# Innovation by Entrants and Incumbents*

Daron Acemoglu
MIT and CIFAR

Dan Cao
Georgetown University

April 2, 2011

**Abstract**

We extend the basic Schumpeterian endogenous growth model by allowing incumbents to undertake innovations to improve their products, while entrants engage in more "radical" innovations to replace incumbents. Our model provides a tractable framework for the analysis of growth driven by both entry of new firms and productivity improvements by continuing firms. Unlike in the basic Schumpeterian models, subsidies to potential entrants might decrease economic growth because they discourage productivity improvements by incumbents in response to increased entry, which may outweigh the positive effect of greater creative destruction. As the model features entry of new firms and expansion and exit of existing firms, it also generates a non-degenerate equilibrium firm size distribution. We show that, when there is also costly imitation preventing any sector from falling too far below the average, the stationary firm size distribution is Pareto with an exponent approximately equal to one (the so-called "Zipf distribution").

*We thank Sam Kortum, Erzo Luttmer, Ariel Pakes, and seminar participants at MIT and Toulouse Network on Information Technology Conference at Seattle for useful comments. We are particularly grateful to Xavier Gabaix for numerous useful suggestions at the early stages of this project. Financial support from the Toulouse Network on Information Technology is gratefully acknowledged.

# 1 Introduction

The endogenous technological change literature provides a coherent and attractive framework for modeling productivity growth at the industry and the aggregate level. It also enables a study of how economic growth responds to policies and market structure. A key aspect of the growth process is the interplay between innovations and productivity improvements by existing firms on the one hand and entry by more productive, new firms on the other. Existing evidence suggests that this interplay is important for productivity growth. For example, Bartelsman and Doms (2000) and Foster and Krizan (2000), among others, document that entry of new establishments (plants) accounts for about 25% of average TFP growth at the industry level, with the remaining productivity improvements accounted for by continuing establishments (Lentz and Mortensen (2008) find an even more important role for entry).

These issues, however, are difficult to address with either of the two leading approaches to endogenous technological change, the expanding variety models, e.g., Romer (1990), Grossman and Helpman (1991b), Jones (1995), and the Schumpeterian quality ladder models, e.g., Segerstrom and Dinopoloulos (1990), Aghion and Howitt (1992), Grossman and Helpman (1991a).[1] The expanding variety models do not provide a framework for directly addressing these questions.[2] The Schumpeterian models are potentially better suited to studying the interplay between incumbents and entrants as they focus on the process of creative destruction and entry. Nevertheless, because of Arrow's replacement effect (Arrow 1962), these baseline models predict that all innovation should be undertaken by new firms and thus does not provide a framework for the analysis of the remaining 75% of the productivity growth due to innovation by existing firms and establishments.[3] In fact, Schumpeter's own work not only emphasized the role of creative destruction in economic growth, but also the importance of large (here continuing) firms in innovation (see Schumpeter 1934, and Schumpeter 1942).

This paper provides a simple framework that combines these two ideas by Schumpeter and involves simultaneous innovation by new and existing establishments.[4] The model is a

[1] Klette and Kortum (2004) is an exception and will be discussed below.

[2] In the expanding variety models, the identity of the firms that are undertaking the innovation does not matter, so one could assume that it is the existing producers that are inventing new varieties, though this will be essentially determining the distribution of productivity improvements across firms by assumption.

[3] Models of step-by-step innovation, such as Aghion et al. (2001), Aghion et al. (2005a), and Acemoglu and Akcigit (2006), allow innovation by incumbents, but fix the set of firms competing within an industry, and thus do not feature entry. Aghion et al. (2005b) consider an extension of these models in which there is entry, but focus on how the threat of entry may induce incumbents to innovate.

[4] In the model, each firm will consist of a single plant, thus the terms "establishment," "plant" and "firm" can be used interchangeably. Clearly, models that distinguish between plants and firms made up of multiple plants would be better suited to empirical analysis of industry dynamics, but would also be more complex. We are following the bulk of the endogenous technological change literature in abstracting from

tractable (and minimal) extension of the textbook multisector Schumpeterian growth model. A given number of sectors produce inputs (machines) for the unique final good of the economy. In each sector, there is a quality ladder, and at any point in time, a single firm has access to the highest quality input (machine). This firm can increase its quality continuously by undertaking "incremental" R&D in order to increase productivity and profits. These R&D investments generate productivity growth by continuing firms. At the same time, potential entrants undertake "radical" R&D in order to create a better input and replace the incumbent.[5] A large case study literature on the nature of innovation, for example, Freeman (1982), Pennings and Buitendam (1987), Tushman and Anderson (1986) and Scherer (1984), documents how established firms are the main source of innovations that improve existing products, while new firms invest in more radical and "original" innovations (see also the discussion in Arrow 1974). Recent work by Akcigit and Kerr (2010) provides empirical evidence from the US Census of Manufacturers that large firms engage more in "exploitative" R&D, while small firms perform "exploratory" R&D (defined similarly to the notions of "incremental" and "radical" R&D here).

The dynamic equilibrium of this economy can be characterized in closed-form and leads to a number of interesting comparative static results. It generates endogenous growth in a manner similar to the standard endogenous technological change models, but the contribution of incumbent (continuing plants) and new firms to growth is determined in equilibrium. Although the parameters necessary for a careful calibration of the model are not currently available, the model can plausibly generate about 75% of productivity growth from innovations by continuing firms.

Despite the Schumpeterian character of the model, there may be a negative relationship between the rate of entry of new firms and the rate of aggregate productivity growth. This reflects the importance of productivity growth by incumbents. In particular, more entry makes incumbents less profitable and they respond by reducing their R&D investments. The resulting lower productivity growth by incumbents may outweigh the higher growth due to entry. Consequently, taxes or entry barriers on potential entrants may *increase* economic growth (while taxes on existing firms unambiguously reduce productivity growth). This result is particularly surprising since the underlying model is only a small variant of

this important distinction.

[5] Continuing firms do not invest in radical R&D because of Arrow's replacement effect, but generate productivity growth as they have access to a technology for improving the quality of their machines/products and have the incentives to do so. Etro (2004) provides an alternative model in which incumbents invest in R&D because, as in Aghion et al. (2001), Aghion et al. (2005a), and Acemoglu and Akcigit (2006), they are engaged in a patents race against entrants. He shows that the Arrow replacement effect disappears when incumbents are Stackelberg leaders in this race.

the baseline Schumpeterian growth model, where growth is entirely driven by entry of new firms. Of course, this result does not imply that entry barriers would be growth-enhancing in practice, but isolates a new effect of entry on productivity growth.

Since existing firms are involved in innovation and expand their sizes as they increase their productivity and there is entry and exit of firms, the model, despite its simplicity, also generates rich firm dynamics and an endogenous distribution of firm sizes. The available evidence suggests that firm size distribution, or its tail for firms above a certain size, can be approximated by a Pareto distribution with a shape coefficient close to one (i.e., the so-called "Zipf's distribution," where the fraction of firms with size greater than $S$ is proportional to $1/S$, e.g., Lucas 1978, Gabaix 1999, Axtell 2001).[6] We show that a slight variant of the model where costly imitation is also allowed (so that new firms enter into sectors that fall significantly below the average in terms of quality, ensuring an endogenous lower bound to quality), the stationary firm size distribution has a Pareto tail with a shape coefficient approximately equal to one.[7]

Our paper is most closely related to Klette and Kortum (2004). Klette and Kortum construct a rich model of firm and aggregate innovation dynamics. Their model is one of expanding product varieties and the firm size distribution is driven by differences in the number of products that a particular firm produces. Klette and Kortum assume that firms with more products have an advantage in discovering more new products. With this assumption, their model generates the same patterns as here and also matches additional facts about propensity to patent and differential survival probabilities by size. One disadvantage of this approach is that the firm size distribution is not driven by the dynamics of continuing plants (and if new products are interpreted as new plants or establishments, the Klette-Kortum model predicts that all productivity growth will be driven by entry of new plants, though this may be an extreme interpretation, since some new products are produced in existing

[6] Our model also implies that firm growth is consistent with the so-called "Gibrat's Law," which posits a unit root in firm growth and appears to be a good approximation to the data (see, for example, Sutton (1997) and Sutton (1998)), even though detailed analysis shows significant deviations from Gibrat's Law for small firms, see, for example, Hall (1987), Akcigit (2010) and Rossi-Hansberg and Wright (2007b). See also Dunne and Samuelson (1988), Dunne and Samuelson (1989) and Klepper (1996) for patterns of firm entry and exit. For evidence on firm size distribution, see the classic paper by Simon and Bonini (1958) and the recent evidence in Axtell (2001), and Rossi-Hansberg and Wright (2007b). The later finds a slight deviation from Zipf's law for firms with more than 10,000 employees. For the size distribution of cities, see, among others, Dobkins and Ioannides (1998), Gabaix (1999) and Eeckhout and Jovanovic (2002).

[7] Firm dynamics leading to this result are richer than those implied by many existing models, since firm growth is driven both by the expansion of incumbents and entry (both of which are fully endogenous). Nevertheless, as in standard Schumpeterian models, entrants are more productive, and hence larger, than incumbents. This feature can be relaxed by assuming that it takes entrants a while to reach their higher potential productivity, though this extension would significantly complicate the analysis (e.g., Freeman (1982)). See Luttmer (2010a) for a model with this feature.

plants). The current model is best viewed as complementary to Klette and Kortum (2004), and focuses on innovations by existing firm in the same line of business instead of the introduction of new products. In practice, both types of innovations appear to be important and it is plausible that existing large firms might be more successful in locating new product opportunities.[8] Nevertheless, both qualitative and some recent quantitative evidence suggest that innovation by existing firms and existing lines of products are more important. Abernathy (1980), Lieberman (1984), and Scherer (1984), among others, provide various case studies documenting the importance of innovations by existing firms and establishments in the same line of business (for example, Abernathy stresses the role of innovations by General Motors and Ford in the car industry). Empirical work by Bartelsman and Doms (2000) and Foster and Krizan (2000) also suggests that productivity growth by continuing establishments plays a major role in industry productivity growth, while Broda and Weinstein (1996) provide empirical evidence on the importance of improvements in the quality of products in international trade.

Other related papers include Lentz and Mortensen (2008), Klepper (1996) and Atkeson and Burstein (2010). Lentz and Mortensen (2008) extend Klette and Kortum's model by introducing additional sources of heterogeneity and estimate the model on Danish data. Klepper (1996) documents various facts about firm size, entry and exit decisions, and innovation, and provides a simple descriptive model that can account for these facts. The recent paper by Atkeson and Burstein (2010) also incorporates innovations by existing firms, but focuses on the implications for the relationship between trade opening and productivity growth. Atkeson and Burstein (2010) also discuss the interactions between incumbents and entrants, and the firm size distribution. There are some notable differences, however. First, in their model, as in Luttmer (2007), the main interaction between entrants and incumbents is through spillovers. Second, they characterize the stationary equilibrium and the firm size distribution numerically (while we focus on analytical characterization). None of these papers consider a Schumpeterian growth model with innovation both by incumbents and entrants that can be easily mapped to decomposing the contribution of new and continuing plants (firms) to productivity growth.

Another set of related papers include Jovanovic (1982), Hopenhayn (1992), Ericson and Pakes (1995), Melitz (2003), Rossi-Hansberg and Wright (2007b,a), Lagos (2001), and Luttmer (2004, 2007, 2010a,b). Many of these papers generate realistic firm size distributions based on productivity heterogeneity (combined with fixed costs of operation). They

[8] Scherer (1984), for example, emphasizes both the importance of innovation by continuing firms (and establishments) and that larger firms produce more products.

typically take the stochastic productivity growth process of firms as exogenous, whereas our focus here is on understanding how R&D decisions of firms shape the endogenous process of productivity growth. Luttmer's recent papers (2010a, 2010b) are particularly notable, as they also incorporate exit and entry decisions, and generate empirically realistic firm size distributions under the assumption that there are knowledge spillovers across firms (and with exogenous aggregate growth). Our model, despite being highly tractable and only a small deviation from the textbook Schumpeterian model, also generates a realistic firm size distribution (with a Pareto tail as in Luttmer (2007) for firm sizes, Gabaix (1999) and Cordoba (2008) for cities, or Benhabib et al. (2010) for wealth distribution). To the best of our knowledge, ours is the first paper to analytically characterize the stationary firm size distribution with fully endogenous growth rates (of both continuing firms and entrants).[9]

The rest of the paper is organized as follows. Section 2 presents the basic environment and characterizes the equilibrium. Section 3 looks at the effects of policy on equilibrium growth and briefly characterizes the Pareto optimal allocation in this economy and compares it to the equilibrium. In Section 4, we characterize the equilibrium firm size distribution. Section 5 presents some numerical simulations of the model and shows that under some plausible parameterizations the model generates a large fraction of productivity growth driven by incumbents. Section 6 concludes, while the Appendix contains several proofs omitted from the text and some additional results.

# 2 Baseline Model

## 2.1 Environment

The economy is in continuous time and admits a representative household with the standard constant relative risk aversion (CRRA) preference

$$\int_0^\infty e^{-\rho t} \frac{C(t)^{1-\theta} - 1}{1-\theta} dt,$$

where $\theta$ is the coefficient of relative risk aversion or the inverse of the intertemporal elasticity of substitution.

Population is constant at $L$ and labor is supplied inelastically. The resource constraint at time $t$ takes the form

$$C(t) + X(t) + Z(t) \leq Y(t), \tag{1}$$

where $C(t)$ is consumption, $X(t)$ is aggregate spending on machines, and $Z(t)$ is total expenditure on R&D at time $t$.

[9] See Luttmer (2010c) for a review of the current literature.

There is a continuum of machines (inputs) normalized to 1 used in the production of a unique final good. Each machine line is denoted by $\nu \in [0,1]$. The production function of the unique final good is given by:

$$Y(t) = \frac{1}{1-\beta}\left[\int_0^1 q(\nu,t)^{\beta} x(\nu,t|q)^{1-\beta} d\nu\right] L^{\beta}, \tag{2}$$

where $x(\nu,t|q)$ is the quantity of the machine of type $\nu$ of quality $q$ used in the production process. This production process implicitly imposes that only the highest quality machine will be used in production for each type of machine $\nu \in [0,1]$.

Throughout, the price of the final good at each point is normalized to 1 (relative prices of final goods across different periods being determined by the interest rate).

The engine of economic growth here will be two forms of process innovations that lead to quality improvements: (1) Innovation by incumbents. (2) Creative destruction by entrants. Let $q(\nu,t)$ be the quality of machine line $\nu$ at time $t$. We assume the following "quality ladder" for each machine type:

$$q(\nu,t) = \lambda^n q(\nu,s) \text{ for all } v \text{ and } t,$$

where $\lambda > 1$ and $n$ is the number of incremental innovations on this machine line between time $s \leq t$ and time $t$, where time $s$ is the date at which this particular machine type was first invented and $q(\nu,s)$ refers to its quality at that point. The incumbent has a fully enforced patent on the machines that it has developed (though this patent does not prevent entrants leapfrogging the incumbent's quality). We assume that at time $t = 0$, each machine line starts with some quality $q(\nu,0) > 0$ owned by an incumbent with a fully-enforced patent on this initial machine quality.

Incremental innovations can only be performed by the incumbent producer.[10] So we can think of those as "tinkering" innovations that improve the quality of the machine. If the current incumbent spends an amount $z(\nu,t)\, q(\nu,t)$ of the final good for this type of innovation on a machine of current quality $q(\nu,t)$, it has a flow rate of innovation equal to $\phi(z(\nu,t))$, where $\phi(z)$ is strictly increasing, concave in $z$ and satisfies the following Inada-type assumption:[11]

$$\phi(0) = 0 \text{ and } \phi'(0) = \infty.$$

[10] This is similar to the assumption in section 7.3 in Barro and Sala-i Martin (2004) that the incumbents have some cost advantage in conducting research. Here we make an extrem assumption that the cost of incremental research is infinite to the entrants. However, incumbents and entrants have equal cost in performing radical research.

[11] More formally, this implies that for any interval $\Delta t > 0$, the probability of one incremental innovation is $\phi(z(\nu,t))\,\Delta t$ and the probability of more than one incremental innovation is $o(\Delta t)$ with $o(\Delta t)/\Delta t \to 0$ as $\Delta t \to 0$.

Recall that such an innovation results in a proportional improvement in quality and the resulting new machine will have quality $\lambda q(\nu, t)$.

The alternative to incremental innovations are *radical innovations*. A new firm (entrant) can undertake R&D to innovate over the existing machines in machine line $\nu$ at time $t$.[12] If the current quality of machine is $q(\nu, t)$, then by spending one unit of the final good, this new firm has a flow rate of innovation equal to $\frac{\eta(\widehat{z}(\nu,t))}{q(\nu,t)}$, where $\eta(\cdot)$ is a strictly decreasing, continuously differentiable function, and $\widehat{z}(\nu, t)\, q(\nu, t)$ is the total amount of R&D by new entrants towards machine line $\nu$ at time $t$. The presence of the strictly decreasing function $\eta$ captures the fact that when many firms are undertaking R&D to replace the same machine line, they are likely to try similar ideas, thus there will be some amount of "external" diminishing returns (new entrants will be "fishing out of the same pond"). Since each entrant attempting R&D on this line is potentially small, they will all take $\widehat{z}(\nu, t)$ as given. Given the total amount of R&D,$\widehat{z}(\nu, t)\, q(\nu, t)$, by new entrants and the asumption on the flow rate of innovation, the rate of radical innovations realized in machine $\nu$ at time $t$ is $\widehat{z}(\nu, t)\, \eta(\widehat{z}(\nu, t))$. This implicit assumption that the cost of radical innovations is proportional to the current quality level in a sector is consistent with the assumption on the innovation technology, $\phi(.)$ of the incumbents.

Throughout we assume that $z\eta(z)$ is *strictly increasing* in $z$ so that greater aggregate R&D towards a particular machine line increases the overall likelihood of discovering a superior machine. We also suppose that $\eta(z)$ satisfies the following Inada-type assumptions

$$\lim_{z\to 0} z\eta(z) = 0 \text{ and } \lim_{z\to\infty} z\eta(z) = \infty.$$

An innovation by an entrant leads to a new machine of quality $\kappa q(\nu, t)$, where $\kappa > \lambda$. This is the sense in which innovation by entrants are more "radical" than those of incumbents. Existing empirical evidence from studies of innovation support the notion that innovations by new entrants are more significant or radical than those of incumbents.[13] We assume that whether the entrant was a previous incumbent or not does not matter for its technology of innovation or for the outcome of its innovation activities.

Simple examples of functions $\phi(\cdot)$ and $\eta(\cdot)$ that satisfy the requirements above are

$$\phi(z) = Az^{1-\alpha} \text{ and } \eta(z) = Bz^{-\gamma}, \tag{3}$$

[12] Incumbents could also access the technology for radical innovations, but would choose not to. Arrow's replacement effect implies that since entrants make zero or negative profits from this technology (because of free entry), the profits of incumbents, who would be replacing their own product, would be negative. Incumbents will still find it profitable to use the technology for incremental innovations, which is not available to entrants.

[13] However, it may take a while for the successful entrants to realize the full productivity gains from these innovations (e.g., Freeman 1982). We are abstracting from this aspect.

with $\alpha, \gamma \in (0,1)$. We will sometimes use these functional forms to illustrate some of the conditions and the results we present below.

Now we turn to describing the production technology. Once a particular machine of quality $q(\nu, t)$ has been invented, any quantity of this machine can be produced at constant marginal cost $\psi$. We normalize $\psi \equiv 1 - \beta$ without loss of any generality, which simplifies the expressions below. This implies that the total amount of expenditure on the production of intermediate goods at time $t$ is

$$X(t) = (1-\beta) \int_0^1 x(\nu, t)\, d\nu, \tag{4}$$

where $x(\nu, t)$ is the quantity of this machine used in final good production. Similarly, the total expenditure on R&D is

$$Z(t) = \int_0^1 \left[ z(\nu, t) + \widehat{z}(\nu, t) \right] q(\nu, t)\, d\nu, \tag{5}$$

where $q(\nu, t)$ refers to the highest quality of the machine of type $\nu$ at time $t$. Notice also that total R&D is the sum of R&D by incumbents and entrants ($z(\nu, t)$ and $\widehat{z}(\nu, t)$ respectively). Finally, define $p^x(\nu, t|q)$ as the price of machine type $\nu$ of quality $q(\nu, t)$ at time $t$. This expression stands for $p^x(\nu, t|q(\nu, t))$, but there should be no confusion in this notation since it is clear that $q$ refers to $q(\nu, t)$, and we will use this notation for other variables as well (and moreover, we also write $z(\nu, t)$ and $\widehat{z}(\nu, t)$ without conditioning on the type and quality of the machine at which R&D is directed, since this will not cause any confusion and simplifies the notation).

## 2.2 Equilibrium Definitions

An allocation in this economy consists of time paths of consumption levels, aggregate spending on machines, and aggregate R&D expenditure $[C(t), X(t), Z(t)]_{t=0}^{\infty}$, time paths for R&D expenditure by incumbents and entrants $[z(\nu, t), \widehat{z}(\nu, t)]_{v \in [0,1], t=0}^{\infty}$, time paths of prices and quantities of each machine and the net present discounted value of profits from that machine, $[p^x(\nu, t|q), x(\nu, t), V(\nu, t|q)]_{v \in [0,1], t=0}^{\infty}$, and time paths of interest rates and wage rates, $[r(t), w(t)]_{t=0}^{\infty}$. An equilibrium is an allocation where R&D decisions by entrants maximize their net present discounted value; pricing, quantity and R&D decisions by incumbents maximize their net present discounted value; the representative household maximizes utility; final good producers maximize profits; and the labor and final good markets clear.

Let us start with the aggregate production function for the final good production. Profit maximization by the final good sector implies that the demand for the highest available quality of machine $\nu \in [0,1]$ at time $t$ is given by

$$x(\nu,t) = p^x(\nu,t|q)^{-1/\beta} q(\nu,t) L \text{ for all } \nu \in [0,1] \text{ and all } t. \tag{6}$$

The price $p^x(\nu,t|q)$ will be determined by the profit maximization of the monopolist holding the patent for machine of type $\nu$ and quality $q(\nu,t)$. Note that the demand from the final good sector for machines in (6) is iso-elastic, so the unconstrained monopoly price is given by the usual formula with a constant markup over marginal cost. Throughout, we assume that

$$\kappa \geq \left(\frac{1}{1-\beta}\right)^{\frac{1-\beta}{\beta}}, \tag{7}$$

so that after an innovation by an entrant, there will not be limit pricing. Instead, the entrant will be able to set the unconstrained profit-maximizing (monopoly) price. By implication, an entrant that innovates further after its own initial innovation will also be able to set the unconstrained monopoly price.[14] Condition (7) also implies that, when the highest quality machine is sold at the monopoly price, the final good sector will only use this machine type and thus justifies the form of the final good production function in (2) which imposes that only the highest quality machine in each line will be used.

Since the demand for machines in (6) is iso-elastic and $\psi \equiv 1-\beta$, the profit-maximizing monopoly price is

$$p^x(\nu,t|q) = 1. \tag{8}$$

Combining this with (6) implies

$$x(\nu,t|q) = qL. \tag{9}$$

Consequently, the flow profits of a firm with the monopoly rights on the machine of quality $q$ can be computed as:

$$\pi(\nu,t|q) = \beta qL. \tag{10}$$

[14] In this analysis, we are ignoring the incumbent's potential incentives, after being replaced, to continue to invest in "incremental" innovations with the hope of eventually catching up with a new entrant. This is without much loss of generality, since such investment is unlikely to be profitable. In particular, let $\widetilde{V}(\nu,t|q)$ denote the value of a just replaced incumbent. Then its optimal investment in incremental innovation is given by $\widetilde{z} = \arg\max_{z\geq 0} \phi(z)\left(\widetilde{V}(\nu,t|\lambda q) - \widetilde{V}(\nu,t|q)\right) - zq$. We have $\widetilde{V}(\nu,t|\lambda q) - \widetilde{V}(\nu,t|q) \leq \widetilde{V}(\nu,t|\lambda q) \leq V(\nu,t|\lambda q) = v\lambda q$, where $V$ is defined in (15) as the value function of an incumbent which has not been replaced by an entrant; and $V(\nu,t|\lambda q) = v\lambda q$ in a balanced growth path. From the first-order condition on $z$ we have $\widetilde{z} = (\phi')^{-1}\left(\frac{q}{\widetilde{V}(\nu,t|\lambda q) - \widetilde{V}(\nu,t|q)}\right) \leq (\phi')^{-1}\left(\frac{1}{v\lambda}\right) = \overline{z}$. Hence, the condition

$$\phi(\overline{z}) \leq \eta(\widehat{z})$$

is sufficient (though of course not necessary) to ensure that it is more profitable to invest in R&D for radical innovations (where $\eta(\widehat{z})$ is the equilibrium rate of success in such innovations derived below). This condition is thus also sufficient to ensure that it is not profitable for just-replaced incumbents to invest in incremental innovations.

Next, substituting (9) into (2), we obtain that total output is given by

$$Y(t) = \frac{1}{1-\beta} Q(t) L, \tag{11}$$

where

$$Q(t) \equiv \int_0^1 q(\nu, t)\, d\nu \tag{12}$$

is the average total quality of machines and will be the only state variable in this economy. Since we have assumed that $q(\nu, 0) > 0$ for all $\nu$, (12) also implies $Q(0) > 0$ as the relevant initial condition of our economy.[15]

As a byproduct, we also obtain that the aggregate spending on machines is

$$X(t) = (1-\beta) Q(t) L. \tag{13}$$

Moreover, since the labor market is competitive, the wage rate at time $t$ is

$$w(t) = \frac{\partial Y}{\partial L} = \frac{\beta}{1-\beta} Q(t). \tag{14}$$

To characterize the full equilibrium, we need to determine R&D effort levels by incumbents and entrants. To do this, let us write the net present value of a monopolist with the highest quality of machine $q$ at time $t$ in machine line $\nu$:

$$V(v, t|q) = \mathbb{E}_t \left[ \int_t^{T(\nu,t)} e^{-\int_t^{t+s} r(t+\widetilde{s}) d\widetilde{s}} \left( \pi(\nu, t+s|q) - z(\nu, t+s)\, q(t+s) \right) ds \right], \tag{15}$$

where the quality $q(\nu, t+s)$ follows a Poisson process such that $q(\nu, t+s+\Delta s) = \lambda q(\nu, t+s)$ with probability $\phi(z(\nu, t+s))\,\Delta s$ (obviously with $\Delta s$ infinitesimal), and $T(\nu, t)$ is a stopping time where a new entrant enters into the sector $\nu$. So if the R&D of the entrants into the sector is $\widehat{z}(\nu, t+s_1)$, then the distribution of $T(\nu, t)$ is

$$\Pr(T(\nu, t) \geq t+s) = \mathbb{E}_t \left[ e^{-\int_0^s \widehat{z}(\nu, t+s_1) \eta(\widehat{z}(\nu, t+s_1)) ds_1} \right].$$

Under optimal R&D choice of the incumbents, their value function $V(\nu, t|q)$ defined in (15) satisfies the standard Hamilton-Jacobi-Bellman equation:

$$\begin{aligned} r(t) V(\nu, t|q) - \dot{V}(\nu, t|q) = \max_{z(\nu,t)\geq 0} \{ & \pi(\nu, t|q) - z(\nu, t)\, q(\nu, t) \\ & + \phi(z(\nu, t)) \left( V(\nu, t|\lambda q) - V(\nu, t|q) \right) - \widehat{z}(\nu, t)\, \eta(\widehat{z}(\nu, t))\, V(\nu, t|q) \}, \end{aligned} \tag{16}$$

[15] One might be worried about whether the average quality $Q(t)$ in (12) is well-defined, since we do not know how $q(\nu, t)$ will look like as a function of $\nu$ and the function $q(\cdot, t)$ may not be integrable. This is not a problem in the current context, however. Since the index $\nu$ has no intrinsic meaning, we can rank the $\nu$'s such that $\nu \mapsto q(\nu, t)$ is nondecreasing. Then the average in (12) exists when defined as a Lebesgue integral.

where $\widehat{z}(\nu,t)\,\eta(\widehat{z}(\nu,t))$ is the rate at which radical innovations by entrants occur in sector $\nu$ at time $t$ and $\phi(z(\nu,t))$ is the rate at which the incumbent improves its technology. The first term in (16), $\pi(\nu,t|q)$, is flow of profit given by (10), while the second term is the expenditure of the incumbent for improving the quality of its machine. The second line includes changes in the value of the incumbent due to innovation either by itself (at the rate $\phi(z(v,t))$, the quality of its product increases from $q$ to $\lambda q$) or by an entrant (at the rate $\widehat{z}(\nu,t)\,\eta(\widehat{z}(\nu,t))$, the incumbent is replaced and receives zero value from then on).[16] The value function is written with a maximum on the right hand side, since $z(\nu,t)$ is a choice variable for the incumbent.

Free entry by entrants implies that we must have:[17]

$$\begin{aligned} \eta(\widehat{z}(\nu,t))\,V(\nu,t|\kappa q(\nu,t)) &\leq q(\nu,t)\,, \text{ and} \\ \eta(\widehat{z}(\nu,t))\,V(\nu,t|\kappa q(\nu,t)) &= q(\nu,t) \text{ if } \widehat{z}(\nu,t) > 0, \end{aligned} \tag{17}$$

which takes into account the fact that by spending an amount $q(\nu,t)$, the entrant generates a flow rate of innovation of $\eta(\widehat{z})$, and if this innovation is successful (flow rate $\eta(\widehat{z}(\nu,t))$), then the entrant will end up with a machine of quality $\kappa q$, thus earning the (net present discounted) value $V(\nu,t|\kappa q)$. The free entry condition is written in complementary-slackness form, since it is possible that in equilibrium there will be no innovation by entrants.

Finally, maximization by the representative household implies the familiar Euler equation,

$$\frac{\dot{C}(t)}{C(t)} = \frac{r(t)-\rho}{\theta}, \tag{18}$$

and the transversality condition takes the form

$$\lim_{t\to\infty} e^{-\int_0^t r(s)ds}\left[\int_0^1 V(\nu,t|q)\,d\nu\right] = 0. \tag{19}$$

This transversality condition follows because the total value of corporate assets is $\int_0^1 V(\nu,t|q)\,d\nu$. Even though the evolution of the quality of each machine is line is stochastic, the value of a machine of type $\nu$ of quality $q$ at time $t$, $V(\nu,t|q)$ is non-stochastic. Either $q$ is not the highest quality in this machine line, in which case the value function of the firm with a machine of quality $q$ is 0, or alternatively, $V(\nu,t|q)$ is given by (15).

We summarize the conditions for an equilibrium as follows:

[16] The fact that the incumbent has zero value from then on follows from the assumption that, after being replaced, a previous incumbent has no advantage relative to other entrants (see footnote 14).

[17] Since there is a continuum of machines $\nu \in [0,1]$, all optimality conditions should be more formally stated as "for all $\nu \in [0,1]$ except subsets of $[0,1]$ of zero Lebesgue measure" or as "almost everywhere". We will not add this qualification to simplify the notation and the exposition.

**Definition 1** *An equilibrium is given by time paths of* $\{C(t), X(t), Z(t)\}_{t=0}^{\infty}$ *that satisfy (1), (5), (13) and (19); time paths for R&D expenditure by incumbents and entrants,* $\{z(\nu,t), \hat{z}(\nu,t)\}_{v\in[0,1],t=0}^{\infty}$ *that satisfy (16) and (17); time paths of prices and quantities of each machine and the net present discounted value of profits,* $\{p^x(\nu,t|q), x(\nu,t|q), V(\nu,t|q)\}_{t=0}^{\infty}$ *given by (8), (9) and (16); and time paths of wage and interest rates,* $\{w(t), r(t)\}_{t=0}^{\infty}$ *that satisfy (14) and (18).*

In addition, we define a BGP (balanced growth path) as an equilibrium path in which innovation, output and consumption grow at a constant rate. Notice that in such a BGP, aggregates grow at the constant rate, but there will be firm deaths and births, and the firm size distribution may change. We will discuss the firm size distribution in Section 4 and will refer to a BGP equilibrium with a stationary (constant) distribution of normalized firm sizes as "a stationary BGP equilibrium". For now, we refer to an allocation as a BGP regardless of whether the distribution of (normalized) firm sizes is stationary.

**Definition 2** *A balanced growth path (hereafter BGP) is an equilibrium path in which innovation, output and consumption grow at a constant rate.*

In what follows, we will focus on *linear BGP*, where the value function of a firm with quality $q$ is linear in $q$, and often refer to it simply as the "BGP". In particular:[18]

**Definition 3** *A linear BGP is a BGP where* $V(\nu,t|q) = vq$ *for all* $\nu$, $t$ *(for some* $v > 0$*).*

## 2.3 Existence and Characterization

While a (linear) BGP always exists, because innovation by incumbents may increase the demand for the inputs of other incumbents (through what is sometimes referred to as "aggregate demand externalities"), there is a force pushing towards multiple BGPs. Counteracting this, greater innovation, by increasing the growth rate, also increases the interest rate and thus makes further innovation less profitable. In the remainder of the analysis, we focus on the case where the BGP is unique. The following is a sufficient condition for this.

**Assumption 1** *The intertemporal elasticity of substitution of the representative household,* $\theta$*, is sufficiently high, i.e.,*

$$\theta \geq \frac{1}{1 + \min_{z\geq 0}\left\{\frac{(\phi'(z))^2}{-\phi''(z)\phi(z)}\right\}}.$$

[18] We conjecture that all BGPs are in fact linear, but we are unable to prove this except when $\phi$ is linear.

Intuitively, when the intertemporal elasticity of substitution is higher, from the Euler equation, (18), an increase in growth rate generated by an increase in innovation leads to a greater rise in interest rate. This makes innovation by other incumbents less profitable, ensuring that the second force mentioned above dominates the first one.

Assumption 1 is not very restrictive. For example, if $\phi(\cdot)$ is linear, this assumption simply requires $\theta \geq 0$, which is of course always satisfied. If, on the other hand, $\phi(\cdot)$ has the functional form in (3), this assumption requires $\theta \geq \alpha$.

The requirement that consumption grows at a constant rate in the BGP implies that $r(t) = r$, from (18). Now focusing on linear BGP, where $V(q) = vq$, we have that $\dot{V}(\nu, t|q) = 0$. Hence the functional equation that determines the value of incumbent firms (16) can be written as

$$rv = \beta L + \max_{z \geq 0} \{\phi(z)(\lambda - 1)v - z\} - \widehat{z}\eta(\widehat{z})v \tag{20}$$

and assuming positive entry rate, the free-entry condition (17) can be written as

$$\eta(\widehat{z})\kappa v = 1. \tag{21}$$

Let $z(v) \equiv \arg\max_{z \geq 0} \phi(z)(\lambda - 1)v - z$ and $\widehat{z}(v) \equiv \eta^{-1}\left(\frac{1}{\kappa v}\right)$. Then clearly $z(v)$ is strictly increasing in $v$ (recall that $\phi(z)$ is strictly concave) and $\widehat{z}(v)$ is strictly increasing in $v$ (recall that $\eta(z)$ is decreasing in $z$). Moreover, since $z\eta(z)$ is strictly increasing in $z$, $\widehat{z}(v)\eta(\widehat{z}(v))$ is strictly increasing in $v$. From the Euler equation (18), we also have $\dot{C}(t)/C(t) = (r - \rho)/\theta = g$, where $g$ is the growth rate of consumption and output.

From (11), the growth rate of output can be expressed as

$$\frac{\dot{Y}(t)}{Y(t)} = \frac{\dot{Q}(t)}{Q(t)}.$$

From (20) and (21), in a linear BGP, for all machines, incumbents and entrants will undertake constant R&D $z(v)$ and $\widehat{z}(v)$, respectively. Consequently, in a small interval of time $\Delta t$, there will be $\phi(z(v))\Delta t$ sectors that experience one innovation by the incumbent (increasing their productivity by $\lambda$) and $\widehat{z}(v)\eta(\widehat{z}(v))\Delta t$ sectors that experience replacement by new entrants (increasing productivity by factor of $\kappa$). The probability that there will be two or more innovations of any kind within an interval of time $\Delta t$ is $o(\Delta t)$. Therefore, we have

$$\begin{aligned} Q(t + \Delta t) &= (\lambda\phi(z(v))\Delta t) Q(t) + (\kappa\widehat{z}(v)\eta(\widehat{z}(v))\Delta t) Q(t) \\ &\quad + (1 - \phi(z(v))\Delta t - \widehat{z}(v)\eta(\widehat{z}(v))\Delta t) Q(t) + o(\Delta t). \end{aligned}$$

Now substracting $Q(t)$ from both sides, dividing $\Delta t$ and taking the limit as $\Delta t \to 0$, we obtain

$$\frac{\dot{Q}(t)}{Q(t)} = (\lambda - 1)\phi(z(v)) + (\kappa - 1)\widehat{z}(v)\eta(\widehat{z}(v)).$$

Thus

$$g = \phi(z(v))(\lambda - 1) + \widehat{z}(v)\eta(\widehat{z}(v))(\kappa - 1). \tag{22}$$

Finally, equations (18), (20) and (22) together give us a single equation that determines $v$ and thus the key value function $V(\nu, t|q)$ in (16):

$$\beta L = \rho v + (\theta - 1)\phi(z(v))(\lambda - 1)v + z(v) + (\theta(\kappa - 1) + 1)\widehat{z}(v)\eta(\widehat{z}(v))v \tag{23}$$

The right-hand side of this equation is equal to 0 at $v = 0$ and goes to $+\infty$ as $v$ goes to $+\infty$. Thus, a value of $v^* > 0$ satisfying this equation, and thus a linear BGP, always exists. Moreover, Assumption 1 implies that the right-hand side is strictly increasing, so that this $v^* > 0$, and thus the BGP, is unique. Given $v^*$, the other equilibrium objects are easy to compute. The R&D rates of incumbents and entrants are simply given by

$$z^* = z^*(v^*) \tag{24}$$

$$\widehat{z}^* = \widehat{z}^*(v^*), \tag{25}$$

and the GDP growth rate is

$$g^* = \phi(z^*)(\lambda - 1) + \widehat{z}^*\eta(\widehat{z}^*)(\kappa - 1), \tag{26}$$

while the BGP interest rate, again from the Euler equation (18), is obtained as

$$r^* = \rho + \theta g^*. \tag{27}$$

Notice that equation (26) shows the decomposition of the aggregate growth rate, $g^*$, into two components on the right hand side. The first term, $\phi(z^*)(\lambda - 1)$, comes from innovation by incumbents. The second term, $\widehat{z}^*\eta(\widehat{z}^*)(\kappa - 1)$, comes from the innovation of the entrants.

The final step is to verify that the transversality condition of the representative household, (19) is satisfied. The condition for this is $r^* > g^*$ which is also satisfied if $\theta \geq 1$ (or less stringently, if $\rho > [\phi(z^*)(\lambda - 1) + \widehat{z}^*\eta(\widehat{z}^*)(\kappa - 1)]/(1 - \theta)$). This discussion thus establishes the following proposition (proof in the text).

**Proposition 1** *Suppose that Assumption 1 holds and* $\rho > [\phi(z^*)(\lambda - 1) + \widehat{z}^*\eta(\widehat{z}^*)(\kappa - 1)]/(1 - \theta)$. *Then there exists a unique linear BGP with the value function of an incumbent with quality* $q$ *given by* $V(q) = v^* q$, *where* $v^*$ *is the unique solution to (23), the aggregate growth rate* $g^*$ *is given by (26), and the interest rate* $r^*$ *is given by (27). Starting with any initial condition, the economy immediately jumps to this BGP (i.e., always grows at the rate* $g^*$*).*

Another set of interesting implications of this model concerns firm size dynamics. The size of a firm can be measured by its sales, which is equal to $x(\nu, t \mid q) = qL$ for all $\nu$ and $t$. We have seen that the quality of an incumbent firm increases at the flow rate $\phi(z^*)$, with $z^*$ given by (24), while the firm is replaced at the flow rate $\hat{z}^*\eta(\hat{z}^*)$. Hence, for $\Delta t$ sufficiently small, the stochastic process for the size of a particular firm is given by

$$x(\nu, t + \Delta t \mid q) = \begin{cases} \lambda x(\nu, t \mid q) & \text{with probability } \phi(z^*)\Delta t + o(\Delta t) \\ 0 & \text{with probability } \hat{z}^*\eta(\hat{z}^*)\Delta t + o(\Delta t) \\ x(\nu, t \mid q) & \text{with probability } (1 - \phi(z^*)\Delta t - \hat{z}^*\eta(\hat{z}^*)\Delta t) + o(\Delta t) \end{cases} \tag{28}$$

for all $\nu$ and $t$. Firms therefore have random growth, and surviving firms expand on average. However, firms also face a probability of bankruptcy (extinction). In particular, denoting the probability that a particular incumbent firm that started production in machine line $\nu$ at time $s$ will be bankrupt by time $t \geq s$ by $P(t \mid s, \nu)$, we clearly have $\lim_{t\to\infty} P(t \mid s, \nu) = 1$, so that each firm will necessarily die eventually. The implications of equation (28) for the stationary firm size distribution will be discussed in Section 4. For now it suffices to say that this equation satisfies Gibrat's Law, which postulates that firm growth is independent of size (e.g., Sutton 1997, Gabaix 1999).[19]

## 2.4 Special Case: Linear $\phi(\cdot)$

The limiting environment where $\phi(z)$ is linear, i.e., $\phi(z) = \phi z$, is a useful special case. In this case, equation (16) implies $\phi(V(\lambda q) - V(q)) = 1$, otherwise the incumbents will undertake an infinite amount of R&D or no R&D at all. Therefore, the value of an incumbent with quality $q$ simplifies to

$$V(q) = \frac{q}{\phi(\lambda - 1)}. \tag{29}$$

Moreover, from the free-entry condition (again holding as equality from the fact that the equilibrium is interior), we have $\eta(\widehat{z})V(\kappa q) = q$. This equation implies that the BGP R&D level by entrants $\widehat{z}^*$ is implicitly defined by

$$\widehat{z}(q) = \widehat{z}^* = \eta^{-1}\left(\frac{\phi(\lambda - 1)}{\kappa}\right) \text{ for all } q > 0, \tag{30}$$

where $\eta^{-1}$ is the inverse of the $\eta(z)$ function. Since $\eta(\cdot)$ is strictly decreasing, so is $\eta^{-1}(\cdot)$. In a linear BGP, the fact that $V(\nu, t|q) = vq$ for all $\nu, t$ and $q$ together with (20) also implies

$$V(q) = \frac{\beta L q}{r^* + \widehat{z}\eta(\widehat{z})}. \tag{31}$$

[19] The most common form of Gibrat's Law involves firm sizes evolving according to the stochastic process $S_{t+1} = \gamma_t S_t + \varepsilon_t$, where $\gamma_t$ is a random variable with mean 1 and $\varepsilon_t$ is a random variable with mean zero. Both variables are orthogonal to $S_t$. The law of motion (28) is a special case of this general form.

Next, combining this equation with (29) we obtain the BGP interest rate as

$$r^* = \phi(\lambda - 1)\beta L - \widehat{z}^* \eta(\widehat{z}^*).$$

Therefore, the BGP growth rate of consumption and output in this case is obtained as:

$$g^* = \frac{1}{\theta}\left(\phi(\lambda - 1)\beta L - \widehat{z}^* \eta(\widehat{z}^*) - \rho\right). \tag{32}$$

Equation (32) already has some interesting implications. In particular, it determines the relationship between the rate of innovation by entrant $\widehat{z}^*$ and the BGP growth rate $g^*$. In standard Schumpeterian models, this relationship is positive. In contrast, here, since $\widehat{z}\eta(\widehat{z})$ is strictly increasing in $\widehat{z}$, we have:

**Remark 1** *There is a negative relationship between $\widehat{z}^*$ and $g^*$.*

We will see next that one of the implications of Remark 1 will be a positive relationship between entry barriers and growth (when $\phi(\cdot)$ is linear).

# 3 The Effects of Policy on Growth

We now briefly study the effects of different policies on equilibrium productivity growth and also characterize the Pareto optimal allocation in this economy.

## 3.1 Taxes and Entry Barriers

Since the model has a Schumpeterian structure (with quality improvements as the engine of growth and creative destruction playing a major role), it may be conjectured that entry barriers (or taxes on potential entrants) will have the most negative effect on economic growth. To investigate whether this is the case, let us suppose that there is a tax (or an entry barrier) $\tau_e$ on R&D expenditures by entrants and a tax $\tau_i$ on R&D expenditure by incumbents. Tax revenues are not redistributed back to the representative household (for example, they finance an additive public good). Alternatively, $\tau_e$ can also be interpreted not only as a tax or an entry barrier, but also as a more strict patent policy. To keep the analysis brief, we only focus on the case in which tax revenues are collected by the government rather than being rebated back to incumbents as patent fees.

Repeating the analysis in subsection 2.3 for the case of nonlinear return to R&D by the incumbents, we obtain the following equilibrium conditions

$$z_\tau(v) = \arg\max_{z \geq 0} \phi(z)(\lambda - 1)v - (1 + \tau_i)z \tag{33}$$

and

$$\widehat{z}_{\tau}(v)=\eta^{-1}\left(\frac{1+\tau_e}{\kappa v}\right), \tag{34}$$

where $z_{\tau}(v)$ and $\widehat{z}_{\tau}(v)$ are the incumbent and entrant R&D rates when the policy vector is $\tau=(\tau_i,\tau_e)$. Combining these equations with (18) and (22), we obtain again an equation that uniquely determines $v$:

$$\begin{aligned}\beta L &= \rho v+(\theta-1)\,\phi\left(z_{\tau}(v)\right)(\lambda-1)\,v+(1+\tau_i)z_{\tau}(v) \\ &\quad +\left(\theta\left(\kappa-1\right)+1\right)\widehat{z}_{\tau}(v)\,\eta\left(\widehat{z}_{\tau}(v)\right)v.\end{aligned} \tag{35}$$

The corresponding equations for the case with linear $\phi(\cdot)$ are similar and we omit them to save space. Using this characterization, we now establish:

**Proposition 2** 1. *The BGP growth rate is strictly decreasing in $\tau_i$ (the tax on incumbents).*

2. *If $|\phi''(\widehat{z}^*_{\tau})|<\frac{(\kappa-1)(\lambda-1)(\beta L)^2}{\rho(1+\tau_i)^2}$, then the BGP growth rate of the economy is strictly increasing in the tax rate on entrants, i.e., $dg^*/d\tau_e>0$. In general, however, an increase in R&D tax or entry barrier on entrants has ambiguous effects on growth.*

**Proof.** See the Appendix. ■

## 3.2 Pareto Optimal Allocations

We now briefly discuss the Pareto optimal allocation, which will maximize the utility of the representative household starting with some initial value of the average quality of machines $Q(0)>0$. As usual, we can think of this allocation as resulting from an optimal control problem by a social planner. There will be two differences between the decentralized equilibrium and the Pareto optimal allocation. The first is that the social planner will not charge a markup for machines. This will increase the value of machines and innovation to society. Second, the social planner will not respond to the same incentives in inducing entry (radical innovation). In particular, the social planner will not be affected by the "business stealing" effect, which makes entrants more aggressive because they wish to replace the current monopolist, and she will also internalize the negative externalities in radical research captured by the decreasing function $\eta$.

Let us first observe that the social planner will always "price" machines at marginal cost, thus in the Pareto optimal allocation, the quantities of machine used in final good production will be given by

$$x^S(v,t|q)=\psi^{-\frac{1}{\beta}}qL=(1-\beta)^{-\frac{1}{\beta}}qL.$$

Substituting this into (2), we obtain output in the Pareto optimal allocation as

$$Y^S(t) = (1-\beta)^{-\frac{1}{\beta}} Q^S(t) L,$$

where the superscript $S$ refers to the social planner's allocation and $Q^S(t)$ is the average quality of machines at time $t$ in this allocation. Part of this output will be spent on production machines and is thus not available for consumption or research. For this reason, it is useful to compute net output as

$$\begin{aligned}\overline{Y}^S(t) &= Y^S(t) - X^S(t) = (1-\beta)^{-\frac{1}{\beta}} Q^S(t) L - \psi (1-\beta)^{-\frac{1}{\beta}} Q^S(t) L \\ &= \beta (1-\beta)^{-\frac{1}{\beta}} Q^S(t) L.\end{aligned}$$

Given that the specification of the innovation possibilities frontier above consists of radical and incremental innovations, the evolution of average quality of machines is

$$\frac{\dot{Q}^S(t)}{Q^S(t)} = (\lambda-1)\phi\left(z^S(t)\right) + (\kappa-1)\hat{z}^S(t)\eta\left(\hat{z}^S(t)\right), \tag{36}$$

where $z^S(t)$ is the average rate of incumbent R&D and $\hat{z}^S(t)$ is the rate of entrant R&D chosen by the social planner. The total cost of R&D to the society is:[20]

$$\left(\phi\left(z^S(t)\right) + \hat{z}^S(t)\eta\left(\hat{z}^S(t)\right)\right) Q^S(t).$$

The maximization of the social planner can then be written as

$$\max \int_0^\infty e^{-\rho t} \frac{C^S(t)^{1-\theta} - 1}{1-\theta} dt$$

subject to (36) and the resource constraint, which can be written as

$$C^S(t) + \left(z^S(t) + \widehat{z}^S(t)\right) Q^S(t) \leq \beta (1-\beta)^{-\frac{1}{\beta}} Q^S(t) L.$$

We show in the Appendix that the social planner's problem always satisfies

$$(\lambda-1)\phi'\left(z^S\right) = (\kappa-1)\left(\eta\left(\widehat{z}^S\right) + \widehat{z}^S\eta'\left(\widehat{z}^S\right)\right). \tag{37}$$

Equations (37) shows that the trade-off between radical and incremental innovations for the social planner is different from the allocation in the competitive equilibrium[21]

$$(\lambda-1)\phi'(z) = \kappa\eta(\widehat{z}),$$

[20] Because of convexity, it is optimal for the social planner to choose the same (proportional) level of R&D investment in each sector, which we have imposed in writing this expression.

[21] This equation is derived by combining the first order condition in (20) and equation (21) in the competitive equilibrium.

because the social planner internalizes the negative effect that one more unit of R&D creates on the success probability of other firms performing radical R&D on the same machine line. This is reflected by the negative term $\widehat{z}^S \eta' \left(\widehat{z}^S\right)$ on the right-hand side of (37). This effect implies that the social planner will tend to do less radical innovation than the decentralized equilibrium. Since $z^S$ and $\widehat{z}^S$ are constant, consumption growth rate is also constant in the optimal allocation (thus no transitional dynamics). This Pareto optimal consumption growth rate can be greater than or less than the equilibrium BGP growth rate, $g^*$, because there are two counteracting effects. One the one hand, the social planner uses machines more intensively (because she avoids the monopoly distortions), and this tends to increase $g^S$ above $g^*$. This same effect can also encourage more radical R&D. On the other hand, the social planner also has a reason for choosing a lower rate of radical R&D because she internalizes the negative R&D externalities in research and the business stealing effect. One can construct examples in which the growth rate of the Pareto optimal allocation is greater or less than that of the decentralized equilibrium (though only in relatively rare cases is the equilibrium rate of Pareto optimal allocation smaller than that of the decentralized equilibrium). The following proposition illustrates these intuitions:

**Proposition 3** *There exists an $\epsilon > 0$ such that if $|\phi''(\cdot)| < \epsilon$, then the growth rate of the Pareto optimal allocation is greater than that of the BGP growth rate, while the R&D rate of entrants is lower than in the decentralized equilibrium. In general, however, the comparison is ambiguous.*

**Proof.** See the Appendix. ■

Proposition 3 suggests that when $\phi(\cdot)$ is close to linear, i.e., $|\phi''(\cdot)|$ is small, taxing entrants, by reducing their R&D, might move the decentralized equilibrium toward the social planner's allocation. This is investigated in the next proposition. In light of the importance of creative destruction in Schumpeterian models, the following can again be viewed as a paradoxical result.

**Proposition 4** *There exists an $\epsilon > 0$ such that, if $|\phi''(\widehat{z}^*_\tau)| < \epsilon$ and $\rho < \epsilon$, the welfare of the representative household is strictly increasing in $\tau_e$.*

**Proof.** See the Appendix. ■

Once again, we do not believe that erecting entry barriers or taxing entrants would be welfare improving in practice. Instead, we highlight this result to emphasize that the endogenous innovation responses by incumbents are potentially important.

# 4 Stationary BGP Equilibrium

As we pointed out above, the baseline model (cfr. Proposition 1) generates firm size dynamics due to growth by continuing firms and entry by new firms. So far, we have focused on the behavior of aggregate variables such as output or average quality. We now study the "stationary" distribution of firm sizes. We first show that in the baseline model, even though aggregate output is well behaved and grows at a constant rate, a stationary distribution does not exist (as time goes infinity, a vanishingly small fraction of firms become arbitrarily large, making the remaining firms arbitrarily small relative to average firm size in the economy). This is simply a consequence of the fact that because firm size growth follows Gibrat's Law, as shown by equation (28), the distribution of firm sizes will continuously expand. Yet this is partly an artificial result due to the fact that there is no lower (or upper bound) on firm size relative to average firm size in the economy. In practice, there will be several economic forces that pull up sectors that fall significantly below the average productivity in the economy. In subsection 4.2, we incorporate one such mechanism, (costly) *entry by imitation*: potential entrants can pay some cost to copy a technology with quality proportional to the current average quality in the economy. This implies that when a sector falls significantly below average quality, entry by imitation becomes profitable. We then show that the economy incorporating entry by imitation has a well-defined equilibrium. Moreover, when the imitation technology is not very productive (such that firms entering using the technology are initially sufficiently small), this equilibrium is arbitrarily close to the one that is characterized in Proposition 1 and the stationary firm size distribution has a Pareto tail with a shape parameter approximately equal to one (i.e., the so-called "Zipf distribution," which, as discussed above, appears to be a good approximation to US data).

## 4.1 Firm Size Distributions

Let us study the distribution of firm sizes as measured by sales, $x(\nu, t \mid q)$. Since the average firm size grow in this economy, we will focus on the behavior of firm sizes normalized by average size. In particular, let $X_1(t) \equiv X(t)/(1-\beta) = \int_0^1 x(\nu, t \mid q)\, d\nu$, where $X(t)$ was defined in (4) as total expenditures on machines/inputs.[22] Then, normalized firm size is

$$\widetilde{x}(\nu, t \mid q) = \frac{x(\nu, t \mid q)}{X_1(t)}.$$

[22]Sales normalized by the equilibrium wage rate, $w(t)$, have exactly the same behavior, since the equilibrium wage rate also scales with average quality, $Q(t)$.

Because sales are proportional to quality, we also have

$$\widetilde{x}(\nu, t \mid q) = \widetilde{x}(t) = \widetilde{q}(t) \equiv \frac{q(t)}{Q(t)},$$

where the first equality makes the dependence on sector and quality implicit to simplify notation and the last equality defines $\widetilde{q}(t)$. The law of motion of normalized firm size is similar to (28) introduced above, except that we are now keeping track of the (leading) firm producing in a given sector rather than a given firm. Hence, after entry (the second line), the relevant firm size is not zero, but is equal to the size of the new entrant. Noting that in BGP, $\dot{X}_1(t)/X_1(t) = g^* > 0$, for $\Delta t$ small, the law of motion of firm sizes can be written as

$$\widetilde{x}(t+\Delta t) = \begin{cases} \frac{\lambda}{1+g^*\Delta t}\widetilde{x}(t) & \text{with probability } \phi(z^*)\Delta t \\ \frac{\kappa}{1+g^*\Delta t}\widetilde{x}(t) & \text{with probability } \widehat{z}^*\eta(\widehat{z}^*)\Delta t \\ \frac{1}{1+g^*\Delta t}\widetilde{x}(t) & \text{with probability } 1-\phi(z^*)\Delta t - \widehat{z}^*\eta(\widehat{z}^*)\Delta t. \end{cases} \tag{38}$$

**Proposition 5** *In the baseline economy studied, there exists no stationary distribution.*

**Proof.** See the Appendix. ■

The essence of Proposition 5 is that with the random growth process in (38), the distribution of firm sizes will continuously expand.[23]

## 4.2 The Economy with Imitation

To ensure the existence of a stationary distribution with a minimal modification to the baseline economy, we introduce a third type of innovation, "imitation". A new firm can enter in any sector $\nu \in [0,1]$ with a technology $q^e(\nu, t) = \omega Q(t)$, where $\omega \geq 0$ and $Q(t)$ is average quality of machines in the economy given by (12) and we think of $\omega$ as small, capturing the fact that such imitation should only be profitable if the sector in question has fallen significantly below average quality in the economy. The cost of this type of innovation is assumed to be $\mu_e \omega Q(t)$. The fact that the cost should be proportional to average quality is in line with the structure of the model so far.[24]

Firm value is again given by (15) except that $T(\nu, t)$ in this equation is now the stopping time where either an entrant or an imitator enters and replaces the monopolist. Put differently, all firms solve the maximization problem as in (16), but they also take into account

[23] The nature of the "limiting distribution" is therefore similar to the "immiserization" result for income distribution in Atkeson and Lucas (1992) economy with dynamic hidden information; in the limit, all firms have approximately zero size relative to the average $X_1(t)$ and a vanishingly small fraction of firms become arbitrarily large (so that average firm size $X_1(t)$ remains large and continues to grow).

[24] This imitation technology captures the knowledge spillover channel as in Romer (1990). There are also alternative ways of ensuring a stationary firm size distribution. For example, in a companion paper we consider the case in which each firm has to pay a small fixed costs in terms of labor to operate. See also Luttmer (2007, 2010b).

the possibility of entry by imitation as well. It is then straightforward that there will exist some $\epsilon > 0$ such that entry by imitation is profitable if

$$q(\nu, t) \leq \epsilon Q(t),$$

and thus there will be no sectors with quality less than $\epsilon$ times average quality (clearly, $\epsilon$ is a function of $\omega$, and of course, $\epsilon = 0$ when $\omega = 0$). The baseline model is then the special case where $\omega = \epsilon = 0$, i.e., where there is no imitation. We will show that a BGP equilibrium and the stationary distribution in this economy with imitation are well-defined, and as $\omega \to 0$, the value function, the innovation decisions and the growth rate converge to those characterized in the baseline economy (cfr. Proposition 1). Moreover, for $\omega > 0$ but small, the stationary firm size distribution has a Pareto tail.

In order to prove the existence of a stationary BGP in the economy with imitation, we need a slightly stronger condition on $\theta$ than in Assumption 1

**Assumption 1b** $\theta \geq 1$.

We also impose the following technical condition.

**Assumption 2** *Let $\varepsilon_\eta(z) \equiv -z\eta'(z)/\eta(z)$ be the elasticity of the entry function $\eta(z)$. Then*

$$\max_{z>0} \varepsilon_\eta(z) \leq 1 - \frac{1}{\kappa^\theta}.$$

Under functional form (3) for $\eta$, Assumption 2 is equivalent to $\gamma \leq 1 - \kappa^{-\theta}$, and implies that the entry function is not too "elastic". This assumption is used in Lemma 4 in the Appendix to ensure boundedness of the value function of incumbent firms when both types of entry are present. When there is no entry by creative destruction, i.e., the economy with only the incumbents and the imitators, the same description of the stationary BGP goes through without Assumption 2. We also note for future reference that the Pareto distribution takes the form

$$\Pr[\tilde{x} \leq y] = 1 - \Gamma y^{-\chi}$$

with $\Gamma > 0$ and $y \geq \Gamma$; $\chi$ is the shape parameter (exponent) of the Pareto distribution. We say that a distribution has a Pareto tail if its behavior for $y$ large can be approximated by $\Pr[\tilde{x} \geq y] \propto \Gamma y^{-\chi}$.

**Proposition 6** *Suppose the BGP equilibrium in the baseline economy is described by Proposition 1, in particular with $v^*$, $g^*$ and $r^*$ as given by to (23), (26) and (27), and Assumptions 1b and 2 are satisfied. Then there exist $0 < \underline{\mu} < \overline{\mu}$, $\Delta > 0$ and $\overline{\omega} > 0$ such that for any $\mu_e \in (\underline{\mu}, \overline{\mu})$ and $\omega \in (0, \overline{\omega})$, there exists a BGP with the following properties:*

1. *There is entry by imitation whenever* $q(\nu, t) \leq \epsilon(\omega) Q(t)$, *where* $0 < \epsilon(\omega) \leq \omega(1-\beta)^{\frac{1-\beta}{\beta}}$.

2. *The equilibrium growth rate is* $g(\omega) \in (g^*, g^* + \Delta)$ *and satisfies*

$$\lim_{\omega \to 0} g(\omega) = g^*.$$

3. *The value function of the incumbents above the exit threshold* $\epsilon(\omega) Q(t)$ *in this economy, normalized by quality,* $V_\omega(q)/q$, *converges uniformly to the value function in the baseline economy normalized by quality,* $V(q)/q = v^*$. *Formally, for any* $\delta > 0$

$$\lim_{\omega \to 0} \sup_{q \geq \delta Q} \left| \frac{V_\omega(q) - V(q)}{q} \right| = 0.$$

This proposition ensures that the growth rate of the economy with imitation is well behaved and it is "close" to the equilibrium of the baseline economy when $\omega$ is small. Note also that our requirement $\mu_e \geq \underline{\mu}$ (together with $\omega \leq \overline{\omega}$) ensures that entry by imitation is not profitable when the entrants charge a limit price in the competition against the incumbent. As a result, entrants use imitation to enter in a sector only when the quality of the sector falls sufficiently below the average quality so that these entrants can charge the monopoly price after entry. The condition $\mu_e < \overline{\mu}$ ensures that this type of entry is not too costly so that there will be some imitation in equilibrium.

We next show that this economy admits a stationary distribution of normalized firm sizes with a Pareto tail with the shape parameter approaching 1 as $\omega$ becomes small.

**Proposition 7** *The stationary equilibrium distribution of firm sizes in the economy with imitation (characterized in Proposition 6) exists and has a Pareto tail with the shape parameter* $\chi = \chi(\omega) > 1$ *in the sense that for any* $\xi > 0$ *there exist* $\overline{B}$, $\underline{B}$ *and* $\widetilde{x}_0$ *such that the density function of the firm size distribution,* $f(\widetilde{x})$, *satisfies*

$$f(\widetilde{x}) < 2\overline{B}\widetilde{x}^{-(\chi-1-\xi)}, \text{ for all } \widetilde{x} \geq \widetilde{x}_0, \text{ and}$$

$$f(\widetilde{x}) > \frac{1}{2}\underline{B}\widetilde{x}^{-(\chi-1+\xi)}, \text{for all } \widetilde{x} \geq \widetilde{x}_0.$$

*In other words,* $f(\widetilde{x}) = \widetilde{x}^{-\chi-1}\varphi(\widetilde{x})$, *where* $\varphi(\widetilde{x})$ *is a slow-varying function. Moreover*

$$\lim_{\omega \to 0} \chi(\omega) = 1.$$

**Proof.** See the Appendix. ■

This result on the stationary firm size distribution has several parallels with existing results in the literature, for example, Gabaix (1999) and Luttmer (2007, 2010a,b). In particular, as in these papers, the stationary firm size distribution is obtained by combining firm growth following Gibrat's Law with a lower bound on (relative) firm size. These papers also have limiting results such that when the lower bound becomes negligible the size distribution converges to Zipf's law, i.e., the tail index converges to 1 as in our proposition.[25] There are also important differences, however. First, Gibrat's Law is derived endogenously here from the innovation decisions of continuing firms and entrants, and in fact, the growth rate of output in the aggregate is endogenously determined. Second, the equilibrium is obtained from the optimization problem of firms that recognize the possibility that there will be entry by imitation if their quality falls significantly relative to the average.

We next provide a sketch of the proof of Proposition 6.

## 4.3 Sketch of the Proof of Proposition 6

The proof consists of showing that for each $\mu_e \in (\underline{\mu}, \overline{\mu})$ and $\omega \in (0, \overline{\omega})$, there exists a BGP with the growth rate given by $g(\omega)$. The first step proves the existence of the value function of the incumbents under the threat of entry by imitation. In this step we show that the relevant state variable is the relative quality of the incumbents $q(\nu, t) / Q(t)$. The second step establishes the existence of and characterizes the form of the stationary firm size distribution when the incumbents and the entrants follow the strategies determined using the value function in the first step. Finally, the last step establishes the existence of a BGP with the value and investment functions derived from the first step and the stationary firm size distribution derived in the second step. Luttmer (2007, 2010b) follow similar steps in proving the existence of a BGP for an economy with heterogeneous firms, but relying on a specific closed-form of the value function and the stationary distribution. These closed-forms in turn exploit the fact that growth is exogenous, whereas the growth rate is determined endogenously in our economy.

There are two difficulties we must overcome in the first step. The first one is that the value functions are given by a differential equation with deviating (advanced) arguments because the right-hand side involves $V$ evaluated at $\lambda q$ and $\kappa q$. As a result, we cannot apply standard existence proofs from the theory of ordinary differential equations. Instead, we use techniques developed in the context of monotone iterative solution methods, see, for

[25] The gamma distribution used in Luttmer (2007) has a Pareto tail according to the definition in Proposition 7.

example, Jankowski (2005). The second difficulty arises because we need to show that the value function satisfies some properties at infinity. This non-standard boundary problem is solved following the approach in Staikos and Tsamatos (1985).

**Step 1**: For each $g \in (g^*, g^* + \Delta)$, we show the existence of a value function of an incumbent in sector $\nu$ at time $t$ which takes the form

$$V_g(\nu, t|q) = Q(t)\,\widehat{V}_g\left(\frac{q(\nu,t)}{Q(t)}\right), \tag{39}$$

and a threshold $\epsilon_g(\omega)$ such that an imitator will pay the cost $\mu_e \omega Q_t$ to imitate and enter with quality $\omega Q_t$ into sector $\nu$ at time $t$ and replace an incumbent if $q(\nu,t) \le \epsilon_g(\omega) Q(t)$. The value of the incumbent depends only on the current average quality, $Q(t)$, and the gap between its current quality and the average quality, $q(\nu,t)/Q(t)$. Plugging (39) in (16), and using the fact that

$$\dot{V}_g(\nu, t|q) = gQ(t)\,\widehat{V}_g\left(\frac{q(\nu,t)}{Q(t)}\right) - gQ(t)\,\widehat{V}_g'\left(\frac{q(\nu,t)}{Q(t)}\right),$$

we obtain that

$$\begin{aligned}(r-g)\,\widehat{V}_g(\widetilde{q}) - g\widehat{V}_g'(\widetilde{q}) \;&=\; \beta L\widetilde{q} + \max_{z(\nu,t)\ge 0}\left\{\phi(z(\nu,t))\left(\widehat{V}_g(\lambda\widetilde{q}) - \widehat{V}_g(\widetilde{q})\right) - z(\nu,t)\,\widetilde{q}\right\}\\ &\quad -\widehat{z}(\nu,t)\,\eta(\widehat{z}(\nu,t))\,\widehat{V}_g(\widetilde{q}),\end{aligned} \tag{40}$$

where $r = \rho + \theta g$ and $\widetilde{q}(\nu,t) = q(\nu,t)/Q(t)$. The free-entry condition for radical innovation, (17), can then be written as

$$\eta(\widehat{z}(\nu,t))\,\widehat{V}_g(\kappa\widetilde{q}(\nu,t)) = \widetilde{q}(\nu,t).$$

Moreover, the free-entry condition for imitation implies

$$\widehat{V}_g(\omega) = \mu_e \omega. \tag{41}$$

Since imitators will replace the incumbent in sector $\nu$ at time $t$ if $q(\nu,t) \le \epsilon_g Q(t)$, we also have the following boundary condition

$$\widehat{V}_g(\epsilon_g) = 0. \tag{42}$$

In the Appendix, we show that when $\mu_e \in \left(\underline{\mu}, \overline{\mu}\right)$, there is imitation in equilibrium but only when imitators can charge monopoly price after entry. Equilibrium innovation rates, $z_g(\widetilde{q})$ and $\widehat{z}_g(\widetilde{q})$, can then be derived from the solution to (40).

To establish the existence of a solution $\widehat{V}_g(\widetilde{q})$ to the functional equation (40), we first construct functional bounds, $\underline{V}_g$ and $\overline{V}_g$, such that $\underline{V}_g(\widetilde{q}) \leq \widehat{V}_g(\widetilde{q}) \leq \overline{V}_g(\widetilde{q})$. The result on uniform convergence of $\widehat{V}_g(\widetilde{q})$ then follows by establishing that $\underline{V}_g$ and $\overline{V}_g$ converge uniformly to $V(q)$ as $g$ goes to $g^*$.

**Step 2:** The innovation rates, $z_g(\widetilde{q})$ and $\widehat{z}_g(\widetilde{q})$, together with the entry rule of the imitators and the growth rate $g(\omega)$ of the average quality yields a stationary distribution over the normalized sizes $\widetilde{q}$ with distribution function $F(\cdot)$ satisfying the following conditions: If $y \geq \omega$, then

$$0 = F'(y)\, yg - \int_{\frac{y}{\lambda}}^{y} \phi(z(\widetilde{q}))\, dF(\widetilde{q}) - \int_{\frac{y}{\kappa}}^{y} \widehat{z}(\widetilde{q})\, \eta(\widehat{z}(\widetilde{q}))\, dF(\widetilde{q})\,. \tag{43}$$

If $y < \omega$, then

$$0 = F'(y)\, yg - F'(\epsilon)\, \epsilon g - \int_{\frac{y}{\lambda}}^{y} \phi(z(\widetilde{q}))\, dF(\widetilde{q}) - \int_{\frac{y}{\kappa}}^{y} \widehat{z}(\widetilde{q})\, \eta(\widehat{z}(\widetilde{q}))\, dF(\widetilde{q}) \tag{44}$$

and

$$F(y) = 0 \text{ for } y \leq \epsilon.$$

We will derive these expressions formally in the Appendix, and to make the dependence on the growth rate of average quality explicit, we will write the solution as $F_g$. Intuitively, given $y > 0$, the mass of firms with size moving out of the interval $(\epsilon, y)$ consists of firms (sectors) with size between $\left(\frac{y}{\lambda}, y\right)$ that are successful in incremental innovation, $\int_{\frac{y}{\lambda}}^{y} \phi(z(\widetilde{q}))\, dF(\widetilde{q})$, and firms (sectors) with size between $\left(\frac{y}{\kappa}, y\right)$, where there is a radical innovation, $\int_{\frac{y}{\kappa}}^{y} \widehat{z}(\widetilde{q})\, \eta(\widehat{z}(\widetilde{q}))\, dF(\widetilde{q})$. When $y < \omega$, we must also add the mass of firms being replaced by imitators with relative quality $\omega$. This mass consists of firms that are in the neighborhood of $\epsilon$, do not experience any innovation, and are therefore drifted to below $\epsilon$ due to the growth rate $g$ of the average quality $Q$; it is equal to $F'(\epsilon)\, \epsilon g$. By definition of a stationary distribution, the total mass of firms moving out of the interval $(\epsilon, y)$ must be equal to the mass of firms moving into the interval. This mass consists of firms around $y$ that do not experience any innovation and thus drift into this interval due to growth at the rate $g$ (given by $F'(y)\, yg$).

**Step 3:** From this analysis, we obtain an implied growth rate of the average product quality $g' = \dot{Q}/Q$ as a function of the current growth rate $g$, from the innovation rates, $z_g(\widetilde{q})$, and $\widehat{z}_g(\widetilde{q})$, imitation threshold $\epsilon_g$, and the equilibrium stationary distribution $F_g$. In particular,

$$g'(g) = \frac{(\lambda - 1)\, \mathbb{E}_{F_g}\left[\phi(z(\widetilde{q}))\, \widetilde{q}\right] + (\kappa - 1)\, \mathbb{E}_{F_g}\left[\widehat{z}_g(\widetilde{q})\, \eta(\widehat{z}_g(\widetilde{q}))\, \widetilde{q}\right]}{1 - \epsilon_g F_g'(\epsilon_g)\,(\omega - \epsilon_g)\,.}. \tag{45}$$

This formula, derived formally in the Appendix, is similar to the decomposition of growth in (26). The numerator combines the innovation rates of incumbents and entrants, respectively $(\lambda - 1)\,\mathbb{E}_{F_g}\left[\phi\left(z\left(\widetilde{q}\right)\right)\widetilde{q}\right]$ and $(\kappa - 1)\,\mathbb{E}_{F_g}\left[\widehat{z}_g\left(\widetilde{q}\right)\eta\left(\widehat{z}_g\left(\widetilde{q}\right)\right)\widetilde{q}\right]$, where $\mathbb{E}_{F_g}$ is used as a shorthand for the integrals using the density $dF_g\left(\widetilde{q}\right)$ as in (43) and (44). The denominator, on the other hand, is the contribution of imitation to growth. The higher is the gap $\omega - \epsilon_g$, the more important is this component. Finally the equilibrium growth rate $g^*\left(\omega\right)$ is a solution to the equation

$$D\left(g\right) \equiv g'\left(g\right) - g = 0,$$

where $g'\left(g\right)$ is given by (45). In the Appendix, we establish the existence of a solution $g^*\left(\omega\right)$ to this equation.

# 5 Simulations

## 5.1 Growth Decompositions

The explicit characterization of equilibrium enables us to obtain simple expressions for how much of productivity growth is driven by creative destruction (innovation by entrants) and how much of it comes from productivity improvements by incumbents. In particular, we can use equation (26), which decomposes growth into the component coming from incumbent firms (the first term) and that coming from new entrants (the second term).

Unfortunately, some of the parameters of the current model are difficult to pin down with our current knowledge of the technology of R&D. Hence, instead of a careful calibration exercise, here we provide some illustrative numbers. Let us normalize population to $L = 1$ and choose the following standard numbers:

$$g^* = 0.02,\ \rho = 0.01,\ r^* = 0.05 \text{ and } \theta = 2,$$

where $\theta$, the intertemporal elasticity of substitution, is pinned down by the choice of the other three numbers. The first three numbers refer to annual rates (implicitly defining $\Delta t = 1$ as one year). The remaining variables will be chosen so as to ensure that the equilibrium growth rate is indeed $g^* = 0.02$. As a benchmark, let us take

$$\beta = 2/3,$$

which implies that two thirds of national income accrues to labor and one third to profits. The requirement in (7) then implies that $\kappa > 1.7$. We will use the benchmark value of $\kappa = 3$ so that entry by new firms is sufficiently "radical" as suggested by some of the qualitative accounts of the innovation process (e.g., Freeman 1982, Scherer 1984). Innovation by incumbents is taken to be correspondingly smaller, in particular $\lambda = 1.5$, so that productivity gains from a radical innovation is about four times that of a standard "incremental" innovation by incumbents (i.e., $(\kappa - 1) / (\lambda - 1) = 4$). For the functions $\phi(z)$ and $\eta(z)$, we adopt the functional form in (3) and choose the benchmark values of $\alpha = 0.95$ and $\gamma = 0.5$. The remaining two parameters $A$ and $B$ will be chosen to ensure $g^* = 0.02$ with two third coming from the innovation of the incumbents and one third coming from the entrants, i.e., the firm term in (22) $\phi(z^*)(\lambda - 1)$ equals 0.0133 and the second term $\widehat{z}^* \eta(\widehat{z}^*)(\kappa - 1)$ equals 0.0067. Given the value of $\kappa$, we obtain $\widehat{z}^* \eta(\widehat{z}^*)$ equals 0.0033. However, varying these parameters shows that the model can lead to quite different decompositions of productivity growth between incumbents and entrants, and a more careful empirical investigation of the fit of the model is necessary (though the parameters that would be required for this need to be estimated).

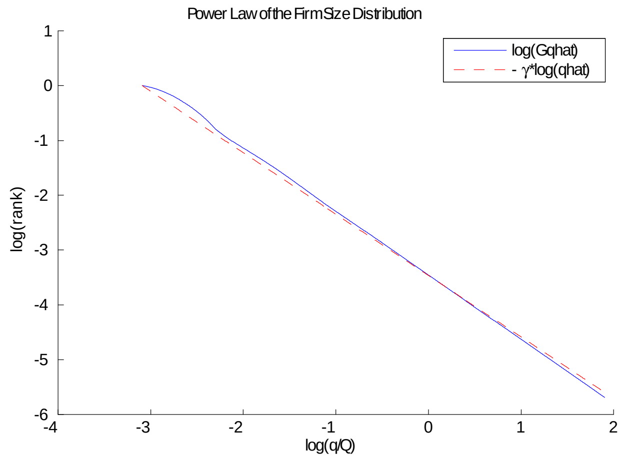
## 5.2 Stationary Distribution in the Economy with Imitation

We now present some simulations to illustrate the form of the stationary distribution of firm sizes in the economy with imitation. We will see that the firm size distribution is indeed well approximated by a Pareto distribution with a coefficient close to 1.

In addition to the parameters from the last subsection, let us set $\mu_e = 15$ and $\omega = 0.1$. This leads to a BGP growth rate of $g(\omega) = 0.0205 > g^* = 0.02$. The threshold for imitation is $\epsilon(\omega) = 0.045 < \omega(1 - \beta)^{\frac{1-\beta}{\beta}}$, which ensures that, following entry, imitators can charge the monopoly price.

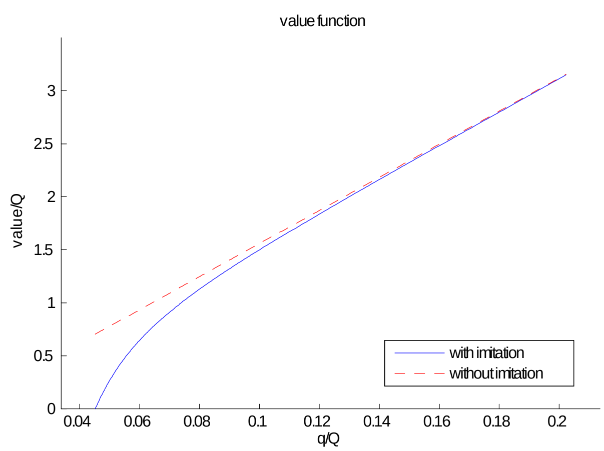
We can then compute that the firm size distribution will have a Pareto tail coefficient given by $\chi(\omega) = 1.12$. Figure 5.2 depicts the stationary firm size distribution by plotting the following relationship (similar to the one in Gabaix (1999)):

$$\log(rank) = C - \chi \log(size).$$




Stationary Distribution of Firm Size

Figure 5.2 then presents the value functions of the incumbents in the economy with imitation (solid line), and also for reference, it plots the value function in the baseline economy without imitation (dashed line). Under entry by imitation, the value of the incumbents is zero if $\widehat{q} = q/Q \leq \epsilon$. We can see that the value of incumbent firms without the imitation is everywhere above the value function in the economy with imitation. Though intuitive, this is not a general feature, because creative destruction may also decline in the presence of entry by imitation, and this may increase the value of incumbents.

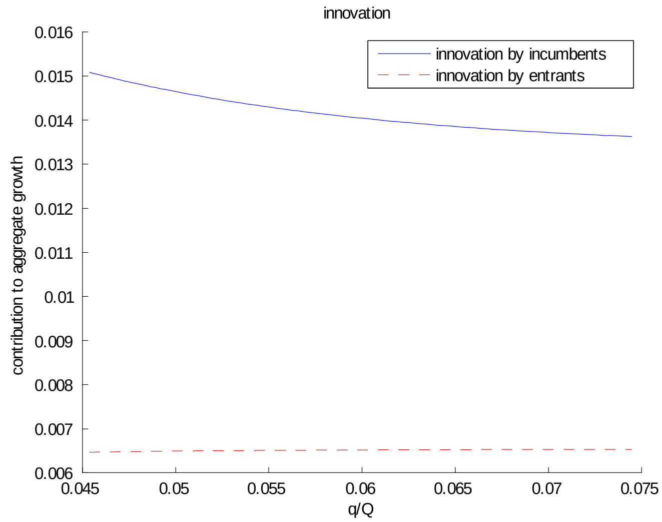



Value Function

Finally, Figure 5.2 presents the contributions of the incumbents and the entrants to the aggregate growth of product quality. About two-thirds of the aggregate growth is still due to incumbents. Notice that the incumbents with lower quality invest more because of the threat of entry by imitation.[26] This threat also makes the radical innovation (creative destruction)

[26] This is similar to the "escape competition" effect in Aghion et al. (2001), Aghion et al. (2005a) and Acemoglu and Akcigit (2006).

less profitable.




Innovations by Entrants and Incumbents.

# 6 Conclusion

A large fraction of US industry-level productivity growth is accounted for by existing firms and continuing establishments. Standard growth models either predict that most growth should be driven by new innovations brought about by entrants (and creative destruction) or do not provide a framework for decomposing the contribution of incumbents and entrants to productivity growth. In this paper, we proposed a simple modification of the basic Schumpeterian endogenous growth models that can address these questions. The main departure from the standard models is that incumbents have access to a technology for incremental innovations and can improve their existing machines (products). A different technology can then be used to generate more radical innovations. Arrow's replacement effect implies that only entrants will undertake R&D for radical innovations, while incumbents will invest in incremental innovations. This general pattern is in line with qualitative and quantitative evidence on the nature of innovation.

The model is not only consistent with the broad evidence but also provides a tractable framework for the analysis of productivity growth and of the entry of new firms and the expansion of existing firms. It yields a simple equation that decomposes productivity growth between continuing establishments and new entrants. Although the parameters to compute the exact contribution of different types of establishments to productivity growth have not yet been estimated, the use of plausible parameter values suggests that, in contrast to basic endogenous technological change models and consistent with the US data, a large fraction—but not all—of productivity growth is accounted by continuing establishments.

The comparative static results of this model are also potentially different from those of existing growth models, because innovation by incumbents also responds to changes in parameters and policy. For example, despite the presence of entry and creative destruction, the model shows that entry barriers or taxes on potential entrants may increase the equilibrium growth rate of the economy. This is because, in addition to their direct negative effects, such taxes create a positive impact on productivity growth by making innovation by incumbents more profitable.

Finally, because the model features entry by new firms and expansion and exit of existing firms, it also generates an equilibrium firm size distribution. The resulting stationary distribution of firm sizes approximates the Pareto distribution with an exponent of one (the so-called "Zipf distribution") observed in US data (e.g., Axtell 2001).

The model presented in this paper should be viewed as a first step in developing tractable models with endogenous productivity processes for incumbents and entrants (which take place via innovation and other productivity-increasing investments). It contributes to the literature on endogenous technological change by incorporating additional industrial organization elements in the study of economic growth. An important advantage of the approach developed here is that it generates predictions not only about the decomposition of productivity growth between incumbents and entrants, but also about the process of firm growth, entry and exit, and the equilibrium distribution of firm sizes. The resulting stochastic process for firm size is rather simple and does not incorporate rich firm dynamics that have been emphasized by other work, for example, by Klette and Kortum (2004), who allow firms to operate multiple products, or by Hopenhayn (1992), Melitz (2003) and Luttmer (2007), who introduce a nontrivial exit decision (due to the presence of fixed costs of operation) and also allow firms to learn about their productivity as they operate. Combining these rich aspects of firm entry and exit dynamics with innovation decisions that endogenize the stochastic processes of productivity growth of incumbents and entrants appears to be an important area for future theoretical research. A more important line of research, would be a more detailed empirical analysis of the predictions of these various approaches using data on productivity growth, exit and entry of firms. The relatively simple structure of the model presented in this paper should facilitate these types of empirical exercises. For example, a version of the current model, enriched with additional heterogeneity in firm growth, can be estimated using firm-level data on innovation (patents), sales, entry and exit.

# 7 Appendix

**Proof of Proposition 2. Part 1.** We need to show that $dg^*/d\tau_i < 0$, where the growth rate $g^*$ is given by (22). Differentiating this expression with respect to $\tau_i$, we obtain:

$$\begin{aligned} \frac{dg^*}{d\tau_i} &= (\lambda-1)\,\phi'\left(z_\tau\left(v^*\right)\right)\left(\frac{\partial z_\tau\left(v^*\right)}{\partial \tau_i}+\frac{\partial z_\tau\left(v^*\right)}{\partial v}\frac{dv^*}{d\tau_i}\right)+(\kappa-1)\,\frac{d}{dv}\left(\widehat{z}_\tau\left(v\right)\eta\left(\widehat{z}_\tau\left(v\right)\right)\right)\frac{dv^*}{d\tau_i} \\ &= (\lambda-1)\,\phi'\left(z_\tau\left(v^*\right)\right)\frac{\partial z_\tau\left(v^*\right)}{\partial \tau_i}+\left((\lambda-1)\,\phi'\left(z_\tau\left(v^*\right)\right)\frac{\partial z_\tau\left(v^*\right)}{\partial v}+(\kappa-1)\,\frac{\partial}{\partial v}\left(\widehat{z}_\tau\left(v\right)\eta\left(\widehat{z}_\tau\left(v\right)\right)\right)\right)\frac{dv^*}{d\tau_i} \end{aligned} \tag{46}$$

The first-order condition for the optimal investment decision of the incumbents, (33), implies that

$$\phi'\left(z_\tau\left(v\right)\right)(\lambda-1)\,v = 1+\tau_i. \tag{47}$$

Differentiating (47) with respect to $\tau_i$, we have $\phi''\left(z_\tau\left(v\right)\right)(\lambda-1)\,v\cdot\partial z_\tau\left(v^*\right)/\partial\tau_i = 1$, which implies that

$$\frac{\partial z_\tau\left(v^*\right)}{\partial \tau_i} = \frac{1}{\phi''\left(z_\tau\left(v\right)\right)(\lambda-1)\,v} < 0.$$

So a tax on the incumbents' investment will directly reduce their investment, which contributes to the reduction in the aggregate growth rate, i.e., the first term in (46) is negative. However, we need to ensure that the indirect effect on aggregate growth resulting from the equilibrium change in the value of the incumbents $v^*$, i.e., the second term in (46), does not offset the direct effect.

To study the second term in (46), we need to understand how the equilibrium value $v^*$ changes due to $\tau_i$, i.e., $dv^*/d\tau_i$. Notice that $v^*$ is determined in (35) as

$$\beta L = \Phi\left(\tau, v\right),$$

where we define

$$\Phi\left(\tau, v\right) \equiv \rho v + (\theta-1)\,\phi\left(z_\tau\left(v\right)\right)(\lambda-1)\,v + (1+\tau_i)z_\tau\left(v\right) + \left(\theta\left(\kappa-1\right)+1\right)\widehat{z}_\tau\left(v\right)\eta\left(\widehat{z}_\tau\left(v\right)\right)v. \tag{48}$$

Therefore, by the implicit function theorem

$$\frac{dv^*}{d\tau_i} = -\frac{\partial\Phi}{\partial\tau_i}\Big/\frac{\partial\Phi}{\partial v},$$

Plugging this identity in (46) we obtain

$$\begin{aligned} \frac{dg^*}{d\tau_i} &= (\lambda-1)\,\phi'\left(z_\tau\left(v^*\right)\right)\frac{\partial z_\tau\left(v^*\right)}{\partial \tau_i} \\ &\quad -\left((\lambda-1)\,\phi'\left(z_\tau\left(v^*\right)\right)\frac{\partial z_\tau\left(v^*\right)}{\partial v}+(\kappa-1)\,\frac{\partial}{\partial v}\left(\widehat{z}_\tau\left(v\right)\eta\left(\widehat{z}_\tau\left(v\right)\right)\right)\right)\frac{\partial\Phi}{\partial\tau_i}\Big/\frac{\partial\Phi}{\partial v}, \end{aligned}$$

where

$$\begin{aligned} \frac{\partial\Phi}{\partial v} &= \rho+(\theta-1)\,\phi\left(z_\tau\left(v\right)\right)(\lambda-1)+(\theta-1)\,\phi'\left(z_\tau\left(v\right)\right)v\frac{\partial z_\tau\left(v\right)}{\partial v}(\lambda-1)+(1+\tau_i)\frac{\partial z_\tau\left(v\right)}{\partial v} \\ &\quad +\left(\theta\left(\kappa-1\right)+1\right)\widehat{z}_\tau\left(v\right)\eta\left(\widehat{z}_\tau\left(v\right)\right)+\left(\theta\left(\kappa-1\right)+1\right)\frac{\partial\left(\widehat{z}_\tau\left(v\right)\eta\left(\widehat{z}_\tau\left(v\right)\right)\right)}{\partial v}v > 0, \end{aligned} \tag{49}$$

and
$$\frac{\partial\Phi}{\partial\tau_i}=(\theta-1)\,\phi'(z_\tau(v))\,\frac{\partial z_\tau(v)}{\partial\tau_i}(\lambda-1)\,v+(1+\tau_i)\frac{\partial z_\tau(v)}{\partial\tau_i}+z_\tau(v)\,.$$
The desired result, $dg^*/d\tau_i<0$ is thus established if
$$(\lambda-1)\,\phi'(z_\tau(v^*))\,\frac{\partial z_\tau(v^*)}{\partial\tau_i}\frac{\partial\Phi}{\partial v}<\left((\lambda-1)\,\phi'(z_\tau(v^*))\,\frac{\partial z_\tau(v^*)}{\partial v}+(\kappa-1)\,\frac{\partial}{\partial v}\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right)\right)\frac{\partial\Phi}{\partial\tau_i}.$$
Since $\partial z_\tau(v^*)/\partial v>0$,$\partial\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right)/\partial v>0$, $\partial z_\tau(v)/\partial\tau_i<0$ and $z_\tau>0$, it is then sufficient to show that
$$\begin{aligned}(\lambda-1)\,\phi'(z_\tau(v^*))\,\frac{\partial\Phi}{\partial v} &> \left((\lambda-1)\,\phi'(z_\tau(v^*))\,\frac{\partial z_\tau(v^*)}{\partial v}+(\kappa-1)\,\frac{\partial}{\partial v}\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right)\right)\\ &\quad\times\left((\theta-1)\,\phi'(z_\tau(v))\,(\lambda-1)\,v+(1+\tau_i)\right) \qquad (50)\end{aligned}$$
From (49), we have that
$$\frac{\partial\Phi}{\partial v}>(\theta-1)\,\phi'(z_\tau(v))\,v\frac{\partial z_\tau(v)}{\partial v}(\lambda-1)+(1+\tau_i)\frac{\partial z_\tau(v)}{\partial v}+(\theta(\kappa-1)+1)\,\frac{\partial\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right)}{\partial v}v. \qquad (51)$$
Combining (51) with (50), we see that this is equivalent to
$$(\theta(\kappa-1)+1)\,\frac{\partial\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right)}{\partial v}>(\kappa-1)\,\frac{\partial}{\partial v}\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right)+(\theta-1)\,(\kappa-1)\,\frac{\partial}{\partial v}\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right).$$
This is always true as $\theta(\kappa-1)+1>(\kappa-1)+(\theta-1)\,(\kappa-1)\,.$

**Part 2.** We need to show that $\frac{dg^*}{d\tau_e}>0$ when $\left|\phi''(z_\tau(v))\right|<\frac{(\kappa-1)(\lambda-1)(\beta L)^2}{\rho(1+\tau_i)^2}$. Once again differentiating the expression for $g^*$ in (22) with respect to $\tau_e$, we obtain
$$\begin{aligned}\frac{dg^*}{d\tau_e} &= (\lambda-1)\,\phi'(z_\tau(v^*))\,\frac{\partial z_\tau(v^*)}{\partial v}\frac{dv^*}{d\tau_e}\\ &\quad+(\kappa-1)\,\frac{\partial\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right)}{\partial v}\frac{dv^*}{d\tau_e}+(\kappa-1)\,\frac{\partial\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right)}{\partial\tau_e}\\ &= (\kappa-1)\,\frac{\partial\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right)}{\partial\tau_e}+\left((\lambda-1)\,\phi'(z_\tau(v^*))\,\frac{\partial z_\tau(v^*)}{\partial v}+(\kappa-1)\,\frac{\partial}{\partial v}\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right)\right)\frac{dv^*}{d\tau_e}.\end{aligned} \qquad (52)$$

Similarly to the first part, we decompose the change in the aggregate growth $g^*$, due to a tax $\tau_e$ on the entrants, into two components. The first component, which corresponds to the first term in (52), is the direct effect of the tax on the investment of the entrants. (34) implies that $\partial\left(\widehat{z}_\tau(v)\,\eta\left(\widehat{z}_\tau(v)\right)\right)/\partial\tau_e<0$, and thus the tax on entrants will reduce the entrants' investment. However, we will show that the indirect effect resulting from an increase in the incumbents' value $v^*$, i.e., the second term in (52), will more than offset the direct effect. To do so we also need to understand how the equilibrium value $v^*$ changes due to $\tau_e$, i.e., $dv^*/d\tau_e$.
Using the implicit function theorem again, we obtain
$$\frac{dv^*}{d\tau_e}=-\frac{\partial\Phi}{\partial\tau_e}/\frac{\partial\Phi}{\partial v},$$

where the expression for $\Phi$ in (48) implies that $\frac{\partial \Phi}{\partial \tau_e} \equiv (\theta(\kappa-1)+1) \frac{\partial(\widehat{z}_\tau(v)\eta(\widehat{z}_\tau(v)))}{\partial \tau_e} v$. Given that $\partial(\widehat{z}_\tau(v)\eta(\widehat{z}_\tau(v)))/\partial \tau_e < 0$ we have

$$
\begin{aligned}
\frac{dg^*}{d\tau_e} &\propto -(\kappa-1)\frac{\partial \Phi}{\partial v} + \\
&\quad (\theta(\kappa-1)+1) v^* \left( (\lambda-1)\phi'(z_\tau(v^*)) \frac{\partial z_\tau(v^*)}{\partial v} + (\kappa-1)\frac{\partial}{\partial v}(\widehat{z}_\tau(v^*)\eta(\widehat{z}_\tau(v^*))) \right) \\
&= -(\kappa-1)\frac{\partial \Phi}{\partial v} + \\
&\quad (\theta(\kappa-1)+1) \left( (1+\tau_i)\frac{\partial z_\tau(v^*)}{\partial v} + (\kappa-1)\frac{\partial}{\partial v}(\widehat{z}_\tau(v^*)\eta(\widehat{z}_\tau(v^*)))v^* \right) \qquad (53)
\end{aligned}
$$

in which the second equality follows from (47). We also have

$$
\begin{aligned}
v\frac{\partial \Phi}{\partial v} &= \Phi + (\theta-1)\phi'(z_\tau(v))v^2 \frac{\partial z_\tau(v)}{\partial v}(\lambda-1) + (1+\tau_i)v\frac{\partial z_\tau(v)}{\partial v} - (1+\tau_i)z_\tau(v) \\
&\quad + (\theta(\kappa-1)+1)\frac{\partial(\widehat{z}_\tau(v)\eta(\widehat{z}_\tau(v)))}{\partial v}v^2 \\
&= \Phi + \theta(1+\tau_i)v\frac{\partial z_\tau(v)}{\partial v} - (1+\tau_i)z_\tau(v) + (\theta(\kappa-1)+1)\frac{\partial(\widehat{z}_\tau(v)\eta(\widehat{z}_\tau(v)))}{\partial v}v^2.
\end{aligned}
$$

The last equality also follows from (47). Differentiating (47) with respect to $v$, we have

$$
v\frac{\partial z_\tau(v)}{\partial v} = \frac{\phi'(z_\tau)}{-\phi''(z_\tau)} = \frac{1+\tau_i}{(\lambda-1)v(-\phi''(z_\tau(v)))}.
$$

This can now be combined with (53) to yield

$$
\frac{dg^*}{d\tau_e} \propto -(\kappa-1)\beta L + (1+\tau_i)\left( \frac{1+\tau_i}{(\lambda-1)v(-\phi''(z_\tau(v)))} + (\kappa-1)z_\tau(v) \right).
$$

Because from (35), $v^* < \beta L/\rho$, the right-hand side of this expression is greater than

$$
-(\kappa-1)\beta L + \frac{(1+\tau_i)^2}{(\lambda-1)\frac{\beta L}{\rho}(-\phi''(z_\tau(v)))}.
$$

This implies that $dg^*/d\tau_e > 0$ provided that

$$
-\phi''(z_\tau(v)) < \frac{(\kappa-1)(\lambda-1)(\beta L)^2}{\rho(1+\tau_i)^2},
$$

establishing the desired result. ■

## Proof of Proposition 3.

Following from the analysis in the text, the current-value Hamiltonian for the social planner is

$$
\widehat{H}(Q^S, z^S, \widehat{z}^S, \mu^S) = \frac{\left(\beta(1-\beta)^{-\frac{1}{\beta}}Q^S L - (z^S+\widehat{z}^S)Q^S\right)^{1-\theta} - 1}{1-\theta} + \mu^S\left((\lambda-1)\phi(z^S) + (\kappa-1)\widehat{z}^S\eta(\widehat{z}^S)\right)Q^S.
$$

The necessary conditions for a candidate interior solution are given by

$$
\begin{aligned}
\frac{\partial \widehat{H}}{\partial z^S} &= -Q^S\left(\beta(1-\beta)^{-\frac{1}{\beta}}Q^S L - (z^S+\widehat{z}^S)Q^S\right)^{-\theta} + \mu^S(\lambda-1)\phi'(z^S)Q^S \\
\frac{\partial \widehat{H}}{\partial \widehat{z}^S} &= -Q^S\left(\beta(1-\beta)^{-\frac{1}{\beta}}Q^S L - (z^S+\widehat{z}^S)Q^S\right)^{-\theta} + \mu^S(\kappa-1)\left(\eta(\widehat{z}^S) + \widehat{z}^S\eta'(\widehat{z}^S)\right)Q^S,
\end{aligned}
$$

and

$$\begin{aligned}
\rho\mu^S - \dot{\mu}^S &= \frac{\partial \widehat{H}}{\partial Q^S} \\
&= \left(\beta (1-\beta)^{-\frac{1}{\beta}} L - \left(z^S + \widehat{z}^S\right)\right) \left(\beta (1-\beta)^{-\frac{1}{\beta}} Q^S L - \left(z^S + \widehat{z}^S\right) Q^S\right)^{-\theta} \\
&\quad + \mu^S \left((\lambda - 1) \phi \left(z^S\right) + (\kappa - 1) \widehat{z}^S \eta \left(\widehat{z}^S\right)\right), \qquad (54)
\end{aligned}$$

as well as a relevant transversality condition. $\frac{\partial \widehat{H}}{\partial z^S} = 0$ implies

$$\mu^S (\lambda - 1) \phi' \left(z^S\right) = \left(\beta (1-\beta)^{-\frac{1}{\beta}} Q^S L - \left(z^S + \widehat{z}^S\right) Q^S\right)^{-\theta}. \qquad (55)$$

Differentiating both sides with respect to $t$, we obtain:

$$\begin{aligned}
\dot{\mu}^S (\lambda - 1) \phi' \left(z^S\right) &= \left(\beta (1-\beta)^{-\frac{1}{\beta}} L - \left(z^S + \widehat{z}^S\right)\right)^{-\theta} \theta \left(Q^S\right)^{-\theta - 1} \dot{Q}^S (t) \\
&= \left(\beta (1-\beta)^{-\frac{1}{\beta}} L - \left(z^S + \widehat{z}^S\right)\right)^{-\theta} \theta \left(Q^S\right)^{-\theta} \\
&\quad \times \left((\lambda - 1) \phi \left(z^S\right) + (\kappa - 1) \widehat{z}^S (t) \eta \left(\widehat{z}^S\right)\right). \qquad (56)
\end{aligned}$$

Plugging in $\dot{\mu}^S$ and $\mu^S$ from (56) and (55) into (54) and dividing both sides by $\left(\beta (1-\beta)^{-\frac{1}{\beta}} L - \left(z^S + \widehat{z}^S\right)\right)^{-\theta}$ we obtain

$$\begin{aligned}
&\theta \left((\lambda - 1) \phi \left(z^S\right) + (\kappa - 1) \widehat{z}^S \eta \left(\widehat{z}^S\right)\right) + \rho = \\
&\left(\beta (1-\beta)^{-\frac{1}{\beta}} L - \left(z^S + \widehat{z}^S\right)\right) (\lambda - 1) \phi' \left(z^S\right) + \left((\lambda - 1) \phi \left(z^S\right) + (\kappa - 1) \widehat{z}^S \eta \left(\widehat{z}^S\right)\right).
\end{aligned}$$

Moreover, from the two first-order conditions $\frac{\partial \widehat{H}}{\partial z^S} = 0$ and $\frac{\partial \widehat{H}}{\partial \widehat{z}^S} = 0$, we have (37)

$$(\lambda - 1) \phi' \left(z^S\right) = (\kappa - 1) \left(\eta \left(\widehat{z}^S\right) + \widehat{z}^S \eta' \left(\widehat{z}^S\right)\right).$$

These two equations determine $z^S$ and $\widehat{z}^S$, and thus the Pareto allocation. The two equations above also give an expression for the Pareto optimal growth rate

$$g^S = \frac{\left(\beta (1-\beta)^{-\frac{1}{\beta}} L - z^S\right) (\lambda - 1) \phi' \left(z^S\right) + (\lambda - 1) \phi \left(z^S\right) - (\kappa - 1) \left(\widehat{z}^S\right)^2 \eta \left(\widehat{z}^S\right) - \rho}{\theta}.$$

When $\phi(\cdot)$ is linear, this growth rate is equal to

$$\begin{aligned}
g^S &= \frac{\beta (1-\beta)^{-\frac{1}{\beta}} L (\lambda - 1) \phi - (\kappa - 1) \left(\widehat{z}^S\right)^2 \eta' \left(\widehat{z}^S\right) - \rho}{\theta} > \frac{\beta L (\lambda - 1) \phi - \rho}{\theta} \\
&> \frac{\beta L (\lambda - 1) \phi - \widehat{z}^* \eta \left(\widehat{z}^*\right) - \rho}{\theta} = g^*,
\end{aligned}$$

so that the growth rate in the socially planned allocation is greater than the BGP growth rate, $g^*$, as defined in (32). The remark in the text that the social planner uses machines more intensively (because she avoids the monopoly distortions) can be seen by the fact that the first term in $g^S$, $(1-\beta)^{-\frac{1}{\beta}} L (\lambda - 1) \phi$, is strictly greater than the first term in $g^*$ in (32), since $(1-\beta)^{-\frac{1}{\beta}} > 1$). Notice also that (37) for linear $\phi$ implies

$$\eta \left(\widehat{z}^S\right) > \frac{(\lambda - 1) \phi}{\kappa - 1} > \frac{(\lambda - 1) \phi}{\kappa} = \eta \left(\widehat{z}^*\right).$$

Therefore, $\widehat{z}^* > \widehat{z}^S$, i.e., entry is too high in the decentralized equilibrium compared to the Pareto optimal level of entry. These inequalities also hold when $\left|\phi''(\cdot)\right| < \epsilon$ for $\epsilon$ sufficiently small, completing the proof. ∎

**Proof of Proposition 4.** The net present discounted utility of the representative household is given by

$$W = \int_0^\infty e^{-\rho t} \frac{C(t)^{1-\theta} - 1}{1-\theta} dt \propto \frac{1}{1-\theta} \frac{C(0)}{\rho - g(1-\theta)}$$

Therefore:

$$\frac{dW}{d\tau_e} \propto \frac{1}{1-\theta} \frac{dC(0)}{d\tau_e} \left(\rho - g(1-\theta)\right) + \frac{dg}{d\tau_e} C(0).$$

We can choose $\epsilon_1$ from Proposition 2 such that when $\left|\phi''(z_\tau(v))\right| < \epsilon_1$, we have $dg/d\tau_e > 0$. Therefore, $dW/d\tau_e > 0$ provided that $\rho - g(1-\theta) < \epsilon_2$ for some $\epsilon_2 > 0$. Setting $\epsilon = \min\{\epsilon_1, \epsilon_2\}$ establishes the desired result. ∎

**Derivation of the functional qquation for the stationary distribution .** Consider the evolution of the highest quality machine in each sector. In the case of entry without imitation, we have

$$q(t+\Delta t) = \begin{cases} q(t) & \text{with probability } 1 - \phi(z(\widetilde{q}(t)))\Delta t - \widehat{z}(\widetilde{q}(t))\eta(\widehat{z}(\widetilde{q}(t)))\Delta t + o(\Delta t) \\ \lambda q(t) & \text{with probability } \phi(z(\widetilde{q}(t)))\Delta t + o(\Delta t) \\ \kappa q(t) & \text{with probability } \widehat{z}(\widetilde{q}(t))\eta(\widehat{z}(\widetilde{q}(t)))\Delta t + o(\Delta t), \end{cases}$$

where $\widetilde{q}(t) \equiv q(t)/Q(t)$ is the normalized quality, and we have used the fact that average quality $Q(t)$ grows at a constant rate $g$. Moreover, because of imitation when $q(t) \leq \epsilon Q(t)$, $q(t+)$ jumps to $\omega Q(t)$, $\omega > \epsilon$. Therefore the evolution of the normalized quality, $\widetilde{q}(t)$, can be expressed as

$$\widetilde{q}(t+\Delta t) = \begin{cases} q(t)(1-g\Delta t) + o(\Delta t) \\ \text{with probability } 1 - \phi(z(\widetilde{q}(t)))\Delta t - \widehat{z}(\widetilde{q}(t))\eta(\widehat{z}(\widetilde{q}(t)))\Delta t + o(\Delta t) \\ \lambda q(t)(1-g\Delta t) + o(\Delta t) \;\; \text{with probability } \phi(z(\widetilde{q}(t)))\Delta t + o(\Delta t) \\ \kappa q(t)(1-g\Delta t) + o(\Delta t) \;\; \text{with probability } \widehat{z}(\widetilde{q}(t))\eta(\widehat{z}(\widetilde{q}(t)))\Delta t + o(\Delta t), \end{cases}$$

and whenever $\widetilde{q}(t) \leq \epsilon$, it jumps immediately to $\widetilde{q}(t+) = \omega > \epsilon$. Denoting the stationary distribution of normalized quality by $F(y)$, we have that, for $y > \omega$,

$$F(y) = \Pr(\widetilde{q}(t+\Delta t) \leq y) = \mathbb{E}\left[\mathbf{1}_{\{\widetilde{q}(t+\Delta t)\leq y\}}\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\widetilde{q}(t+\Delta t)\leq y\}} | \widetilde{q}(t)\right]\right].$$

We rewrite the iterated expectation as

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[\begin{array}{c} \mathbf{1}_{\{\widetilde{q}(t)(1-g\Delta t)\leq y, 1-\phi(z(\widetilde{q}(t)))\Delta t - \widehat{z}(\widetilde{q}(t))\eta(\widehat{z}(\widetilde{q}(t)))\Delta t, \widetilde{q}(t)(1-g\Delta t)>\epsilon\}} \\ +\mathbf{1}_{\{\lambda\widetilde{q}(t)(1-g\Delta t)\leq y, \phi(z(\widetilde{q}(t)))\Delta t, \widetilde{q}(t)(1-g\Delta t)>\epsilon\}} \\ +\mathbf{1}_{\{\kappa\widetilde{q}(t)(1-g\Delta t)\leq y, \widehat{z}(\widetilde{q}(t))\eta(\widehat{z}(\widetilde{q}(t)))\Delta t, \widetilde{q}(t)(1-g\Delta t)>\epsilon\}} \\ +\mathbf{1}_{\{\widetilde{q}(t)(1-g\Delta t)\leq\epsilon\}} \end{array} \middle| \widetilde{q}(t)\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\begin{array}{c} \mathbf{1}_{\{\widetilde{q}(t)\leq y(1+g\Delta t), 1-\phi(z(\widetilde{q}(t)))\Delta t - \widehat{z}(\widetilde{q}(t))\eta(\widehat{z}(\widetilde{q}(t)))\Delta t, \widetilde{q}(t)>\epsilon(1+g\Delta t)\}} \\ +\mathbf{1}_{\left\{\widetilde{q}(t)\leq\frac{y}{\lambda}(1+g\Delta t), \phi(z(\widetilde{q}(t)))\Delta t, \widetilde{q}(t)>\epsilon(1+g\Delta t)\right\}} \\ +\mathbf{1}_{\left\{\widetilde{q}(t)\leq\frac{y}{\kappa}(1+g\Delta t), \widehat{z}(\widetilde{q})\eta(\widehat{z}(\widetilde{q}))\Delta t, \widetilde{q}(t)>\epsilon(1+g\Delta t)\right\}} \\ +\mathbf{1}_{\{\widetilde{q}(t)\leq\epsilon(1+g\Delta t)\}} \end{array} \middle| \widetilde{q}(t)\right]\right] \\
&= \mathbb{E}\left[\begin{array}{c} \left(1-\phi(z(\widetilde{q}(t)))\Delta t - \widehat{z}(\widetilde{q})\eta(\widehat{z}(\widetilde{q}))\Delta t\right)\mathbf{1}_{\{q(t)\leq y(1+g\Delta t), \widetilde{q}(t)>\epsilon(1+g\Delta t)\}} \\ +\phi(z(\widetilde{q}(t)))\Delta t \mathbf{1}_{\left\{\widetilde{q}(t)\leq\frac{y}{\lambda}(1+g\Delta t), \widetilde{q}(t)>\epsilon(1+g\Delta t)\right\}} \\ +\widehat{z}(\widetilde{q})\eta(\widehat{z}(\widetilde{q}))\Delta t \mathbf{1}_{\left\{\widetilde{q}(t)\leq\frac{y}{\kappa}(1+g\Delta t), \widetilde{q}(t)>\epsilon(1+g\Delta t)\right\}} \\ +\mathbf{1}_{\{\widetilde{q}(t)\leq\epsilon(1+g\Delta t)\}}. \end{array}\right]
\end{aligned} \tag{57}$$

Replacing the last expectations by integrals and using the fact that $F(\widetilde{q})=0$ for $\widetilde{q}\leq\epsilon$, we obtain

$$\begin{aligned}F(y)=&\int_{\epsilon(1+g\Delta t)}^{y(1+g\Delta t)}\left(1-\phi\left(z\left(\widetilde{q}\right)\right)\Delta t-\widehat{z}\left(\widetilde{q}\right)\eta\left(\widehat{z}\left(\widetilde{q}\right)\right)\Delta t\right)dF\left(\widetilde{q}\right)\\&+\int_{\epsilon(1+g\Delta t)}^{\frac{y}{\lambda}(1+g\Delta t)}\phi\left(z\left(\widetilde{q}\right)\right)\Delta t dF\left(\widetilde{q}\right)+\int_{\epsilon(1+g\Delta t)}^{\frac{y}{\kappa}(1+g\Delta t)}\widehat{z}\left(\widetilde{q}\right)\eta\left(\widehat{z}\left(\widetilde{q}\right)\right)\Delta t dF\left(\widetilde{q}\right)\\&+F\left(\epsilon\left(1+g\Delta t\right)\right).\end{aligned}\tag{58}$$

For $\Delta t$ small, the right-hand side can be written as

$$\begin{aligned}F(y)=&F(y)+F'(y)\,yg\Delta t\\&-\int_{\epsilon}^{y}\phi\left(z\left(\widetilde{q}\right)\right)dF\left(\widetilde{q}\right)\Delta t-\int_{\epsilon}^{y}\widehat{z}\left(\widetilde{q}\right)\eta\left(\widehat{z}\left(\widetilde{q}\right)\right)dF\left(\widetilde{q}\right)\Delta t\\&+\int_{\epsilon}^{\frac{y}{\lambda}}\phi\left(z\left(\widetilde{q}\right)\right)dF\left(\widetilde{q}\right)\Delta t+\int_{\epsilon}^{\frac{y}{\kappa}}\widehat{z}\left(\widetilde{q}\right)\eta\left(\widehat{z}\left(\widetilde{q}\right)\right)dF\left(\widetilde{q}\right)\Delta t.\end{aligned}$$

Now eliminating $F(y)$ from both sides of the last equation, dividing $\Delta t$ and taking the limit as $\Delta t\to 0$, we obtain

$$0=F'(y)\,yg-\int_{\frac{y}{\lambda}}^{y}\phi\left(z\left(\widetilde{q}\right)\right)dF\left(\widetilde{q}\right)-\int_{\frac{y}{\kappa}}^{y}\widehat{z}\left(\widetilde{q}\right)\eta\left(\widehat{z}\left(\widetilde{q}\right)\right)dF\left(\widetilde{q}\right)$$

as in (43). For $y<\omega$, we proceed exactly as above except that now the terms $\mathbf{1}_{\{\widetilde{q}(t)(1-g\Delta t)\leq\epsilon\}}$ do not appear in (57) due to the fact that all imitating firms will have normalized quality $\omega$ exceeding $y$. We thus obtain

$$0=F'(y)\,yg-F'(\epsilon)\,\epsilon g-\int_{\frac{y}{\lambda}}^{y}\phi\left(z\left(\widetilde{q}\right)\right)dF\left(\widetilde{q}\right)-\int_{\frac{y}{\kappa}}^{y}\widehat{z}\left(\widetilde{q}\right)\eta\left(\widehat{z}\left(\widetilde{q}\right)\right)dF\left(\widetilde{q}\right)$$

as in (44). ■

**Proof of Proposition 5.** Evaluating (44) for the special case in which $\epsilon=\omega=0$, $g=g^*$, $z(\widetilde{q})\equiv z^*$ and $\widehat{z}(\widetilde{q})\equiv\widehat{z}^*$, we obtain the functional equation determining the stationary distribution in the baseline economy:

$$0=F'(y)\,yg^*-\phi\left(z^*\right)\left(F(y)-F\left(\frac{y}{\lambda}\right)\right)-\widehat{z}^*\eta\left(\widehat{z}^*\right)\left(F(y)-F\left(\frac{y}{\kappa}\right)\right).\tag{59}$$

The only possible solution to this equation is $F(y)=1-\left(\frac{\Gamma}{y}\right)^{\chi}$. Substituting this into (59) gives

$$\phi\left(z^*\right)\left(\lambda^{\chi}-1\right)+\widehat{z}^*\eta\left(\widehat{z}^*\right)\left(\kappa^{\chi}-1\right)-g^*\chi=0.$$

Since $\frac{\lambda^{\chi}-1}{\chi}$ and $\frac{\kappa^{\chi}-1}{\chi}$ are strictly increasing in $\chi$ and by definition of $g^*$, we have equality at $\chi=1$. Therefore $\chi=1$ is the unique solution. In some ways, this result looks quite remarkable, since it generates a stationary firm size distribution given by a Pareto distribution with an exponent of one, i.e., $\Pr\left[\widetilde{x}\leq y\right]=1-\frac{\Gamma}{y}$ with $\Gamma>0$. But the Pareto distribution is only defined for all $y\geq\Gamma$, thus $\Gamma$ should be the minimum normalized firm size. However (38) shows that it is possible for the normalized size of a firm $\widetilde{x}$ to tend to 0. Therefore $\Gamma$ should be equal to 0, which implies that there does not exist a stationary firm size distribution. ■

**Derivation of the Growth Equation.** The growth of the average product quality $Q_t$ comes from three sources: innovation from the incumbent firms, from the innovative entrants and from the imitators. Recall the definition of $Q_t$

$$Q_t = \int_0^1 q(v,t)\,dv$$

where $q(v,t)$ is the highest quality in sector $v$. We suppose that the investment of the incumbents in each sector is $z(\widetilde{q})$ and of the entrants is $\widehat{z}(\widetilde{q})$, where $\widetilde{q}$ is the quality relative to the average quality that grows at the rate $g$ from time $t$ to $t+\Delta t$. Then we have

$$\begin{aligned}
Q_{t+\Delta t} &= \int_0^1 q(v,t+\Delta t)\,dv \\
&= \int_{0,q(v,t)\geq \epsilon Q(t)(1+g\Delta t)}^1 \begin{pmatrix} \phi\left(z\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t\lambda q(v,t) + \widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t\kappa q(v,t) \\ +\left(1-\phi\left(z\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t - \widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t\right) q(v,t) \end{pmatrix} dv \\
&+ \int_{0,\epsilon Q(t)<q(v,t)<\epsilon Q(t)(1+g\Delta t)}^1 \begin{pmatrix} \phi\left(z\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t\lambda q(v,t) + \widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t\kappa q(v,t) \\ +\left(1-\phi\left(z\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t - \widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t\right) \omega Q(t) \end{pmatrix} dv.
\end{aligned}$$

Expanding the right hand side around $\Delta t = 0$, we have

$$\begin{aligned}
Q_{t+\Delta t} &= \lambda Q_t \Delta t \int_0^1 \phi\left(z\left(\frac{q(v,t)}{Q_t}\right)\right)\frac{q(v,t)}{Q_t}dv + \kappa Q_t\Delta t\int_0^1 \widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\right)\frac{q(v,t)}{Q_t}dv \\
&+ \int_{0,q(v,t)\geq \epsilon Q(t)(1+g\Delta t)}^1 \left(1-\phi\left(z\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t - \widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t\right) q(v,t)\,dv \\
&+ \int_{0,\epsilon Q(t)<q(v,t)<\epsilon Q(t)(1+g\Delta t)}^1 \left(1-\phi\left(z\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t - \widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t\right)\omega Q(t)\,dv.
\end{aligned}$$

We can rearrange to decompose the growth of average quality into three different components: innovation from incumbents, from entrants, and from imitators:

$$\begin{aligned}
Q_{t+\Delta t} = Q(t) &+ \underbrace{(\lambda-1)Q_t\Delta t\int_0^1 \phi\left(z\left(\frac{q(v,t)}{Q_t}\right)\right)\frac{q(v,t)}{Q_t}dv}_{\text{Innovation from Incumbents}} \\
&+ \underbrace{(\kappa-1)Q_t\Delta t\int_0^1 \widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\right)\frac{q(v,t)}{Q_t}dv}_{\text{Innovation from Entrants}} \\
&+ \underbrace{\int_{0,\epsilon Q(t)<q(v,t)<\epsilon Q(t)(1+g\Delta t)}^1 \begin{pmatrix} 1-\phi\left(z\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t \\ -\widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q_t}\right)\right)\Delta t \end{pmatrix}(\omega Q(t) - q(v,t))\,dv.}_{\text{Innovation from Imitators}}
\end{aligned}$$

We rewrite this growth accounting in term of stationary distribution with cumulative distribution function $F(\cdot)$ over $\widehat{q} = \frac{q}{Q} > \epsilon$ and probability density function $f(\cdot)$

$$\begin{aligned}
Q(t+\Delta t) = Q(t) &+ (\lambda-1)Q(t)\int \phi(z(\widetilde{q}))\widetilde{q}dF(\widetilde{q})\Delta t \\
&+ (\kappa-1)Q(t)\int \widehat{z}(\widetilde{q})\eta(\widehat{z}(\widetilde{q}))\widetilde{q}dF(\widetilde{q})\Delta t \\
&+ F(\epsilon(1+g\Delta t))(\omega-\epsilon)Q(t).
\end{aligned}$$

So

$$g=(\lambda-1)\,\mathbb{E}_F\left[\phi\left(z\left(\widetilde{q}\right)\right)\widetilde{q}\right]+(\kappa-1)\,\mathbb{E}_F\left[\widehat{z}\left(\widetilde{q}\right)\eta\left(\widehat{z}\left(\widetilde{q}\right)\right)\widetilde{q}\right]+\epsilon g f\left(\epsilon\right)\left(\omega-\epsilon\right).$$

Equivalently

$$g=\frac{(\lambda-1)\,\mathbb{E}_F\left[\phi\left(z\left(\widetilde{q}\right)\right)\widetilde{q}\right]+(\kappa-1)\,\mathbb{E}_F\left[\widehat{z}\left(\widetilde{q}\right)\eta\left(\widehat{z}\left(\widetilde{q}\right)\right)\widetilde{q}\right]}{1-\epsilon f\left(\epsilon\right)\left(\omega-\epsilon\right)}$$

as in (45). When $\omega=0$ we have

$$g=(\lambda-1)\,\mathbb{E}_F\left[\phi\left(z\left(\widetilde{q}\right)\right)\widetilde{q}\right]+(\kappa-1)\,\mathbb{E}_F\left[\widehat{z}\left(\widetilde{q}\right)\eta\left(\widehat{z}\left(\widetilde{q}\right)\right)\widetilde{q}\right],$$

and when $z\left(\widetilde{q}\right)\equiv z^*$ and $\widehat{z}\left(\widetilde{q}\right)\equiv\widehat{z}^*$

$$g=(\lambda-1)\,\phi\left(z^*\right)+(\kappa-1)\,\widehat{z}^*\eta\left(\widehat{z}^*\right)$$

as in (26), given that $E\left[\widetilde{q}\right]=1$. ■

**Proof of Proposition 6**:

Let us first define $I_i\left(\widetilde{v}\right)\equiv\max_{z\geq 0}\phi\left(z\right)\widetilde{v}-z$ and $I_e\left(u\right)\equiv\frac{1}{\kappa u}\eta^{-1}\left(\frac{1}{\kappa u}\right)$. Intuitively, $I_i\left(\widetilde{v}\right)$, for $\widetilde{v}=\left(\lambda-1\right)v$, is the value (or proportional to the value) of incumbent firms from undertaking incremental innovation. $I_e\left(u\right)$ is the rate of entry by entrants with radical innovations. There are one-to-one mappings from the investment technologies $\phi$ and $\eta$ to the functions $I_i$ and $I_e$. Using these notations we can also define $v_g$ as a solution of the equation

$$v=\frac{\beta L+I_i\left(\left(\lambda-1\right)v\right)}{r+I_e\left(v\right)},\tag{60}$$

in which $r=\rho+\theta g$. The following lemma establishes some properties of $v_g$ around the equilibrium values $\left(v^*,g^*\right)$. We can easily see that $v_g q$ is the value function of an incumbent with product quality $q$, given the interest rate $r$, the entry behavior of entrants and without imitators, i.e., functional equation (20).

**Lemma 1** *Suppose Assumption 1b is satisfied. There exists $\Delta>0$ such that for each $g\in\left(g^*-\Delta,g^*+\Delta\right)$, there exists a unique $v_g\in\left(v^*-\Delta,v^*+\Delta\right)$ that satisfies equation (60). Moreover $v_g$ is strictly decreasing in $g$.*

**Proof.** We rewrite equation (60) as $\Phi\left(v,g\right)=0$ where

$$\Phi\left(v,g\right)=v\left(\rho+\theta g+I_e\left(v\right)\right)-\beta L-I_i\left(\left(\lambda-1\right)v\right).$$

As we show below, $\partial\Phi\left(v^*,g^*\right)/\partial v>0$, so the implicit function theorem guarantees the existence and uniqueness of $\left(g,v_g\right)$ in the neighborhood of $\left(g^*,v^*\right)$, establishing the first part of the lemma. The second part follows immediately given that $\partial\Phi\left(v^*,g^*\right)/\partial g=v^*\theta>0$ and also by the implicit function theorem

$$\frac{dv_g}{dg}=-\frac{\partial\Phi\left(v^*,g^*\right)/\partial g}{\partial\Phi\left(v^*,g^*\right)/\partial v}.$$

We use direct calculation to show $\partial\Phi\left(v^*,g^*\right)/\partial v>0$. Indeed, we have

$$\frac{\partial\Phi\left(v^*,g^*\right)}{\partial v}=\rho+\theta g^*+\frac{\partial(v^*I_e\left(v^*\right))}{\partial v}-I_i'\left(\left(\lambda-1\right)v^*\right)\left(\lambda-1\right).$$

First of all, by definition of $I_e$, $v^* I_e(v^*) = \frac{1}{\kappa}\eta^{-1}\left(\frac{1}{\kappa v}\right)$, strictly increasing in $v$, so $\frac{\partial(v^* I_e(v^*))}{\partial v} > 0$. Second of all, applying the envelope theorem to $I_i(\widetilde{v})$ implies $I_i'((\lambda-1)v) = \phi(z(v))$. From the definition of $g^*$ in (26) and Assumption 1b, we have

$$\begin{aligned}
\theta g^* &\geq g^* \\
&> \phi(z^*)(\lambda-1) \\
&= I_i'((\lambda-1)v^*)(\lambda-1).
\end{aligned}$$

These two inequalities $\frac{\partial(v^* I_e(v^*))}{\partial v} > 0$ and $\theta g^* > I_i'((\lambda-1)v^*)(\lambda-1)$ imply

$$\frac{\partial \Phi(v^*, g^*)}{\partial v} > \rho > 0.$$

■

We prove Proposition 6 in three steps sketched in the body of the paper:

**Step 1:** We state the existence of a value function $\widehat{V}_g(\widetilde{q})$ in the following lemma

**Lemma 2 (Existence of Value Function)** *Suppose the BGP equilibrium in the baseline economy is described by Proposition 1, in particular with $v^*$, $g^*$ and $r^*$ as given by (23), (26) and (27), and Assumption 1b and Assumption 2 are satisfied. Then there exist $0 < \underline{\mu} < \overline{\mu}$ and $\Delta > 0$ such that for any $\mu_e \in (\underline{\mu}, \overline{\mu})$, $g \in [g^*, g^* + \Delta]$ and $\omega > 0$, we can find $\epsilon_g \leq \omega(1-\beta)^{\frac{1-\beta}{\beta}}$ and a value function $\widehat{V}_g(\widetilde{q})$ that satisfies (40), (41) and (42).*

Below, we show that $\widehat{V}_g(\widetilde{q}) = \widetilde{q}U_g(\ln(\widetilde{q}) - \ln\epsilon_g)$, where $U_g$ is shown to exist using Schauder's fixed point theorem, satisfies these properties. The following lemma shows the existence of $U_g$.

**Lemma 3** *Suppose Assumptions 1b and 2 are satisfied. Let $\Delta \in (0, g^*)$ small enough to apply Lemma 1. Then for each $g \in [g^* - \Delta, g^* + \Delta]$, there is a solution $U_g \geq 0$ to the functional equation*

$$\begin{aligned}
& rU(p) + gU'(p) \\
&= \beta L + \max_{z\geq 0}\{\phi(z)(\lambda U(p+\ln\lambda) - U(p)) - z\} - \widehat{z}(p)\eta(\widehat{z}(p))U(p),
\end{aligned} \tag{61}$$

*where $r = \rho + \theta g$, $\eta(\widehat{z}(p))\kappa U(p+\ln\kappa) = 1$, and $U$ satisfies the boundary conditions*

$$U(0) = 0 \text{ and } \lim_{p\to\infty} U(p) = v_g, \tag{62}$$

*where $v_g$ is defined in Lemma 1. Moreover, $U_g$ is equicontinuous in $g$ over any finite interval.*

Notice that the conditions (40) and (42) on $\widehat{V}_g$ translate into the conditions (61) and the first part of (62) on $U$. In order to apply Schauder's fixed point theorem, we need to find a subset $\digamma$ of continuous functions $U : [g^* - \Delta, g^* + \Delta] \times \mathbb{R}^+ \to \mathbb{R}$ that satisfy the boundary conditions (62), and a continuous mapping $T$ that summarizes the functional equation (61). We need $T(\digamma)$ to be a compact subset of $\digamma$. $\digamma$ and $T$ are constructed in Definitions 5 and 6. Lemmas 4 and 5 show that $T(\digamma)$ is a compact subset of $\digamma$. Lemma 6 shows that the mapping $T$ is continuous. Together with the Schauder's fixed point theorem, these properties ensure the existence of $U$.

**Definition 4** *$C^0([g^* - \Delta, g^* + \Delta] \times \mathbb{R}^+, \mathbb{R})$ denotes the Banach space of continuous functions $U : [g^* - \Delta, g^* + \Delta] \times \mathbb{R}^+ \to \mathbb{R}$ and $U(g, 0) = 0$ for all $g \in [g^* - \Delta, g^* + \Delta]$ with the norm*

$$\|U\| = \sup_{g^*-\Delta \leq g \leq g^*+\Delta} \sup_{0 \leq p \leq \infty} |U(g,p)|.$$

**Definition 5** *Let $\mathcal{F}$ denote the subset of continuous functions $U \in C^0([g^* - \Delta, g^* + \Delta] \times \mathbb{R}^+, \mathbb{R})$ with $U(g, 0) = 0$ and*

$$v_g - v_g e^{-\theta p} \leq U(g,p) \leq v_g + v_g e^{-\theta p} \text{ for all } p \geq 0. \tag{63}$$

**Definition 6** *For each function $u = U(g, .) \in C^0(\mathbb{R}^+, \mathbb{R})$ consider the operator $T_g$*

$$T_g u \in C^0\left(\mathbb{R}^+, \mathbb{R}\right)$$

*satisfies the following ordinary differential equation*[27]

$$\begin{aligned} & g\left(T_g u\right)'(p) + \left(r_g + I_e\left(u\left(p + \ln\kappa\right)\right)\right)\left(T_g u\right)(p) \\ & = \beta L + I_i\left(\lambda u\left(p + \ln\lambda\right) - T_g u(p)\right). \end{aligned} \tag{64}$$

*with the initial condition $T_g u(0) = 0$. Notice that*

$$r_g = \rho + \theta g. \tag{65}$$

*Here $\widehat{z}(p)$ is defined such that $\eta\left(\widehat{z}(p)\right)\kappa u\left(p + \ln\kappa\right) = 1$. The operator $T$ is defined by*

$$TU(g,p) = T_g U(g,p).$$

**Lemma 4** *Suppose Assumptions 1b and 2 are satisfied, then $T(\mathcal{F}) \subset \mathcal{F}$.*

**Proof.** Let $\overline{k}_g(p) = v_g + v_g e^{-\theta p}$ and $\underline{k}_g(p) = v_g + v_g e^{-\theta p}$. By definition, for each $U \in \mathcal{F}$, we have

$$\underline{k}_g(p) \leq U(g,p) \leq \overline{k}_g(p).$$

Let $k_g(p) = T_g U(g,p)$ then, also by definition (64) implies that

$$\begin{aligned} g k_g'(p) &= \beta L - I_i\left(\lambda u\left(p + \ln\lambda\right) - k_g(p)\right) - \left(r_g + I_e\left(u\left(p + \ln\kappa\right)\right)\right)\left(T_g u\right)(p) \\ &\leq \beta L - I_i\left(\lambda \overline{k}_g\left(p + \ln\lambda\right) - k_g(p)\right) - \left(r_g + I_e\left(\underline{k}_g\left(p + \ln\kappa\right)\right)\right)\overline{k}_g(p). \end{aligned}$$

So if

$$g\overline{k}_g'(p) > \beta L - I_i\left(\lambda \overline{k}_g\left(p + \ln\lambda\right) - \overline{k}_g(p)\right) - \left(r_g + I_e\left(\underline{k}_g\left(p + \ln\kappa\right)\right)\right)\overline{k}_g(p), \tag{66}$$

then $k_g(p) < \overline{k}_g(p)$ for all $p > 0$ given that $k_g(0) = 0 < \overline{k}_g(0)$. Similarly, if

$$g\underline{k}_g'(p) < \beta L - I_i\left(\lambda \underline{k}_g\left(p + \ln\lambda\right) - \underline{k}_g(p)\right) - \left(r_g + I_e\left(\overline{k}_g\left(p + \ln\kappa\right)\right)\right)\underline{k}_g(p) \tag{67}$$

[27] Standard results from the theory of ordinary differential equations ensure the existence and uniqueness of $T_g u(p)$ if $u \in \mathcal{F}$ defined below.

then $k_g(p) > \underline{k}_g(p)$ for all $p > 0$ given that $k_g(0) = 0 = \underline{k}_g(0)$.
Below we will use Assumption 2 to show (66) and (67). Indeed, the two inequalities can be re-written as (for all $0 < x \leq v_g$):

$$-g\theta x > \beta L + I_i\left((\lambda-1)v_g + \left(\lambda^{1-\theta}-1\right)x\right) - \left(r_g + I_e\left(v_g - \kappa^{-\theta}x\right)\right)(v_g + x), \tag{68}$$

and

$$g\theta x < \beta L + I_i\left((\lambda-1)v_g - \left(\lambda^{1-\theta}-1\right)x\right) - \left(r_g + I_e\left(v_g + \kappa^{-\theta}x\right)\right)(v_g - x). \tag{69}$$

By definition of $v_g$ in (1), we have equalities at $x = 0$. It is sufficient to show that the derivative of the left hand side of (68) is strictly greater than the derivative of its right hand side. Or equivalently,

$$-g\theta > I_i'\left((\lambda-1)v_g + \left(\lambda^{1-\theta}-1\right)x\right)\left(\lambda^{1-\theta}-1\right) - r_g - I_e\left(v_g - \kappa^{-\theta}x\right) + I_e'\left(v_g - \kappa^{-\theta}x\right)\kappa^{-\theta}v_g.$$

Equation (65) now implies that $r_g > g\theta$ and, from Assumption 1b, $\theta \geq 1$, yields

$$I_i'\left((\lambda-1)v_g + \left(\lambda^{1-\theta}-1\right)x\right)\left(\lambda^{1-\theta}-1\right) \leq 0.$$

It remains to show that

$$I_e\left(v_g - \kappa^{-\theta}x\right)v_g \geq I_e'\left(v_g - \kappa^{-\theta}x\right)\kappa^{-\theta},$$

or

$$\left(v_g - \kappa^{-\theta}x\right) \geq \frac{1}{\min \varepsilon_{I_e}}\kappa^{-\theta}v_g.$$

or

$$\frac{\min \varepsilon_{I_e}}{\min \varepsilon_{I_e} + 1} \geq \kappa^{-\theta} \tag{70}$$

Similarly, it is sufficient to show that the derivative of the left hand side of (69) is strictly greater than the derivative of its right hand side. Or equivalently,

$$g\theta < I_i'\left((\lambda-1)v_g - \left(\lambda^{1-\theta}-1\right)x\right)\left(1-\lambda^{1-\theta}\right) + r_g + I_e\left(v_g + \kappa^{-\theta}x\right) - I_e'\left(v_g + \kappa^{-\theta}x\right)\kappa^{-\theta}v_g.$$

This is true if $\left(v_g + \kappa^{-\theta}x\right) \geq \frac{1}{\min \varepsilon_{I_e}}\kappa^{-\theta}v_g$, or equivalently, if

$$\min \varepsilon_{I_e} \geq \kappa^{-\theta}. \tag{71}$$

Since

$$\epsilon_{I_e} = \frac{1}{\epsilon_\eta} - 1,$$

Assumption 2 implies both (70) and (71).

■

**Lemma 5** *$T(\digamma)$ is a compact subset of $C^0([g^* - \Delta, g^* + \Delta] \times \mathbb{R}^+, \mathbb{R})$.*

**Proof.** Suppose $\{f_n\}_{n=1}^{\infty} \subset \digamma$, we will show that we can extract a subsequence from $\{Tf_n\}_{n=1}^{\infty}$ that converges to $f^* \in \digamma$. First, there exists a constant $K > 0$ such that $\|U\| \leq K$ for all $U \in \digamma$. So (for all $g$ and $p$):

$$\left|\frac{\partial}{\partial p}Tf_n\right| = \left|\frac{d}{dp}T_g f_n(g,p)\right| \leq \frac{\beta L + I_i((\lambda+1)K) + (\rho + \theta(g^*+\Delta) + I_e(K))K}{g^* - \Delta}$$

Second, $D_g(p) = \frac{\partial}{\partial g}(Tf_n(p))$ is the solution of

$$\begin{aligned} & gD_g'(p) + Tf_n(p) + \left(\frac{dr_g}{dg} + I_e(f_n(p+\ln\kappa))\right) Tf_n(p) \\ = & \left(I_i'(\lambda f_n(p+\ln\lambda) - Tf_n(p)) - (r_g + I_e(f_n(p+\ln\kappa)))\right) D_g(p) \end{aligned}$$

So $D_g(p)$ is uniformly bounded over $[g^* - \Delta, g^* + \Delta] \times [0, M]$ for any $M > 0$. Therefore, for each $M = 1, 2, ....,$ we have $\{Tf_n(g,p)\}_{n=1}^{\infty}$ is equicontinuous over

$$C^0([g^* - \Delta, g^* + \Delta] \times [0, M], \mathbb{R}).$$

We construct subsequences $\left(\{Tf_{M_k}\}_{k\geq 1}\right)_{M\geq 1}$ of $\{Tf_n\}_{n\geq 1}$ as follows:

- $M = 1$: Since $\{Tf_n\}_{n\geq 1}$ is equicontinuous over $[g^* - \Delta, g^* + \Delta] \times [0, M]$, there exists a subsequence $\{Tf_{1_k}\}_{k=1}^{\infty}$ that converges uniformly to $f_M^* \in C^0([g^* - \Delta, g^* + \Delta] \times [0, M], \mathbb{R})$ over $[g^* - \Delta, g^* + \Delta] \times [0, M]$.
- $M \Longrightarrow M+1$: Since $\{Tf_{M_k}\}_{k=1}^{\infty}$ is equicontinuous over $[g^* - \Delta, g^* + \Delta] \times [0, M+1]$, there exists a subsequence $\left\{Tf_{(M+1)_k}\right\}_{k=1}^{\infty}$ that converges uniformly to $f_{M+1}^*$ over $[g^* - \Delta, g^* + \Delta] \times [0, M+1]$. Because of the subsequence property: $f_{M+1}^*\big|_{[g^*-\Delta, g^*+\Delta]\times[0,M]} = f_M^*$.

Let $f^* : [g^* - \Delta, g^* + \Delta] \times \mathbb{R}^+ \to R$ be defined by $f^*|_{[g^*-\Delta,g^*+\Delta]\times[0,M]} = f_M^*$ for all $M \in \mathbb{Z}_+$. By definition of $f^*$ we have for each $p \geq 0$ and $g \in [g^* - \Delta, g^* + \Delta]$, $\lim_{M\to\infty} Tf_{M_M}(g,p) = f^*(g,p)$ so $f^* \in \digamma$.
We now show that the subsequence $\{Tf_{M_M}\}_{M=1}^{\infty}$ converges to $f^*$, i.e.,

$$\lim_{M\to\infty} \|Tf_{M_M} - f^*\|_{C^0([g^*-\Delta,g^*+\Delta]\times\mathbb{R}^+,\mathbb{R})} = 0.$$

Indeed, for any $\epsilon > 0$, given (63) in the definition of $\digamma$, there exists a $p_1 > 0$ such that $|Tf_{M_M}(g,p) - v_g| < \frac{\epsilon}{2}$ and $|f^*(g,p) - v_g| < \frac{\epsilon}{2}$ for all $p \geq p_1$. So for all $p \geq p_1$ and $g \in [g^* - \Delta, g^* + \Delta]$, we have $|Tf_{M_M}(g,p) - f^*(g,p)| < \epsilon$. Given $p_1$, there exists an $M_1$ such that $|Tf_{M_M}(g,p) - f^*(g,p)| < \epsilon$ for all $p_1 \geq p \geq 0$,$g \in [g^* - \Delta, g^* + \Delta]$ and $M \geq M_1$ . Therefore, for all $p \geq 0$, $g \in [g^* - \Delta, g^* + \Delta]$ and $M \geq M_1$, we have $|Tf_{M_M}(g,p) - f^*(g,p)| < \epsilon$. ■

**Lemma 6** *The mapping $T$ is continuous over $\digamma$.*

**Proof.** Suppose $f_n \to f$, by the Lebesgue dominated convergence theorem, we have $Tf_n$ converges pointwise toward $Tf$. We next prove that

$$\lim_{n\to\infty} \|Tf_n - Tf\|_{C^0([g^*-\Delta,g^*+\Delta]\times\mathbb{R}^+,\mathbb{R})} = 0.$$

First, notice that $\{Tf_n\}$ is a Cauchy sequence: Because, for any $\epsilon > 0$ $\{Tf_n\}$ it is a Cauchy sequence over any restricted interval $[0, p_1]$ so we can find $M$ such that

$$\|Tf_m - Tf_n\|_{C^0([g^*-\Delta,g^*+\Delta]\times[0,p_1],\mathbb{R})} < \epsilon \text{ for all } m, n \geq M$$

and by definition of $\digamma$ we can choose $p_1$ such that

$$\begin{aligned} \|Tf_m(p,g) - Tf_n(p,g)\| &< \|Tf_m(p,g) - v_g\| + \|Tf_n(p,g) - v_g\| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \text{ for all } p \geq p_1. \end{aligned}$$

Second, by the relative compactness of $T(\digamma)$, from any subsequence of $\{Tf_n\}$ there is subsequence $\{h_M\}$ of $\{Tf_n\}$ that converges to $h$ over $C^0([g^* - \Delta, g^* + \Delta] \times \mathbb{R}^+, \mathbb{R})$. Since $\{h_M\}$ also converges pointwise to $Tf$ we have $h = Tf$. Therefore

$$\lim_{M\to\infty} \|h_M - Tf\|_{C^0([g^*-\Delta,g^*+\Delta]\times\mathbb{R}^+,\mathbb{R})} = 0.$$

Thus

$$\lim_{n\to\infty} \|Tf_n - Tf\|_{C^0([g^*-\Delta,g^*+\Delta]\times\mathbb{R}^+,\mathbb{R})} = 0.$$

■

**Proof of Lemma 3.** Given Lemma 4, 5, 6 we can apply the Schauder's fixed point theorem to show that $T$ admits a fixed point $U$ in $\digamma$: $TU = U$. Or equivalently for each $g \in [g^* - \Delta, g^* + \Delta]$, $u(\cdot) = U(g, .)$ satisfies $u(0) = 0$ and (61). The limit at infinity in (62) follows directly from the definition of $\digamma$. Finally, equicontinuity is a consequence of the fact that $U(\cdot) \in C^0([g^* - \Delta, g^* + \Delta] \times \mathbb{R}^+, \mathbb{R})$. ■

Now, we find $\epsilon_g$ such that the value function $\widehat{V}_g(\widetilde{q}) = \widetilde{q}U_g(\ln(\widetilde{q}) - \ln \epsilon_g)$ satisfies (40), (41) and (42).

**Proof of Lemma 2 (Existence of the Value Function).** Let us choose $\underline{\mu} < \overline{\mu}$ such that

$$\underline{\mu} > U_{g^*}\left(\frac{1-\beta}{\beta}\log\left(\frac{1}{1-\beta}\right)\right)$$

and

$$\overline{\mu} < U_{g^*}\left(\frac{1-\beta}{\beta}\log\left(\frac{1}{1-\beta}\right) + \delta\right),$$

where $\delta > 0$. Given that $U(g,p)$ is equicontinuous in $g \in [g^*, g^* + \Delta]$, we can choose $\Delta$ sufficiently small such that we can apply Lemma 3 and moreover $U_g\left(\frac{1-\beta}{\beta}\log\left(\frac{1}{1-\beta}\right)\right) < \underline{\mu} < \overline{\mu} < U_g\left(\frac{1-\beta}{\beta}\log\left(\frac{1}{1-\beta}\right) + \delta\right)$ for all $g \in [g^*, g^* + \Delta]$. Therefore, for any $\mu_e \in (\underline{\mu}, \overline{\mu})$ there exists an $\omega_g \in \left(\frac{1-\beta}{\beta}\log\left(\frac{1}{1-\beta}\right), \frac{1-\beta}{\beta}\log\left(\frac{1}{1-\beta}\right) + \delta\right)$ such that $\mu_e = U_g(\omega_g)$. For each $\omega$, let

$$\epsilon_g = \omega/\exp(\omega_g) < \omega(1-\beta)^{\frac{1-\beta}{\beta}}. \tag{72}$$

and let $\widehat{V}_g(\widetilde{q}) = \widetilde{q}U_g(\ln(\widetilde{q}) - \ln \epsilon_g)$. Then $\widehat{V}_g$ satisfies (40), (41) and (42). ■

Given the existence of $U(g,p)$, for each $g \in [g^*, g^* + \Delta]$, we define

$$z_g(p) = \arg\max_{z\geq 0} (\lambda U_g(p+\lambda) - U_g(p))\,\phi(z) - z$$

and

$$\widehat{z}_g(p) = \eta^{-1}\left(\frac{U_g(p+\ln\kappa)}{\kappa}\right)$$

Since $\lim_{p\to\infty} U_g(p) = v_g$, we have

$$\lim_{p\to\infty} z_g(p) = z(v_g) \text{ and } \lim_{p\to\infty} \widehat{z}_g(p) = \widehat{z}(v_g).$$

Armed with the existence of the value function and the corresponding investment decisions, we are ready to prove the second step

**Step 2:** In Lemma 7, we show the existence of the stationary distribution under the form $F_g'(\widetilde{q}) = f_g(\widetilde{q}) = \frac{h_g(\ln\widetilde{q}-\ln\epsilon_g)}{\widetilde{q}}$. Moreover, in Lemma 10, we show that $f_g$ satisfies the asymptotic Pareto property in Proposition 7. And lastly, for the purpose of the last step in proving the existence of a stationary BGP, in Lemma 12 we show that the mean firm size goes to infinity as $g$ approaches $g^*$, i.e., $\lim_{g\downarrow g^*}\int_0^\infty \widetilde{q}dF_g(\widetilde{q}) = \infty$.

We look for a stationary distribution $F_g(y)$ that solves equations (43) and (44) with $z(\widetilde{q}) = z_g(\ln\widetilde{q}-\ln\epsilon_g)$ and $\widehat{z}(\widetilde{q}) = \widehat{z}_g(\ln\widetilde{q}-\ln\epsilon_g)$. Let $h_g(p) = \epsilon_g e^p F_g'(\epsilon_g e^p)$, the equations (43) and (44) become:

If $p > \ln\omega_g$

$$0 = h_g(p)\,g - \int_{p-\ln\lambda}^{p}\phi(z_g(\widetilde{p}))\,h_g(\widetilde{p})\,d\widetilde{p} - \int_{p-\ln\kappa}^{p}\widehat{z}_g(\widetilde{p})\,\eta(\widehat{z}_g(\widetilde{p}))\,h_g(\widetilde{p})\,d\widetilde{p}. \tag{73}$$

If $p \le \ln\omega_g$

$$0 = h_g(p)\,g - h_g(0)\,g - \int_{p-\ln\lambda}^{p}\phi(z_g(\widetilde{p}))\,h_g(\widetilde{p})\,d\widetilde{p} - \int_{p-\ln\kappa}^{p}\widehat{z}_g(\widetilde{p})\,\eta(\widehat{z}_g(\widetilde{p}))\,h_g(\widetilde{p})\,d\widetilde{p}. \tag{74}$$

We also have $h_g(p) = 0$ for all $p \le 0$. The conditions for $F_g$ to be a well-defined distribution is

$$\int h_g(p)\,dp = 1.$$

The following lemma shows the existence and uniqueness of the stationary distribution.

**Lemma 7** *Given the investment strategies $z_g(p), \widehat{z}_g(p)$, the stationary distribution $h_g(p)$ exists and is unique.*

**Proof.** Differentiate both side of the integral equations on $h_g$, we have

$$\begin{aligned} gh_g'(p) &= \phi(z_g(p))\,h_g(p) - \phi(z_g(p-\ln\lambda))\,h_g(p-\ln\lambda) \\ &\quad + \widehat{z}_g(p)\,\eta(\widehat{z}_g(p))\,h_g(p) - \widehat{z}_g(p-\ln\kappa)\,\eta(\widehat{z}_g(p-\ln\kappa))\,h_g(p-\ln\kappa). \end{aligned}$$

We rewrite this equation as

$$\begin{aligned} gh_g'(p) - (\phi(z_g(p)) + \widehat{z}_g(p)\,\eta(\widehat{z}_g(p)))\,h_g(p) &= -\phi(z_g(p-\ln\lambda))\,h_g(p-\ln\lambda) \\ &\quad - \widehat{z}_g(p-\ln\kappa)\,\eta(\widehat{z}_g(p-\ln\kappa))\,h_g(p-\ln\kappa). \end{aligned}$$

Using the variation of constant formula, this equation yields a unique equation for $0 \leq p < \omega_g$ given $h_g(0)$. For $p \geq \omega_g$ the equation also yields a unique solution, however the initial condition is now

$$h_g(\omega_g) = \frac{1}{g}\int_{\omega_g - \ln\lambda}^{\omega_g} \phi(z_g(\widetilde{p}))\, h_g(\widetilde{p})\, d\widetilde{p} + \frac{1}{g}\int_{\omega_g - \ln\kappa}^{\omega_g} \widehat{z}_g(\widetilde{p})\, \eta(\widehat{z}_g(\widetilde{p}))\, h_g(\widetilde{p})\, d\widetilde{p}.$$

Since the system is linear in the initial condition $h_g(0)$, therefore, there exists a unique $h_g(0)$ such that $\int_0^\infty h_g(p)\, dp = 1$. Notice that Lemma 10 below implies that $\int_0^\infty h_g(p)\, dp < \infty$. ■

Having established the existence and uniqueness of the stationary distribution $h_g$, we now characterize some of its properties. In particular, Lemma 3 suggests that the investment policies $z_g(p), \widehat{z}_g(p)$ are approximately constant as $p$ goes to infinity. So the evolution of firm (sector) size assembles Gibrat's law for large firms. As a result, the stationary distribution $h_g$ should have a tail distribution close to Pareto. The following lemmas prove that conjecture.

Let $z_g^* = z(v_g)$ and $\widehat{z}_g^* = \widehat{z}_g(v_g)$. Then for each $g > g^*$, define the $\chi(g)$ as the unique number $\chi$ satisfying

$$g = \phi(z_g^*)\frac{\lambda^\chi - 1}{\chi} + \widehat{z}_{g^*}^*\eta(\widehat{z}_g^*)\frac{\kappa^\chi - 1}{\chi},$$

because the right hand side is strictly increasing in $\chi$. We will show below that $\chi(g)$ is the Pareto index of the Pareto tail of the stationary distribution $h_g$.

**Lemma 8** *$\chi(g^*) = 1$ and $\chi(g) > 1$ for all $g > g^*$ and in the neighborhood of $g^*$.*

**Proof.** By definition of $g^*$ we have $g^* = \phi(z_{g^*}^*)(\lambda - 1) + \widehat{z}^*\eta(\widehat{z}_{g^*}^*)(\kappa - 1)$, therefore $\chi(g^*) = 1$. For $g > g^*$

$$g > \phi(z_g^*)(\lambda - 1) + \widehat{z}_g^*\eta(\widehat{z}_g^*)(\kappa - 1).$$

Thus $\chi(g) > 1$. To show the previous inequality, notice that the left hand side is strictly increasing in $g$ and the right hand side is strictly increasing in $v_g$. However, Lemma 1 shows that $v_g$ is strictly decreasing in $g$, so the right hand side is strictly decreasing in $g$. Combining this fact with the fact that at $g = g^*$ the two sides are equal, we obtain the desired inequality. ■

**Lemma 9** *For each $\xi > 0$, there exists a $\delta > 0$ such that*

$$\left(\frac{1}{g}\phi(z_g^*) + \delta\right)\frac{\lambda^{\chi-\xi} - 1}{\chi - \xi} + \left(\frac{1}{g}\widehat{z}_g^*\eta(\widehat{z}_g^*) + \delta\right)\frac{\kappa^{\chi-\xi} - 1}{\chi - \xi} < 1$$

*and*

$$\left(\frac{1}{g}\phi(z_g^*) - \delta\right)\frac{\lambda^{\chi+\xi} - 1}{\chi + \xi} + \left(\frac{1}{g}\widehat{z}_g^*\eta(\widehat{z}_g^*) - \delta\right)\frac{\kappa^{\chi+\xi} - 1}{\chi + \xi} > 1.$$

**Proof.** This is true given $\frac{1}{g}\phi(z_g^*)\frac{\lambda^\chi - 1}{\chi} + \frac{1}{g}\widehat{z}_g^*\eta(\widehat{z}_g^*)\frac{\kappa^\chi - 1}{\chi} = 1$ and the functions

$$\frac{\lambda^{\chi-\xi} - 1}{\chi - \xi}, \frac{\kappa^{\chi-\xi} - 1}{\chi - \xi}$$

are strictly increasing in $\chi - \xi$. $\blacksquare$

For each $\xi > 0$ let $\delta > 0$ be such a $\delta$. Given the limit result in Lemma 3, there exists a $p_0 = p_0(\delta) \geq \omega_g$ such that, for all $p \geq p_0$

$$\left| \left( \frac{1}{g}\phi\left(z_g(p)\right) + \frac{1}{g}\widehat{z}_g(p)\,\eta\left(\widehat{z}_g(p)\right)\right) - \left(\frac{1}{g}\phi\left(z_g^*\right) + \frac{1}{g}\widehat{z}_g^*\eta\left(\widehat{z}_g^*\right)\right) \right| < \delta$$

and

$$\left| \frac{1}{g}\phi\left(z_g(p - \ln\lambda)\right) - \frac{1}{g}\phi\left(z_g^*\right) \right| < \delta$$
$$\left| \frac{1}{g}\widehat{z}_g(p - \ln\kappa)\,\eta\left(\widehat{z}_g(p - \ln\kappa)\right) - \frac{1}{g}\widehat{z}_g^*\eta\left(\widehat{z}_g^*\right) \right| < \delta.$$

We will now state and prove a key lemma. Proposition 7 then follows as a corollary of this lemma.

**Lemma 10 (Tail Index)** *For any $\xi > 0$, there exist $\overline{B}$, $\underline{B}$ and $p_0$ such that*

$$h_g(p) < 2\overline{B}e^{-(\chi(g)-\xi)p}, \text{ for all } p \geq p_0$$

*and*

$$h_g(p) > \frac{1}{2}\underline{B}e^{-(\chi(g)+\xi)p}, \text{ for all } p \geq p_0,$$

*In other words, $h_g(p) = e^{-\chi(g)p}\varphi_g(p)$, where $\varphi_g(p)$ is a slow-varying function.*

**Proof of the Tail Index Lemma.** Let us define $\overline{B}(\delta) \equiv \max_{p_0 \leq p \leq p_0 + \ln\kappa} h_g(p)\, e^{(\chi-\xi)p}$ and $\underline{B}(\delta) \equiv \min_{p_0 \leq p \leq p_0 + \ln\kappa} h_g(p)\, e^{(\chi+\xi)p}$. We will show that

$$h_g(p) < 2\overline{B}(\delta)\, e^{-(\chi-\xi)p}, \text{ for all } p \geq p_0$$

and

$$h_g(p) > \frac{1}{2}\underline{B}(\delta)\, e^{-(\chi+\xi)p}, \text{ for all } p \geq p_0.$$

These inequalities hold for $p_0 \leq p \leq p_0 + \ln\kappa$ by definition. We will next show that they also hold for all $p \geq p_0$. To obtain a contradiction, suppose that there is $p > p_0 + \ln\kappa$ such that $h_g(p) \geq 2\overline{B}(\delta)\, e^{-(\chi-\xi)p}$. Consider the infimum of those $p$, then

$$h_g(p) = 2\overline{B}(\delta)\, e^{-(\chi-\xi)p}.$$

In the other hand, the equation determining $h_g$ implies

$$\begin{aligned}
h_g(p) &= \frac{1}{g}\int_{p-\ln\lambda}^{p} \phi\left(z_g(\widetilde{p})\right) h_g(\widetilde{p})\, d\widetilde{p} + \frac{1}{g}\int_{p-\ln\kappa}^{p} \widehat{z}_g(\widetilde{p})\,\eta\left(\widehat{z}_g(\widetilde{p})\right) h_g(\widetilde{p})\, d\widetilde{p} \\
&< \int_{p-\ln\lambda}^{p} \left(\frac{1}{g}\phi\left(z_g^*\right) + \delta\right) 2\overline{B}(\delta)\, e^{-(\chi-\xi)\widetilde{p}} d\widetilde{p} + \int_{p-\ln\kappa}^{p} \left(\frac{1}{g}\widehat{z}_g^*\eta\left(\widehat{z}_g^*\right) + \delta\right) 2\overline{B}(\delta)\, e^{-(\chi-\xi)\widetilde{p}} d\widetilde{p} \\
&= 2\overline{B}(\delta)\left(\frac{1}{g}\phi\left(z_g^*\right) + \delta\right)\frac{\lambda^{\chi-\xi} - 1}{\chi - \xi} e^{-(\chi-\xi)p} + 2\overline{B}(\delta)\left(\frac{1}{g}\widehat{z}_g^*\eta\left(\widehat{z}_g^*\right) + \delta\right)\frac{\kappa^{\chi-\xi} - 1}{\chi - \xi} e^{-(\chi-\xi)p} \\
&< 2\overline{B}(\delta)\, e^{-(\chi-\xi)p}.
\end{aligned}$$

This yields a contraction. Therefore

$$h_g(p) < 2\overline{B}(\delta)\, e^{-(\chi-\xi)p}, \text{ for all } p \geq p_0.$$

Similarly, we can show that

$$h_g(p) > \frac{1}{2}\underline{B}(\delta)\, e^{-(\chi+\xi)p}, \text{ for all } p \geq p_0.$$

$\blacksquare$

As a consequence, if $g > g^*$, then $\chi(g) > 1$, Lemma 10 for $\xi = \frac{\chi(g)-1}{2}$ implies

$$\int h_g(p)\, dp < C \int e^{-\frac{1+\chi(g)}{2}p} dp < \infty.$$

In order to proceed to step 3, we need to show that the mean of firm size converges to infinity as $g$ approaches $g^*$ from above, i.e., $\lim_{g \longrightarrow g^*} \int_0^\infty h_g(p)\, e^p dp = \infty$. This is intuitively true using Lemma 7, because $h_{g^*}(p) \propto e^{-p}$ so $\int_0^\infty h_{g^*}(p)\, e^p dp \propto \int_0^\infty 1 dp = \infty$. Unfortunately, this does not work formally because Lemma 7 only provides $e^{-(1-\xi)p}$ as a lower bound. So, the following lemma gives a better lower bound of $h_{g^*}$ in order to prove the limiting result.

**Lemma 11 (Tail Index at the Limit)** *There exists $\underline{B}$ and $p_0$ such that*

$$h_{g^*}(p) > \frac{1}{2}\underline{B}\frac{e^{-p}}{p}, \text{ for all } p \geq p_0.$$

**Proof.** Let us choose $\underline{B} > 0$ such that the inequality holds for $p_0 \leq p \leq p_0 + \ln \kappa$. We will show that they also hold for all $p \geq p_0$ using contradiction.
First, we choose $p_0$ such that, using Lemma 4 there exists a constant $C > 0$ that satisfies

$$\begin{aligned} \phi(z_{g^*}(\widetilde{p})) &> \phi\left(z^*_{g^*}\right) - g^* C e^{-\theta \widetilde{p}} \\ \widehat{z}_{g^*}(\widetilde{p})\, \eta(\widehat{z}_{g^*}(\widetilde{p})) &> \widehat{z}^*_{g^*}\eta\left(\widehat{z}^*_{g^*}\right) - g^* C e^{-\theta \widetilde{p}} \\ \forall \widetilde{p} &\geq p_0. \end{aligned}$$

Suppose that there is $p > p_0 + \ln \kappa$ such that

$$h_{g^*}(p) < \frac{1}{2}\underline{B}\frac{e^{-p}}{p}.$$

Consider the infimum of those $p$, then

$$h_{g^*}(p) = \frac{1}{2}\underline{B}\frac{e^{-p}}{p}.$$

In the other hand, the equation determining $h_{g^*}$ implies

$$\begin{aligned}
h_{g^*}(p) &= \frac{1}{g^*}\int_{p-\ln\lambda}^{p} \phi\left(z_{g^*}(\widetilde{p})\right) h_{g^*}(\widetilde{p})\, d\widetilde{p} + \frac{1}{g^*}\int_{p-\ln\kappa}^{p} \widehat{z}_{g^*}(\widetilde{p})\, \eta\left(\widehat{z}_{g^*}(\widetilde{p})\right) h_{g^*}(\widetilde{p})\, d\widetilde{p} \\
&> \int_{p-\ln\lambda}^{p} \left(\frac{1}{g^*}\phi\left(z^*_{g^*}\right) - Ce^{-\theta\widetilde{p}}\right)\frac{1}{2}\underline{B}\frac{e^{-\widetilde{p}}}{\widetilde{p}}d\widetilde{p} + \int_{p-\ln\kappa}^{p}\left(\frac{1}{g^*}\widehat{z}^*_{g^*}\eta\left(\widehat{z}^*_{g^*}\right) - Ce^{-\theta\widetilde{p}}\right)\frac{1}{2}\underline{B}\frac{e^{-\widetilde{p}}}{\widetilde{p}}d\widetilde{p} \\
&= \frac{1}{2}\underline{B}\frac{1}{g^*}\phi\left(z^*_{g^*}\right)(\lambda-1)\frac{e^{-p}}{p} + \frac{1}{2}\underline{B}\frac{1}{g^*}\widehat{z}^*_{g^*}\eta\left(\widehat{z}^*_{g^*}\right)(\kappa-1)\frac{e^{-p}}{p} \\
&+ \frac{1}{2}\underline{B}C'\frac{e^{-p}}{p^2} - \frac{1}{2}\underline{B}C''e^{-\left(1+\frac{\theta}{2}\right)p} \\
&> \frac{1}{2}\underline{B}\frac{e^{-p}}{p}.
\end{aligned}$$

(We also choose $p_0$ such that $C'\frac{e^{-p}}{p^2} - \frac{1}{2}C''e^{-\left(1+\frac{\theta}{2}\right)p} > 0$ for all $p \geq p_0$). This yields a contraction. Therefore

$$h_{g^*}(p) > \frac{1}{2}\underline{B}\frac{e^{-p}}{p},\ \forall\ p \geq p_0.$$

■

So, a direct consequence of this lemma is

$$\int_0^\infty h_{g^*}(p)\, e^p dp > \int_{p_0}^\infty \frac{1}{2}\underline{B}\frac{1}{p}dp = \infty.$$

We can use the results above to prove the following property of $h_g$, which will be crucial to show the existence of the equilibrium growth rate of the economy with imitation, i.e., the last step in proving Proposition 6.

**Lemma 12** *$h_g$ is uniformly continuous in $g$. And for $g > g^*$*

$$\Phi(g) = \int_0^\infty h_g(p)\, e^p dp < \infty$$

*is continuous in $g$. Moreover $\lim_{g\downarrow g^*}\Phi(g) = +\infty$.*

**Proof.** The fact that $h_g$ is uniformly continuous in $g$ is a result of uniform continuity of $\{U_g\}$, thus of $z_g(\cdot)$ and $\widehat{z}_g(\cdot)$ as well. $\Phi(g)$ is finite given Proposition 7. $\Phi(g)$ is continuous by the Lebesgue dominated convergence theorem. Finally, as we show above $\Phi(g^*) = +\infty$ and by the uniform continuity of $h_g$, we have $\lim_{g\downarrow g^*}\Phi(g) = +\infty$. ■

For any $\omega > 0$, $\epsilon_g$ is defined as in (72). The corresponding stationary distribution is $f_g(\widetilde{q}) = \frac{h_g(\ln\widetilde{q} - \ln\epsilon_g)}{\widetilde{q}}$ and policy functions are $z(\widetilde{q}) = z_g(\ln\widetilde{q} - \ln\epsilon_g)$ and $\widehat{z}(\widetilde{q}) = \widehat{z}_g(\ln\widetilde{q} - \ln\epsilon_g)$for all $\widetilde{q} \geq \epsilon_g$. Given $g$, $g'$ as defined in (45) can be written using $h_g$ and the change of variable $\widetilde{q} = \epsilon_g e^p$:

$$\begin{aligned}
g' &= \frac{(\lambda-1)\int_{\epsilon_g}^\infty \widetilde{q}\phi\left(z(\widetilde{q})\right) f_g(\widetilde{q})\, d\widetilde{q} + (\kappa-1)\int_{\epsilon_g}^\infty \widetilde{q}\widehat{z}(\widetilde{q})\,\eta\left(\widehat{z}(\widetilde{q})\right) f_g(\widetilde{q})\, d\widetilde{q}}{1-\epsilon_g f_g(\epsilon_g)(\omega-\epsilon_g)} \\
&= \epsilon_g\frac{(\lambda-1)\int_0^\infty e^p\phi\left(z_g(p)\right) h_g(p)\, dp + (\kappa-1)\int_0^\infty e^p\widehat{z}_g(p)\,\eta\left(\widehat{z}_g(p)\right) h_g(p)\, dp}{1-h_g(0)(\omega-\epsilon_g)}
\end{aligned}$$

We obtain an BGP if $g' = g$.

**Step 3:** We next show that there exists $g(\omega)$ such that

$$D(g(\omega)) \equiv g'(g(\omega)) - g(\omega) = 0.$$

Let $\underline{\mu}$, $\overline{\mu}$ and $\Delta$ be chosen as in Lemma 2 and such that Lemma 8 for $g \in [g^*, g^* + \Delta]$. Consider $\overline{\omega} > 0$ sufficiently small such that for all $0 < \omega < \overline{\omega}$, $\epsilon = \frac{\omega}{\omega_{g^*+\Delta}}$ satisfies

$$\begin{aligned} &\epsilon \frac{(\lambda - 1)\int_0^\infty e^p \phi\left(z_{g^*+\Delta}(p)\right) h_{g^*+\Delta}(p)\, dp + (\kappa - 1)\int_0^\infty e^p \widehat{z}_{g^*+\Delta}(p)\, \eta\left(\widehat{z}_{g^*+\Delta}(p)\right) h_{g^*+\Delta}(p)\, dp}{1 - h_{g^*+\Delta}(0)(\omega - \epsilon)} \\ &< g^* + \Delta. \end{aligned} \tag{75}$$

We will show that, for each $\omega$ such that $0 < \omega < \overline{\omega}$, there exists a $g = g(\omega) \in (g^*, g^* + \Delta)$ such that $g' = g$ and that

$$\lim_{\omega \to 0} g(\omega) = g^*.$$

Indeed, as in the proof of Lemma 2 there exists a $\omega_g$ such that $\omega_g \in \left(\frac{1-\beta}{\beta} \log\left(\frac{1}{1-\beta}\right), \frac{1-\beta}{\beta} \log\left(\frac{1}{1-\beta}\right) + \delta\right)$ and $U_g(\omega_g) = \mu_e$. Set $\epsilon_g = \omega / \exp(\omega_g)$ and define

$$D(g) = \epsilon_g \frac{(\lambda - 1)\int_0^\infty e^p \phi(z_g(p))\, h_g(p)\, dp + (\kappa - 1)\int_0^\infty e^p \widehat{z}_g(p)\, \eta(\widehat{z}_g(p))\, h_g(p)\, dp}{1 - h_g(0)(\omega - \epsilon_g)} - g.$$

Using Lemma 12, we can show that $D(g)$ is continuous in $g$. Moreover, by (75), we have

$$D(g^* + \Delta) < 0,$$

and, Lemma 12 implies

$$\lim_{g \to g^*} D(g) = +\infty.$$

Therefore, by the intermediate value theorem, there exists a $g(\omega)$ such that $D(g) = 0$. Moreover if $g(\omega) > g^* + \varpi$ as $\omega \to 0$, we also have

$$\epsilon_g \frac{(\lambda - 1)\int_0^\infty e^p \phi(z_g(p))\, h_g(p)\, dp + (\kappa - 1)\int_0^\infty e^p \widehat{z}_g(p)\, \eta(\widehat{z}_g(p))\, h_g(p)\, dp}{1 - h_g(0)(\omega - \epsilon_g)} \to 0$$

(because $\epsilon_g \to 0$). This implies $D(g(\omega)) < -(g^* + \varpi) < 0$, yielding a contradiction with the fact that $D(g(\omega)) = 0$. Thus

$$\lim_{\omega \to 0} g(\omega) = g^*.$$

The uniform convergences of the value and policy functions are obtained immediately given the bounds on the value function in (63), which themselves converge uniformly. ■

# References

William Abernathy. *The Productivity Dilemma.* Johns Hopkins University Press, Baltimore, 1980.

Daron Acemoglu and Ufuk Akcigit. State dependent intellectual property right policy. NBER Working Paper, No. 12775, 2006.

Philippe Aghion and Peter Howitt. A model of growth through creative destruction. *Econometrica*, 60(2):323–351, 1992. ISSN 00129682.

Philippe Aghion, Christopher Harris, Peter Howitt, and John Vickers. Competition, imitation and growth with step-by-step innovation. *Reviews of Economic Studies*, 68:467–492, 2001.

Philippe Aghion, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. Competition and innovation: An inverted-u relationship. *Quarterly Journal of Economics*, 120(2):701–728, 2005a. doi: 10.1162/0033553053970214. URL `http://www.mitpressjournals.org/doi/abs/10.1162/0033553053970214`.

Philippe Aghion, Robin Burgess, Stephen Redding, and Fabrizio Zilibotti. Entry liberalization and inqequality in industrial performance. *Journal of the European Economic Association Papers and Proceed*, 3:291–302, 2005b.

Ufuk Akcigit. Firm size, innovation dynamics and growth. University of Pennsylvania, working paper, 2010.

Ufuk Akcigit and William R. Kerr. Growth through heterogeneous innovations. University of Pennsylvania, working paper, 2010.

Kenneth Arrow. *The Limits of Organization.* Norton, New York, 1974.

Kenneth J. Arrow. The economic implications of learning by doing. *Review of Economic Studies*, 29:155–173, 1962.

Andrew Atkeson and Ariel Burstein. Innovation, firm dynamics and international trade. 2010.

Andrew Atkeson and Robert Lucas. On efficient distribution with private information. *Review of Economic Studies*, 59:427–53, 1992.

R.L. Axtell. Zipf distribution of us firm sizes. *Science*, 293:1818–1820, 2001.

Robert Barro and Xavier Sala-i Martin. *Economic Growth*. MIT Press, Cambridge, MA, 2004.

Eric J Bartelsman and Mark Doms. Understanding productivity: Lessons from longitudinal microdata. *Journal of Economic Literature*, 38:569–594, 2000.

Jesse Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth and fiscal policy in economies with finitely lived agents. NYU working paper, 2010.

Christian Broda and David E. Weinstein. Globalization and the gains from variety. *Quarterly Journal of Economics*, CXXI(2):541–585, 1996.

Juan Cordoba. On the distribution of city sizes. *Journal of Urban Economics*, 2008.

L. Dobkins and Yannis Ioannides. Dynamic evolution of the u. s. city size distribution. In J. M. Huriot and J. F. Thisse, editors, *The Economics of Cities*. Cambridge University Press, New York, 1998.

Mark J. Roberts Dunne, Timothy and Larry Samuelson. Patterns of firm entry and exit in us manufacturing industries. *Rand Journal of Economics*, 19:495–515, 1988.

Mark J. Roberts Dunne, Timothy and Larry Samuelson. The growth and failure of us manufacturing plants. *Quarterly Journal of Economics*, 104:671–698, 1989.

Jan Eeckhout and Boyan Jovanovic. Knowledge spillovers and inequality. *American Economic Review*, 92:1290–1307, 2002.

Richard Ericson and Ariel Pakes. Markov perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62:53–82, 1995.

Federico Etro. Innovation by leaders. *Economic Journal*, 114:281–303, April 2004.

John Haltiwanger Foster, Lucia and C.J. Krizan. Aggregate procuctivity growth: Lessons from microeconomic evidence. NBER Working Paper No.6803, 2000.

Christopher Freeman. *The Economics of Industrial Innovation*. MIT Press Cambridge, Massachusetts, 1982.

Xavier Gabaix. Zipf's law for cities: An explanation. *Quarterly Journal of Economics*, 114: 739–767, 1999.

Gene Grossman and Elhanan Helpman. Quality ladders in the theory of growth. *Review of Economic Studies*, 68:43–61, 1991a.

Gene Grossman and Elhanan Helpman. *Innovation and Growth in the Global Economy.* Cambridge, MA, MIT Press, 1991b.

Bronwyn Hall. The relationship between firm size and firm growth in the us manufacturing sector. *The journal of industrial economics*, 35, 1987.

Hugo A. Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 65(5):1127–1150, 1992.

Tadeuz Jankowski. Advanced differential equations with nonlinear boundary condition. *Journal of Mathematical Analysis and Applications*, 304:490–503, 2005.

Charles I. Jones. R and d-based models of economic growth. *Journal of Political Economy*, 103:759–784, 1995.

Boyan Jovanovic. Selection and evolution of industry. *Econometrica*, 50:649–670, 1982.

Steven Klepper. Entry, exit, growth, and innovation over the product life cycle. *The American Economic Review*, 86(3):562–583, 1996. ISSN 00028282.

Tor Jakob Klette and Samuel Kortum. Innovating firms and aggregate innovation. *Journal of Political Economy*, 112(5):p986 – 1018, 2004. ISSN 00223808.

Ricardo Lagos. A model of tfp. NYU Working Paper, 2001.

Rasmus Lentz and Dale T. Mortensen. On the size distribution of business firms. *Econometrica*, 76(6):1317–1373, 2008.

M. B. Lieberman. The learning curve and pricing curve and chemical processing industries. *Rand Journal of Economics*, 15:213–228, 1984.

Robert. E. Jr. Lucas. On the size distribution of business firms. *Bell Journal of Economics*, 9(2):508–523, 1978.

Erzo Luttmer. The size distribution of firms in an economy with fixed entry costs. Federal Reserve Bank of Minneapolis, working paper, 2004.

Erzo Luttmer. Selection, growth and the size distribution of firms. *Quarterly Journal of Economics*, 122:1103–1144, 2007.

Erzo Luttmer. On the mechanics of firm growth. Federal Reserve Bank of Minneapolis, working paper, 2010a.

Erzo Luttmer. Technology diffusion and growth. Federal Reserve Bank of Minneapolis, working paper, 2010b.

Erzo Luttmer. Models of firm heterogeneity and growth. *Annual Review of Economics*, 2: 547–576, 2010c.

Mark Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725, 2003.

J. M. Pennings and A. Buitendam. *New Technology As Organizational Innovation: The Development and Diffusion of Microelectronics*. Bollinger, Cambridge, MA, 1987.

Paul Romer. Endogenous technical change. *Journal of Political Economy*, 98(5):71–102, 1990.

Esteban Rossi-Hansberg and Mark L.J. Wright. Urban structure and growth. *Review of Economic Studies*, 74:597–624, 2007a.

Esteban Rossi-Hansberg and Mark L.J. Wright. Establishment size dynamics in the aggregate economy. *American Economic Reviews*, 97(5):1639–1666, 2007b.

Frederick Scherer. *Innovation and Growth: Schumpeterian Perspectives*. MIT Press Cambridge, Massachusetts, 1984.

Joseph A. Schumpeter. *The Theory of Economic Development*. Harvard University Press, Cambridge, MA, 1934.

Joseph A. Schumpeter. *Capitalism, Socialism and Democracy*. Harper and Brothers, New York, 1942.

T.C. A. Anant Segerstrom, Paul S. and Elias Dinopoloulos. A schumpterian model of the product life cycle. *American Economic Reviews*, 80:1077–1091, 1990.

Herbert. Simon and C. P. Bonini. The size distribution of business firms. *American Economic Reviews*, 48:607–617, 1958.

V.A. Staikos and Tsamatos. On the terminal value problem for differential equations with deviating arguments. *Archivum Mathematicum*, 21(1):43–50, 1985.

John Sutton. Gibrat's legacy. *Journal of Economic Literature*, 35:40–59, 1997.

John Sutton. *Technology and Market Structure:Theory and History.* MIT Press, Cambridge, MA, 1998.

M. L. Tushman and P. Anderson. Technological discontinuities and organizational environments. *Administrative Science Quarterly*, 31:439–465, 1986.