)	0.033	0.035	17.	Fixed cost-R&D labor ratio	4.237	5.035
9.	Emp. growth (large-old)	-0.005	-0.005	18.	Aggregate growth	0.022	0.022

B-7 Model with Three Types

TABLE B-14: PARAMETER ESTIMATES

#	Parameter	Description	Value
1.	ϕ	Fixed cost of operation	0.229
2.	θ^H	Innovative capacity of high-type firms	1.802
3.	θ^M	Innovative capacity of medium-type firms	1.753
4.	θ^L	Innovative capacity of low-type firms	1.381
5.	θ^E	Innovative capacity of entrants	0.023
6.	α_H	Probability of being high-type entrant	0.105
7.	α_M	Probability of being medium-type entrant	0.855
8.	ν	Transition rate to low-type	0.215
9.	λ	Innovation step size	0.134
10.	φ	Exogenous destruction rate	0.036

TABLE B-15: MODEL AND DATA MOMENTS

#	Moments	Model	Data	#	Moments	Model	Data
1.	Firm exit (small-young)	0.096	0.107	10.	Sales growth (small-young)	0.100	0.107
2.	Firm exit (small-old)	0.092	0.077	11.	Sales growth (small-old)	0.037	0.024
3.	Firm exit (large-old)	0.035	0.036	12.	Sales growth (large-old)	-0.005	-0.003
4.	Trans. from large to small	0.021	0.010	13.	R&D to sales (small-young)	0.083	0.064
5.	Trans. from small to large	0.037	0.014	14.	R&D to sales (small-old)	0.063	0.059
6.	Prob. of small (cond on entry)	0.849	0.753	15.	R&D to sales (large-old)	0.056	0.037
7.	Emp. growth (small-young)	0.099	0.106	16.	5-year Entrant Share	0.329	0.393
8.	Emp. growth (small-old)	0.038	0.035	17.	Fixed cost-R&D labor ratio	4.386	5.035
9.	Emp. growth (large-old)	-0.005	-0.005	18.	Aggregate growth	0.022	0.022