Innovation by Entrants and Incumbents

Daron Acemoglu[*]

Department of Economics, MIT

77 Massachussets Avenue, Cambridge MA 02147, United States

Dan Cao

Department of Economics, Georgetown University

37th and O Streets, NW, Washington DC 20057, United States

December 15, 2014

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. The model generates a non-degenerate equilibrium firm size distribution driven by entry of new firms and expansion exit of existing firms. 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").

Keywords: economic growth, entry, Gibrat's Law, industry structure, innovation, Pareto distribution, productivity growth, Zipf's Law.
JEL Classification: O31, O34, O41, L16.

---

[*]Department of Economics, MIT. E-mail: daron@mit.edu. Telephone: +1-617-253-1927.

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 et al. (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 (1991c), Jones (1995), and the Schumpeterian quality ladder models, e.g., Segerstrom and Dinopoloulos (1990), Aghion and Howitt (1992), Grossman and Helpman (1991b).[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 bulk of productivity growth accounted for continuing firms and establishments.[3]

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 tractable (and minimal) extension of the textbook multisector Schumpeterian growth model.

---

[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 this important distinction. For models in which dynamics of multi-establishment firms can be studied to see Klette and Kortum (2004) and Acemoglu et al. (2013).

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.

Our model identifies a new effect: the rate of entry of new firms depresses the profitability of incumbent innovation and, via this channel, may reduce aggregate productivity growth. Whether this is the case or not depends on the degree of diminishing returns to incumbent innovation. When incumbent innovation has a linear technology, greater entry or policies that encourage entry reduce overall innovation and growth. But in the more plausible case where incumbent innovation exhibits diminishing returns, the relationship between entry and overall growth is ambiguous, and depends on whether the degree of diminishing returns is above or below a critical threshold. This result highlights that, despite the Schumpeterian character of the model here, policies that encourage entry may create countervailing effects and may have less than expected or little impact on the overall innovation and growth rate of the economy.

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

---

[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.

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), and to a handful of recent papers building on their work, such as Lentz and Mortensen (2008) and Acemoglu et al. (2013). 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 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 success-

---

[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.

ful 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 important in innovation and productivity growth. 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 et al. (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.

Several papers provide models with innovation from both incumbents and entrants, but based on a different approach from ours. Grossman and Helpman (1991a), Segerstrom and Zolnierek (1999), and Denicolo and Zanchettin (2012) generate this result focusing only on "radical innovations" that replace old products, but assume that the incumbents have lower cost of doing innovation compared to the entrants. Peretto (1999), Luttmer (2007), Atkeson and Burstein (2010), and Acemoglu et al. (2013) consider models in which entrants' new products are fully differentiated from the incumbents' product, thus eschewing creative destruction (and incumbents and entrants are equally likely to invest in new products). In contrast to these papers, in our work incumbents specialize on improving their products while entrants engage in radical R&D in order to create higher quality product and overtake product lines previously operated by incumbents.

Other related papers include Lentz and Mortensen (2008), Klepper (1996) and Acemoglu et al. (2013). Lentz and Mortensen (2008) and Acemoglu et al. (2013) extend Klette and Kortum's model by introducing additional sources of heterogeneity and estimate the model's parameters. 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.

Lastly, our paper is also related to the set of papers on firm heterogeneity and firm size distribution which 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 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

---

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

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. In Section 3, we characterize the balanced growth path and equilibrium firm size distribution. Section Section 5 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 θ 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.

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:

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

---

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

where $x(\nu, t|q)$ is the quantity of the machine of type ν 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 ν 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]

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

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 ν at time t.[12] If

---

[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 $\Delta t \to 0$.

[12] Incumbents could also access the technology for radical innovations, but would choose not to. Arrow's

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 ν at time t. The presence of the strictly decreasing function η 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 assumption on the flow rate of innovation, the rate of radical innovations realized in machine ν 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, $\phi(\cdot)$ 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.[14]

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

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

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

---

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.

[14]The model in Chapter 14.3 in Acemoglu (2008) corresponds to a special case of the model presented here with linear ϕ, i.e., $\phi(z) = \phi z$.

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, which we normalize at $1 - \beta$ without loss of any generality in order to simplify 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

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

where $q(\nu, t)$ refers to the highest quality of the machine of type ν 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 ν 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 ν 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:

Assumption 1 $\kappa \geq \left(\frac{1}{1-\beta}\right)^{\frac{1-\beta}{\beta}}$

So 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.[15] By implication, an entrant that innovates further after its own initial innovation will also be able to set the unconstrained monopoly price.[16] Assumption 1 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 given the marginal cost $1 - \beta$, the profit-maximizing monopoly price is

$$p^x(\nu, t|q) = 1. \tag{7}$$

Combining this with (6) implies

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

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

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

---

[15] This assumption is adopted for simplicity. Our results would not be affected if immediately after entry, new firms set a limit price as in Grossman and Helpman (1991a), though in this case we would have to keep track of changing markups for the first few incremental innovations after entry.

[16] 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 (14) 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 (8) into (2), we obtain that total output is given by

$$Y(t) = \frac{1}{1-\beta} Q(t) L, \tag{10}$$

where

$$Q(t) \equiv \int_0^1 q(\nu, t) \, d\nu \tag{11}$$

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 ν, (11) also implies $Q(0) > 0$ as the relevant initial condition of our economy.[17]

As a by-product, we also obtain that the aggregate spending on machines is

$$X(t) = (1-\beta) Q(t) L. \tag{12}$$

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

$$w(t) = \frac{\partial Y}{\partial L} = \frac{\beta}{1-\beta} Q(t). \tag{13}$$

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 ν:

$$V(v, t | q) = \mathbb{E}_t \left[\int_t^{T(\nu, t)} e^{-\int_t^{t+s} r(t+\widetilde{s}) d\widetilde{s}} \left(\pi(\nu, t+s | q) - z(\nu, t+s) q(t+s) \right) ds \right], \tag{14}$$

where $T(\nu, t)$ is a stopping time where a new entrant enters into the sector ν.[18] Under optimal R&D choice of the incumbents, their value function $V(\nu, t | q)$ satisfies the standard Hamilton-Jacobi-Bellman equation:

$$\begin{aligned}
r(t) V(\nu, t | q) - \dot{V}(\nu, t | q) = \max_{z(\nu, t) \geq 0} \{ & \pi(\nu, t | q) - z(\nu, t) q(\nu, t) \\
& + \phi(z(\nu, t)) (V(\nu, t | \lambda q) - V(\nu, t | q)) - \widehat{z}(\nu, t) \eta(\widehat{z}(\nu, t)) V(\nu, t | q) \},
\end{aligned} \tag{15}$$

---

[17] One might be worried about whether the average quality $Q(t)$ in (11) is well-defined, since we do not know how $q(\nu, t)$ will look like as a function of ν and the function $q(\cdot, t)$ may not be integrable. This is not a problem in the current context, however. Since the index ν has no intrinsic meaning, we can rank the ν's such that $\nu \mapsto q(\nu, t)$ is nondecreasing. Then the average in (11) exists when defined as a Lebesgue integral.

[18] 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 Δs infinitesimal), and 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

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

where $\widehat{z}(\nu,t)\,\eta\,(\widehat{z}(\nu,t))$ is the rate at which radical innovations by entrants occur in sector ν at time t and $\phi\,(z\,(\nu,t))$ is the rate at which the incumbent improves its technology. The first term in (15), $\pi\,(\nu,t|q)$, is flow of profit given by (9), 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 λq) 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).[19] 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:[20]

$$
\begin{aligned}
\eta\,(\widehat{z}\,(\nu,t))\,V\,(\nu,t|\kappa q\,(\nu,t)) &\leq q\,(\nu,t)\,,\text{ and} \\
\eta\,(\widehat{z}\,(\nu,t))\,V\,(\nu,t|\kappa q\,(\nu,t)) &= q\,(\nu,t)\text{ if }\widehat{z}\,(\nu,t) > 0,
\end{aligned}
\tag{16}
$$

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 κq, 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,

$$
\frac{\dot{C}\,(t)}{C\,(t)} = \frac{r\,(t) - \rho}{\theta},
\tag{17}
$$

and the transversality condition takes the form

$$
\lim_{t\to\infty} e^{-\int_0^t r(s)ds}\left[\int_0^1 V\,(\nu,t|q)\,d\nu\right] = 0.
\tag{18}
$$

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 ν 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 (14).

We summarize the conditions for an equilibrium as follows:

---

[19] 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 16).

[20] 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), (12) and (18); time paths for R&D expenditure by incumbents and entrants, $\{z(\nu, t), \widehat{z}(\nu, t)\}_{\nu \in [0,1], t=0}^{\infty}$ that satisfy (15) and (16); 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 (7), (8) and (15); and time paths of wage and interest rates, $\{w(t), r(t)\}_{t=0}^{\infty}$ that satisfy (13) and (17).*

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 3 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:[21]

Definition 3 *A linear BGP is a BGP where $V(\nu, t|q) = vq$ for all ν, 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 2 *The intertemporal elasticity of substitution of the representative household, θ, is sufficiently high, i.e.,*

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

---

[21] We conjecture that all BGPs are in fact linear, but we are unable to prove this except when ϕ is linear.

Intuitively, when the intertemporal elasticity of substitution is higher, from the Euler equation, (17), 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 2 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 (17). 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 (15) can be written as

$$rv = \beta L + \max_{z \geq 0} \{\phi(z)(\lambda - 1)v - z\} - \widehat{z}\eta(\widehat{z})v \qquad (19)$$

and assuming positive entry rate, the free-entry condition (16) can be written as

$$\eta(\widehat{z})\kappa v = 1. \qquad (20)$$

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 (17), we also have $\dot{C}(t)/C(t) = (r - \rho)/\theta = g$, where g is the growth rate of consumption and output.

From (10), the growth rate of output can be expressed as

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

From (19) and (20), in a linear BGP, for all machines, incumbents and entrants will undertake constant R&D $z(v)$ and $\widehat{z}(v)$, respectively. Consequently, the growth rate of $Q(t)$ is[22]

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

---

[22]In a small interval of time Δt, there will be $\phi(z(v))\Delta t$ sectors that experience one innovation by the incumbent (increasing their productivity by λ) and $\widehat{z}(v)\eta(\widehat{z}(v))\Delta t$ sectors that experience replacement by new entrants (increasing productivity by factor of κ). The probability that there will be two or more innovations of any kind within an interval of time Δt 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 Δt and taking the limit as $\Delta t \to 0$, we obtain the formula for the growth rate of $Q(t)$.

Combining this equation with equations (17), (19) gives us a single equation that determines v and thus the key value function $V(\nu, t|q)$ in (15):

$$\beta L = \rho v + (\theta - 1) \phi(z(v))(\lambda - 1)v + z(v) + (\theta(\kappa - 1) + 1)\widehat{z}(v)\eta(\widehat{z}(v))v \qquad (21)$$

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 2 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 $z^* = z(v^*)$ and $\widehat{z}^* = \widehat{z}(v^*)$. The GDP growth rate is

$$g^* = \phi(z^*)(\lambda - 1) + \widehat{z}^*\eta(\widehat{z}^*)(\kappa - 1). \qquad (22)$$

The BGP interest rate, again from the Euler equation (17), is obtained as

$$r^* = \rho + \theta g^*. \qquad (23)$$

Notice that equation (22) 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, (18) is satisfied. The condition for this is $r^* > g^*$ which is also satisfied if $\theta \geq 1$ (or less stringently, if $\rho > (1 - \theta)g^*$). This discussion thus establishes the following proposition (proof in the text).

Proposition 1 *Suppose that Assumption 2 holds and $\rho > (1 - \theta)g^*$, 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 (21), the aggregate growth rate g^* is given by (22), and the interest rate r^* is given by (23). Starting with any initial condition, the economy immediately jumps to this BGP (i.e., always grows at the rate g^*).*

It is useful to consider two special cases of our economy to understand the forces driving growth in this proposition. In the first case, we can shut off radical innovations (e.g., by shutting down entry by setting $\eta(\cdot) = 0$). In this case, though the model would lose its Schumpeterian character, there would still be linear growth because of the AK structure of the model coming from incumbent innovation. This growth rate can be obtained by removing the terms with $\widehat{z}(v)$ in (21) and (22). In the second case, we can shut off incremental

14

innovations by incumbents (e.g., by setting $\phi(\cdot) = 0$). This case would of course correspond to the baseline quality ladder model (e.g., Chapter 14.1 in Acemoglu (2008)), and can be obtained from by removing the terms with $z(v)$ in (21) and (22) to determine the unique linear BGP. In the general case considered in Proposition 1, both of these sources of growth are present and contribute to (but also compete in the determination of) overall growth of the economy.

These two sources create competing effects on growth. In particular, if the contribution of entry to growth increases, then, in response, the contribution of incumbent innovations decreases (because greater entry reduces the value of an incumbent and in particular the value v^*, discouraging incremental innovations). In fact, these competing effects imply that, contrary to standard Schumpeterian models, the effect of policies that discourage entry are ambiguous on the BGP growth rate of the economy. It can be shown that such policies (e.g., a tax on entrants) will reduce growth if $|\phi''(\widehat{z}^*)| > (\kappa - 1)(\lambda - 1)(\beta L)^2/\rho$, and increase it otherwise.

The assumptions leading to sustain endogenous growth in this model are also familiar from the literature. First, the final good production function, equation (2), exhibits constant returns to scale in product quality. Second, the R&D production function (innovation possibilities frontier) uses the final good as in the baseline Schumpeterian model considered in Acemoglu (2008) (and this contrasts with the use of labor in Grossman and Helpman (1991a), for example), but does not have major implications for our results.[23]

Equation (22) also shows that the BGP growth rate of the economy can be decomposed into growth emanating from incremental innovations of incumbents, the first term on the RHS, and growth coming from the radical innovations of entrants, given by the second term on the RHS. We return to this decomposition and explore its quantitative implications in Section 4.

Another interesting implication of our 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 ν and t. We have seen that the quality of an incumbent firm increases at the flow rate $\phi(z^*)$, while the firm is replaced at the flow rate $\hat{z}^* \eta(\hat{z}^*)$. Hence, for Δt sufficiently small, the stochastic process

---

[23] As a result of these features, as equation (21) shows, our economy exhibits scale effect: a higher population leads to a higher value for the intermediate-good producers, and thus to more innovation and growth.

for the size of a particular firm is given by

$$x\left(\nu, t+\Delta t \mid q\right) = \begin{cases} \lambda x\left(\nu, t \mid q\right) & \text{with probability } \phi\left(z^*\right)\Delta t + o\left(\Delta t\right) \\ 0 & \text{with probability } \widehat{z}^*\eta\left(\widehat{z}^*\right)\Delta t + o\left(\Delta t\right) \\ x\left(\nu, t \mid q\right) & \text{with probability } \left(1 - \phi\left(z^*\right)\Delta t - \widehat{z}^*\eta\left(\widehat{z}^*\right)\Delta t\right) + o\left(\Delta t\right) \end{cases} \tag{24}$$

for all ν 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 ν at time s will be bankrupt by time $t \geq s$ by $P\left(t \mid s, \nu\right)$, we clearly have $\lim_{t\to\infty} P\left(t \mid s, \nu\right) = 1$, so that each firm will necessarily die eventually. Equation (24) satisfies Gibrat's Law, which postulates that firm growth is independent of size (e.g., Sutton 1997, Gabaix 1999).[24] This random process of firm growth also leads to non-degenerate distribution of firm sizes as measured by firm size (sales) normalized by average size (sales). In particular, letting $X_1\left(t\right) \equiv X\left(t\right)/\left(1-\beta\right) = \int_0^1 x\left(\nu, t \mid q\right)d\nu$, with $X\left(t\right)$ defined in (4),[25] we have

$$\widetilde{x}\left(\nu, t \mid q\right) \equiv \frac{x\left(\nu, t \mid q\right)}{X_1\left(t\right)} = \widetilde{x}\left(t\right) = \widetilde{q}\left(t\right) \equiv \frac{q\left(t\right)}{Q\left(t\right)},$$

where the first equality defines the variable, the last one defines $\widetilde{q}\left(t\right)$. The law of motion of normalized firm size is similar to (24) 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, $\dot{X}_1\left(t\right)/X_1\left(t\right) = g^* > 0$, for Δt small, the law of motion of normalized firm sizes can be written as

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

The growth decomposition (22) implies that $\widetilde{x}\left(t\right)$ is a random walk, i.e., $E_t\left[\widetilde{x}\left(t+\Delta t\right)\right] = \widetilde{x}\left(t\right)$.[26] Moreover, $\widetilde{x}\left(t\right)$ is positive, so by martingale convergence theorem, $\widetilde{x}\left(t\right)$ converges to a random variable. However, $\widetilde{x}\left(t\right)$ must have finite mean - by definition $E_0\left[\widetilde{x}\left(t\right)\right] = 1$, so the limiting random variable only take either zero or infinity values.[27] In other words, the

---

[24]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 γ_t is a random variable with mean 1 and ε_t is a random variable with mean zero. Both variables are orthogonal to S_t. The law of motion (24) is a special case of this general form.

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

[26]We thank an anonymous referee for suggesting this result.

[27]We show this result by contradiction. Suppose that the limiting random variable has positive mass between zero and infinity, then similar to Lemma 11, we can show that its density function, $f\left(q\right)$, satisfies $f\left(q\right) \geq \frac{B}{q\log(q)}$ for some $\underline{B} > 0$ and for all $q \geq \underline{q}$. This inequality implies infinite mean because $\int_{q_0}^{\infty} \frac{B}{q\log(q)}qdq = \infty$.

16

distribution of normalized firm sizes will continuously expand and a stationary distribution does not exist in a linear BGP. This is because 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.[28] In the next section, we show that by adding another type of (imitative) entrants, which act as a reflecting-lower barrier on firm sizes, we obtain a stationary distribution of firm sizes in a BGP. This firm size distribution resembles the one observed in the data. 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).

3   Firm Dynamics and Growth with Knowledge Spillovers

The model so far and the growth decomposition provided in equation (22) assume only one type of new entrants which are entrants with radical innovations. In reality, some entrants committing not with radical innovations but by building on the knowledge spillovers created by existing firms (e.g., see Luttmer (2007) and Acemoglu et al. (2013)). In this section, we extend the model to include knowledge spillovers that allow imitation for (some) entrants, and show that an extended model of this type maintains the qualitative features of the simpler model presented so far but also generates a rich stationary firm size distribution along the BGP.

More specifically, we allow for costly ("imitative") entry building on knowledge spillovers. Potential entrants can pay some cost to obtain a technology with quality proportional to the current average quality in the economy. This implies that when a sector falls significantly below average quality, imitative entry becomes profitable. We then show that the economy incorporating this type of initiative entry has a well-defined equilibrium. Moreover, when the technology of entry by imitation is not very productive, then the BGP equilibrium is arbitrarily close to the one that is characterized in Proposition 1. We also derive a general analog to the growth decomposition in equation (22), now consisting of three components: innovation by incumbents, by radical entrants, and by entrants by knowledge spillover.

---

[28]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).

3.1 The Economy with Imitative Entry

We consider a third type of innovation based on imitation relying on knowledge spillovers from the current knowledge base of the economy. 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 (11).[29] 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.[30]

Firm value is again given in (14) except that $T(\nu, t)$ 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 (15), but they also take into account 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 ϵ times average quality (clearly, ϵ is a function of ω, 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.[31]

In order to prove the existence of a stationary BGP in the economy with imitation, we need a slightly stronger condition on the inverse of the intertemporal elasticity of substitution, θ, than in Assumption 2

Assumption 2b $\theta \geq 1$.

We also impose the following technical condition.

---

[29] For tractablity we assume that ω is constant as opposed to a random variables realized after entry as in Luttmer (2007) and Acemoglu et al. (2013), but this is without any important consequences for the implications of the model.

[30] There are also alternative ways of introducing entrants by knowledge-spillover. An alternative, which leads to very similar results, is an environment where firms have to pay a fixed cost in terms of labor to operate and existing firms are replaced by entrants by knowlege spillover. See also Luttmer (2007, 2010b) and Atkeson and Burstein (2010, 2013).

[31] Note that because this type of imitative entry is costly, and the cost would be the same for incumbents vs. entrants, by a similar reasoning to Arrow's replacement effect,incumbents that become very low productivity would still not find it profitable to use this technology (i.e., before the incumbents do so, entrants would do it).

Assumption 3 *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 η, Assumption 3 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 the 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 3. We also note for future reference that the Pareto distribution takes the form

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

with $\Gamma > 0$ and $y \geq \Gamma$; χ 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 $\Pr[\widetilde{x} \geq y] \propto \Gamma y^{-\chi}$.

Proposition 2 *Suppose the BGP equilibrium in the baseline economy is described by Proposition 1, in particular with v^*, g^* and r^* as given by to (21), (22) and (23), and Assumptions 1b and 3 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 imitative entry 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 ω 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.

The next proposition provides a simple growth decomposition between incumbents, radical entrants, and imitative entrants generalizing the one provided in equation (22) above.

Proposition 3 *In a stationary BGP, the growth rate of the economy has the following decomposition:*

$$
g = \underbrace{(\lambda - 1)\, \mathbb{E}_F\left[\phi\left(z\left(\widetilde{q}\right)\right)\widetilde{q}\right]}_{Innovation\ from\ incumbents}
$$
$$
+ \underbrace{(\kappa - 1)\, \mathbb{E}_F\left[\widehat{z}\left(\widetilde{q}\right)\eta\left(\widehat{z}\left(\widetilde{q}\right)\right)\widetilde{q}\right]}_{Innovation\ from\ radical\ entrants}
$$
$$
+ \underbrace{\epsilon g f\left(\epsilon\right)\left(\omega - \epsilon\right)}_{Innovation\ from\ imitative\ entrants} \tag{25}
$$

Proof. This follows from the proof of Proposition 2. In particular, Step 3 in the sketch of proof for Proposition 2, we show that the implied growth rate $g'\left(g\right)$ satisfies the decomposition (25). In a stationary BGP, $g'\left(g\right) = g$, thus g also satisfies the growth decomposition. ∎

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

Proposition 4 *The stationary equilibrium distribution of firm sizes in the economy with imitation (characterized in Proposition 2) exists and has a Pareto tail with the shape parameter $\chi = \chi\left(\omega\right) > 1$ 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, $f\left(\widetilde{x}\right)$, satisfies*

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

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

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

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

Proof. The detail of the proof in the Appendix and follows from the proof of Proposition 2. ∎

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.[32] 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.[33] 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 2.

3.2   Sketch of the Proof of Proposition 2

The proof consists of showing that for each $\mu_e \in \left(\underline{\mu}, \overline{\mu}\right)$ 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

---

[32] The gamma distribution used in Luttmer (2007) has a Pareto tail according to the definition in Proposition 4.

[33] Footnote 24 in Atkeson and Burstein (2010) provides a heuristic derivation of the relationship between the endogenous investment of incumbents and the endogeous tail index of the stationary distribution of firm sizes. However, a crucial step in this heuristic derivation is that innovation rates are invariant across firms in the presence of fixed cost. This property only holds approximately in the limit when firm size is sufficiently large. Similarly, the threat of imitative entry in our model makes the incumbent firms' innovation rate a function of firm size. Thus a full derivation needs to take this dependence into account, which is what our rigorous approach accomplishes.

endogenously in our economy.[34]

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 λq and κq. 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 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 ν at time t which takes the form

$$V_g(\nu, t|q) = Q(t) \widehat{V}_g \left(\frac{q(\nu, t)}{Q(t)} \right), \tag{26}$$

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 ν at time t and replace an incumbent if $q(\nu, t) \leq \epsilon_g(\omega) Q(t)$. 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 (26) in (15), and using the fact that

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

we obtain that

$$\begin{aligned} (r - g) \widehat{V}_g(\widetilde{q}) + g\widetilde{q}\widehat{V}'_g(\widetilde{q}) &= \beta L\widetilde{q} + \max_{z(\nu, t) \geq 0} \left\{ \phi(z(\nu, t)) \left(\widehat{V}_g(\lambda\widetilde{q}) - \widehat{V}_g(\widetilde{q}) \right) - z(\nu, t)\widetilde{q} \right\} \\ &\quad - \widehat{z}(\nu, t) \eta(\widehat{z}(\nu, t)) \widehat{V}_g(\widetilde{q}), \end{aligned} \tag{27}$$

where $r = \rho + \theta g$ and $\widetilde{q}(\nu, t) = q(\nu, t)/Q(t)$. The free-entry condition for radical innovation, (16), can then be written as

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

---

[34] Atkeson and Burstein (2010, 2013) also present models with endogenous innovation and firm heterogeneity. However, in these models, growth is either exogenous or given as the equilibrium of a semi-endogenous growth model, making it essentially exogenous. This ensures a block-recursive structure in which interest rate and growth rates are independent of firm size distribution. The general equilibrium interactions present in our model arise whenever equilibrium objects, in particular the interest and the growth rates, depend on the endogenous firm size distribution.

Moreover, the free-entry condition for imitation implies

$$\widehat{V}_g(\omega) = \mu_e \omega. \tag{28}$$

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

$$\widehat{V}_g(\epsilon_g) = 0. \tag{29}$$

In the Appendix, we show that when $\mu_e \in (\underline{\mu}, \overline{\mu})$, 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 (27).

To establish the existence of a solution $\widehat{V}_g(\widetilde{q})$ to the functional equation (27), 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

$$0 = F'(y) yg - \int_{\frac{y}{\lambda}}^{y} \phi(z(\widetilde{q})) dF(\widetilde{q}) - \int_{\frac{y}{\kappa}}^{y} \widehat{z}(\widetilde{q}) \eta(\widehat{z}(\widetilde{q})) dF(\widetilde{q}). \tag{30}$$

If $y < \omega$, then

$$0 = F'(y) yg - F'(\epsilon) \epsilon g - \int_{\frac{y}{\lambda}}^{y} \phi(z(\widetilde{q})) dF(\widetilde{q}) - \int_{\frac{y}{\kappa}}^{y} \widehat{z}(\widetilde{q}) \eta(\widehat{z}(\widetilde{q})) dF(\widetilde{q}) \tag{31}$$

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 (ϵ, 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 ω. This mass consists of firms that are in the neighborhood of ϵ, do not experience any innovation, and are therefore drifted to below ϵ due to the growth rate g of the average quality Q; it is equal to $F'(\epsilon) \epsilon g$. By definition of

23

a stationary distribution, the total mass of firms moving out of the interval (ϵ, 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: This analysis gives an implied growth rate of the average product quality $g' = \dot{Q}/Q$ as a function of the current growth rate g, the innovation rates, $z_g(\tilde{q})$, and $\hat{z}_g(\tilde{q})$, the imitation threshold ϵ_g, and the equilibrium stationary distribution F_g. In particular,

$$g'(g) = \frac{(\lambda - 1) \, \mathbb{E}_{F_g}\left[\phi\left(z\left(\tilde{q}\right)\right) \tilde{q}\right] + (\kappa - 1) \, \mathbb{E}_{F_g}\left[\hat{z}_g\left(\tilde{q}\right) \eta\left(\hat{z}_g\left(\tilde{q}\right)\right) \tilde{q}\right]}{1 - \epsilon_g F'_g\left(\epsilon_g\right)\left(\omega - \epsilon_g\right).} \tag{32}$$

This formula, derived formally in the Appendix, is similar to the decomposition of growth in (22). The numerator combines the innovation rates of incumbents and entrants, respectively $(\lambda - 1) \, \mathbb{E}_{F_g}\left[\phi\left(z\left(\tilde{q}\right)\right) \tilde{q}\right]$ and $(\kappa - 1) \, \mathbb{E}_{F_g}\left[\hat{z}_g\left(\tilde{q}\right) \eta\left(\hat{z}_g\left(\tilde{q}\right)\right) \tilde{q}\right]$, where $\mathbb{E}_{F_g}$ is used as a shorthand for the integrals using the density $dF_g\left(\tilde{q}\right)$ as in (30) and (31). 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

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

where $g'(g)$ is given by (32). In the Appendix, we establish the existence of a solution $g^*\left(\omega\right)$ to this equation.

4   Some Suggestive Numbers and Simulations

In this section, we present the quantitative implications of the growth decomposition in equation (22) and use simulations to compute the implied stationary firm size distribution.

4.1   Baseline Parameters

The explicit characterization of equilibrium enables us to obtain simple expressions for how much of productivity growth is driven by innovation by entrants (creative destruction) and how much of it comes from productivity improvements by incumbents. In particular, we can use equation (22), which decomposes growth into the component coming from incumbent firms (the first term in that equation) and that coming from new entrants (the second term).

Our purpose is not to undertake a systematic estimation of the parameters of the model or calibration, but to provide some illustrative numbers. We 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 θ, 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. Assumption 1 then requires 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 $\lambda = 1.2$, so that productivity gains from a radical innovation is about ten times that of a standard "incremental" innovation by incumbents (i.e., $(\kappa - 1)/(\lambda - 1) = 10$). For the functions $\phi(z)$ and $\eta(z)$, we adopt the functional form in (3) and choose the benchmark values of $\alpha = 0.9$ 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 κ, we obtain $\widehat{z}^* \eta(\widehat{z}^*)$ equals 0.0033. Given the interest rate of r^*, this corresponds to radical innovation on average once every 15 years ($r^*/\widehat{z}^* \eta(\widehat{z}^*)$). Similarly, the rate of incremental innovation is once very nine months.

Varying these baseline 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).

Table 1 shows how these numbers change as we vary the parameters β, κ, λ and α. The first five columns of the table give the choices of parameters. The value of B is chosen so that the equilibrium growth rate $g^* = 0.02$. The next two columns report the innovation rate by the entrants,$r^*/\ \hat{z}^* \eta(\hat{z}^*)$, and by the incumbents, $r^*/\phi(z^*)$. The final column reports the fraction of total productivity growth accounted for by entrants, i.e., $(\kappa - 1)\hat{z}^* \eta(\hat{z}^*)/g^*$.

<div align="center">

Table 1

</div>

	Parameter Values					$r^*/\eta\left(\hat{z}^*\right)\hat{z}^*$	$r^*/\phi\left(z^*\right)$	$(\kappa-1)\eta\left(\hat{z}^*\right)\hat{z}^*/g^*$
1.	$\kappa=3$	$\lambda=1.2$	$\beta=2/3$	$A=0.0977$	$\alpha=0.9$	14.96	0.75	0.333
2.	$\kappa=2$	$\lambda=1.2$	$\beta=2/3$	$A=0.0977$	$\alpha=0.9$	7.34	0.75	0.340
3.	$\kappa=4$	$\lambda=1.2$	$\beta=2/3$	$A=0.0977$	$\alpha=0.9$	22.58	0.75	0.332
4.	$\kappa=3$	$\lambda=1.1$	$\beta=2/3$	$A=0.0977$	$\alpha=0.9$	7.15	0.83	0.699
5.	$\kappa=2$	$\lambda=1.1$	$\beta=2/3$	$A=0.0977$	$\alpha=0.9$	3.55	0.84	0.703
6.	$\kappa=3$	$\lambda=1.2$	$\beta=2/3$	$A=0.0488$	$\alpha=0.9$	7.15	1.66	0.699
7.	$\kappa=3$	$\lambda=1.2$	$\beta=1/2$	$A=0.0977$	$\alpha=0.9$	14.03	0.77	0.356
8.	$\kappa=3$	$\lambda=1.2$	$\beta=2/3$	$A=0.0977$	$\alpha=0.8$	9.48	1.05	0.527

Note. This table reports flow rates of innovation by entrants $(r^*/\eta\left(\hat{z}^*\right)\hat{z}^*)$, incumbents $(r^*/\phi\left(z^*\right))$, and the fraction of productivity growth accounted for by entrants $((\kappa-1)\eta\left(\hat{z}^*\right)\hat{z}^*/g^*)$. In each case, $L=1$, $\theta=2$, $\rho=0.01$, and B and A are chosen such that $r^*=0.05$ and $g^*=0.02$.

The first row reports the benchmark we have just discussed. The second row reduces κ to 2 so that productivity gain from a radical innovation is now five times that of incremental innovation $((\kappa-1)/(\lambda-1)=5)$. This changes the rate of innovation of entrants to once every 7 years, and leaves the incremental innovation rate unchanged. The contribution of entrants to productivity growth is slightly higher, 34%. The third row increases κ to 4. In this case, again the only number that changes significantly is the rate of innovation of entrants.

The fourth and fifth row reduces the size of innovations by incumbents to $\lambda=1.1$, and the sixth row reduces the R&D productivity of the incumbent by half. In these cases the contribution of entrants increase to more than two thirds. The remaining rows vary the values of β and α. When β is set to 1/2, the contribution of entrants does not change significantly. However, when α is set to 0.8, the contribution of entrants increases to above one half.

These results therefore show that, in contrast to basic Schumpeterian models, the contribution of incumbents to productivity growth is often quite substantial. Nevertheless, Table 1 also 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).

4.2 Simulations

We next examine the growth decomposition under the presence of imitative entrance, (25), as well as the stationary distribution of firm size. Starting from the baseline parameters in subsection 4.1, we choose the cost of (imitative) entry based on knowledge spillovers to $\mu_e = 15$. The next table show how this decomposition, as well as the stationary firm size distribution change, as we vary ω. At $\omega = 0$, we converge to the baseline economy without knowledge spillover as shown in Proposition 2. As we increase ω, the contribution of imitative entrants to aggregate growth increases and the tail index of the stationary distribution of firm size increases as well, i.e., the distribution has thinner tails as we increase ω.

ω	g	% by incumbents	% by radical entrants	% by imitative entrants	tail index
0	2%	66.7%	33.3%	0%	1
0.1	2.02%	65.7%	32.7%	1.6%	1.05
0.3	2.07%	64.14%	31.23%	4.63%	1.17
0.5	2.11%	62.53%	29.97%	7.50%	1.26

For the case $\omega = 0.1$, Figure 1 shows the distribution of firm size, i.e., log rank as a function of log size. The distribution is close to Zipf's law with Pareto index of the tail around 1.05. Figure 2 shows the rate of innovation of incumbents and entrants (normalized by the innovation size and the aggregate growth rate). The investment in R&D by incumbents is non-linear and exhibit the "escape competition effect" (Aghion et al. (2001)). In particular, close to the threshold where imitative entry takes place, the incumbents invest more into incremental R&D.

5   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
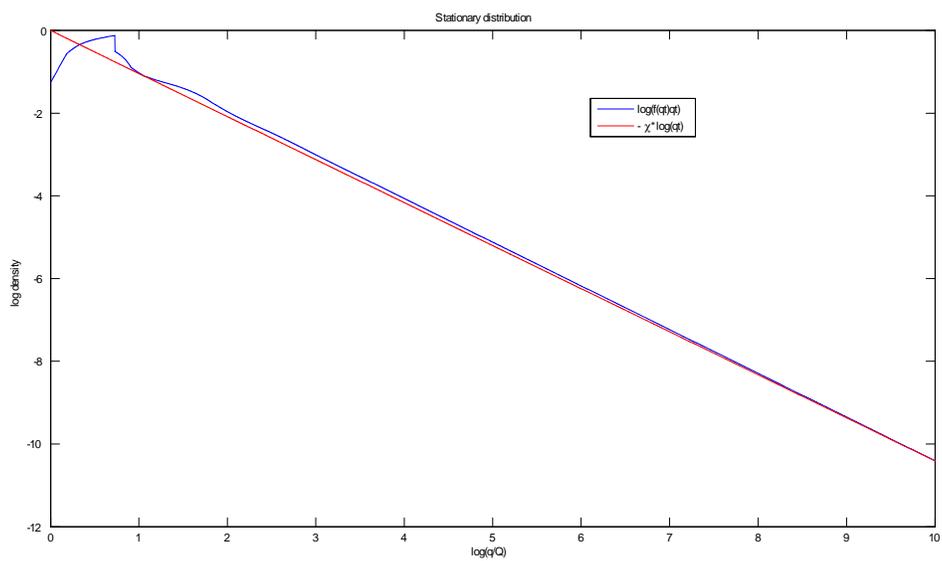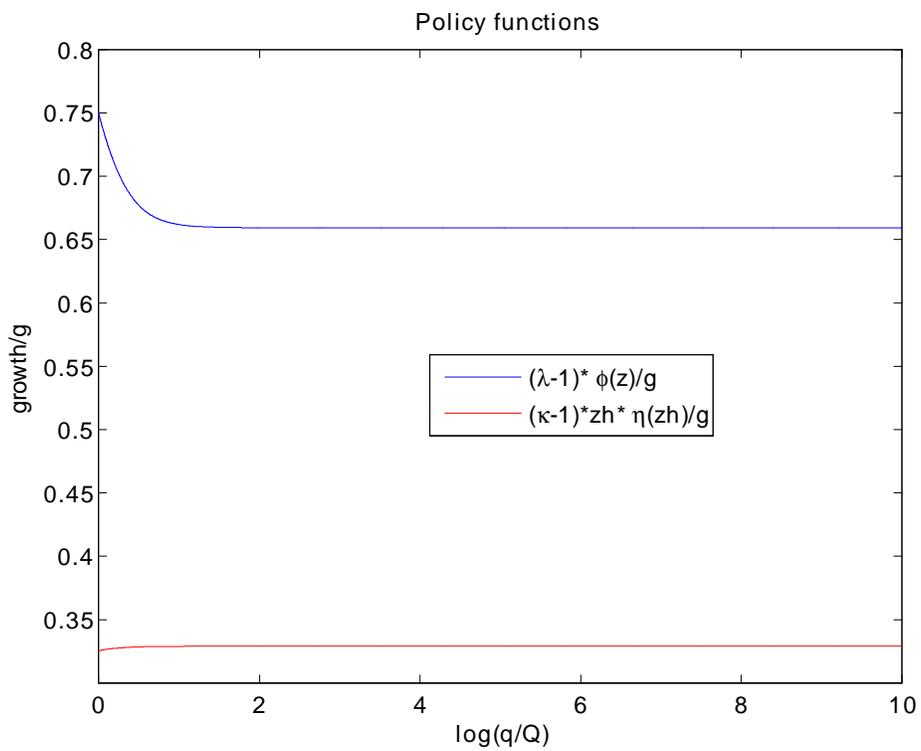
Figure 1: Stationary distribution

Figure 2: Policy functions

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.

Appendix

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 that $Q(t) \equiv \int_0^1 q(v,t)\,dv$, 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

$$Q(t+\Delta t) = \int_0^1 q(v,t+\Delta t)\,dv$$

$$= \int_{0,q(v,t)\geq \epsilon Q(t)(1+g\Delta t)}^1 \left(\begin{array}{c} \phi\left(z\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t \lambda q(v,t) + \widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t \kappa q(v,t) \\ + \left(1 - \phi\left(z\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t - \widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t\right) q(v,t) \end{array} \right) dv$$

$$+ \int_{0,\epsilon Q(t) < q(v,t) < \epsilon Q(t)(1+g\Delta t)}^1 \left(\begin{array}{c} \phi\left(z\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t \lambda q(v,t) + \widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t \kappa q(v,t) \\ + \left(1 - \phi\left(z\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t - \widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t\right) \omega Q(t) \end{array} \right) dv.$$

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

$$Q(t+\Delta t) = \lambda Q(t)\Delta t \int_0^1 \phi\left(z\left(\frac{q(v,t)}{Q(t)}\right)\right)\frac{q(v,t)}{Q(t)}dv + \kappa Q(t)\Delta t \int_0^1 \widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\right)\frac{q(v,t)}{Q(t)}dv$$

$$+ \int_{0,q(v,t)\geq \epsilon Q(t)(1+g\Delta t)}^1 \left(1 - \phi\left(z\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t - \widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t\right) q(v,t)\,dv$$

$$+ \int_{0,\epsilon Q(t) < q(v,t) < \epsilon Q(t)(1+g\Delta t)}^1 \left(1 - \phi\left(z\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t - \widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t\right) \omega Q(t)\,dv.$$

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

$$Q(t+\Delta t) = Q(t) + \underbrace{(\lambda-1)Q(t)\Delta t \int_0^1 \phi\left(z\left(\frac{q(v,t)}{Q(t)}\right)\right)\frac{q(v,t)}{Q(t)}dv}_{\text{Innovation from Incumbents}}$$

$$+ \underbrace{(\kappa-1)Q(t)\Delta t \int_0^1 \widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\right)\frac{q(v,t)}{Q(t)}dv}_{\text{Innovation from Entrants}}$$

$$+ \underbrace{\int_{0,\epsilon Q(t) < q(v,t) < \epsilon Q(t)(1+g\Delta t)}^1 \left(\begin{array}{c} 1 - \phi\left(z\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t \\ -\widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\eta\left(\widehat{z}\left(\frac{q(v,t)}{Q(t)}\right)\right)\Delta t \end{array} \right)(\omega Q(t) - q(v,t))\,dv}_{\text{Innovation from Imitators}}.$$

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 $f(\cdot)$

$$Q(t+\Delta t) = Q(t) + (\lambda-1)Q(t)\int \phi(z(\widetilde{q}))\,\widetilde{q}\,dF(\widetilde{q})\,\Delta t$$

$$+ (\kappa-1)Q(t)\int \widehat{z}(\widetilde{q})\,\eta(\widehat{z}(\widetilde{q}))\,\widetilde{q}\,dF(\widetilde{q})\,\Delta t$$

$$+ F(\epsilon(1+g\Delta t))(\omega - \epsilon)Q(t).$$

31

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

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

as in (32). When $\omega = 0$ we have
$$g = (\lambda - 1) \mathbb{E}_F \left[\phi \left(z \left(\tilde{q} \right) \right) \tilde{q} \right] + (\kappa - 1) \mathbb{E}_F \left[\hat{z} \left(\tilde{q} \right) \eta \left(\hat{z} \left(\tilde{q} \right) \right) \tilde{q} \right],$$

and when $z \left(\tilde{q} \right) \equiv z^*$ and $\hat{z} \left(\tilde{q} \right) \equiv \hat{z}^*$
$$g = (\lambda - 1) \phi \left(z^* \right) + (\kappa - 1) \hat{z}^* \eta \left(\hat{z}^* \right)$$

as in (22), given that $\mathbb{E}[\tilde{q}] = 1$. ∎

Proof of Proposition 2:

Let us first define $I_i \left(\tilde{v} \right) \equiv \max_{z \geq 0} \phi \left(z \right) \tilde{v} - z$ and $I_e \left(u \right) \equiv \frac{1}{\kappa u} \eta^{-1} \left(\frac{1}{\kappa u} \right)$. Intuitively, $I_i \left(\tilde{v} \right)$, for $\tilde{v} = (\lambda - 1) v$, 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 ϕ and η to the functions I_i and I_e. Using these notations we can also define v_g as a solution of the equation

$$v = \frac{\beta L + I_i \left((\lambda - 1) v \right)}{r + I_e \left(v \right)}, \tag{33}$$

in which $r = \rho + \theta g$. The following lemma establishes some properties of v_g around the equilibrium values (v^*, g^*). 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 (19).

Lemma 1 *Suppose Assumption 1b is satisfied. There exists $\Delta > 0$ such that for each $g \in (g^* - \Delta, g^* + \Delta)$, there exists a unique $v_g \in (v^* - \Delta, v^* + \Delta)$ that satisfies equation (33). Moreover v_g is strictly decreasing in g.*

Proof. We rewrite equation (33) as $\Phi \left(v, g \right) = 0$ where
$$\Phi \left(v, g \right) = v \left(\rho + \theta g + I_e \left(v \right) \right) - \beta L - I_i \left((\lambda - 1) v \right).$$

As we show below, $\partial \Phi \left(v^*, g^* \right) / \partial v > 0$, so the implicit function theorem guarantees the existence and uniqueness of (g, v_g) in the neighborhood of (g^*, v^*), 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

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

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 (22) 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 2 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 (21), (22) and (23), and Assumption 1b and Assumption 3 are satisfied. Then there exist $0 < \underline{\mu} < \overline{\mu}$ and $\Delta > 0$ such that for any $\mu_e \in \left(\underline{\mu}, \overline{\mu} \right)$, $g \in [g^*, g^* + \Delta]$ 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 (27), (28) and (29).*

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 3 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*

$$
\begin{aligned}
& rU (p) + gU' (p) \\
&= \beta L + \max_{z \geq 0} \{ \phi (z) (\lambda U (p + \ln \lambda) - U (p)) - z \} - \widehat{z} (p) \eta (\widehat{z} (p)) U (p), \qquad (34)
\end{aligned}
$$

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

$$
U (0) = 0 \text{ and } \lim_{p \to \infty} U (p) = v_g, \qquad (35)
$$

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

Notice that the conditions (27) and (29) on $\widehat{V}_g$ translate into the conditions (34) and the first part of (35) on U. In order to apply Schauder's fixed point theorem, we need to find a subset F of continuous functions $U : [g^* - \Delta, g^* + \Delta] \times \mathbb{R}^+ \to \mathbb{R}$ that satisfy the boundary conditions (35), and a continuous mapping T that summarizes the functional equation (34). We need $T (F)$ to be a compact subset of F. F and T are constructed in Definitions 5 and 6. Lemmas 4 and 5 show that $T (F)$ is a compact subset of F. 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 $C^0\left([g^* - \Delta, g^* + \Delta] \times \mathbb{R}^+, \mathbb{R}\right)$ *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*

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

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

$$v_g - v_g e^{-\theta p} \le U(g, p) \le v_g + v_g e^{-\theta p} \text{ for all } p \ge 0. \tag{36}$$

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

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

satisfies the following ordinary differential equation[35]

$$\begin{aligned} g\left(T_g u\right)'(p) + \left(r_g + I_e\left(u\left(p + \ln \kappa\right)\right)\right)\left(T_g u\right)(p) \\ = \beta L + I_i\left(\lambda u\left(p + \ln \lambda\right) - T_g u(p)\right). \end{aligned} \tag{37}$$

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

$$r_g = \rho + \theta g. \tag{38}$$

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*

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

Lemma 4 *Suppose Assumptions 1b and 3 are satisfied, then* $T(F) \subset F$.

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

$$\underline{k}_g(p) \le U(g, p) \le \bar{k}_g(p).$$

Let $k_g(p) = T_g U(g, p)$ then, also by definition (37) implies that

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

So if

$$g\bar{k}_g'(p) > \beta L - I_i\left(\lambda \bar{k}_g\left(p + \ln \lambda\right) - \bar{k}_g(p)\right) - \left(r_g + I_e\left(\underline{k}_g\left(p + \ln \kappa\right)\right)\right) \bar{k}_g(p), \tag{39}$$

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

$$g\underline{k}_g'(p) < \beta L - I_i\left(\lambda \underline{k}_g\left(p + \ln \lambda\right) - \underline{k}_g(p)\right) - \left(r_g + I_e\left(\bar{k}_g\left(p + \ln \kappa\right)\right)\right) \underline{k}_g(p) \tag{40}$$

---

[35]Standard results from the theory of ordinary differential equations ensure the existence and uniqueness of $T_g u(p)$ if $u \in F$ 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 3 to show (39) and (40). Indeed, the two inequalities can be re-written as (for all $0 < x \le v_g$):

$$-g\theta x > \beta L + I_i\left((\lambda - 1)v_g + \left(\lambda^{1-\theta} - 1\right)x\right) - \left(r_g + I_e\left(v_g - \kappa^{-\theta}x\right)\right)(v_g + x), \qquad (41)$$

and

$$g\theta x < \beta L + I_i\left((\lambda - 1)v_g - \left(\lambda^{1-\theta} - 1\right)x\right) - \left(r_g + I_e\left(v_g + \kappa^{-\theta}x\right)\right)(v_g - x). \qquad (42)$$

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 (41) 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 (38) now implies that $r_g > g\theta$ and, from Assumption 1b, $\theta \ge 1$, yields

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

It remains to show that

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

or

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

or

$$\frac{\min \varepsilon_{I_e}}{\min \varepsilon_{I_e} + 1} \ge \kappa^{-\theta} \qquad (43)$$

Similarly, it is sufficient to show that the derivative of the left hand side of (42) 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 $\left(v_g + \kappa^{-\theta}x\right) \ge \frac{1}{\min \varepsilon_{I_e}}\kappa^{-\theta}v_g$, or equivalently, if

$$\min \varepsilon_{I_e} \ge \kappa^{-\theta}. \qquad (44)$$

Since

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

Assumption 3 implies both (43) and (44).

∎

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

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

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

35

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

$$gD'_g(p) + Tf_n(p) + \left(\frac{dr_g}{dg} + I_e(f_n(p + \ln \kappa))\right)Tf_n(p)$$

$$= \left(I'_i(\lambda f_n(p + \ln \lambda) - Tf_n(p)) - (r_g + I_e(f_n(p + \ln \kappa)))\right)D_g(p)$$

So $D_g(p)$ is uniformly bounded over $[g^* - \Delta, g^* + \Delta] \times [0, M]$ for any $M > 0$. Therefore, for each $M = 1, 2, \ldots$, 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 $[g^* - \Delta, g^* + \Delta] \times [0, M]$.

- $M \implies M+1$: 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 $f^*\big|_{[g^* - \Delta, g^* + \Delta] \times [0, M]} = f_M^*$ 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 $f^* \in F$.
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 (36) in the definition of F, 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 F.*

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 $\mathcal{F}$ 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 $T(\mathcal{F})$, 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 $\mathcal{F}$: $TU = U$. Or equivalently for each $g \in [g^* - \Delta, g^* + \Delta]$, $u(\cdot) = U(g,.)$ satisfies $u(0) = 0$ and (34). The limit at infinity in (35) follows directly from the definition of $\mathcal{F}$. 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 ϵ_g such that the value function $\widehat{V}_g(\widetilde{q}) = \widetilde{q}U_g(\ln(\widetilde{q}) - \ln\epsilon_g)$ satisfies (27), (28) and (29).

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 Δ 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 ω, let

$$\epsilon_g = \omega/\exp(\omega_g) < \omega(1-\beta)^{\frac{1-\beta}{\beta}}. \tag{45}$$

and let $\widehat{V}_g(\widetilde{q}) = \widetilde{q}U_g(\ln(\widetilde{q}) - \ln\epsilon_g)$. Then $\widehat{V}_g$ satisfies (27), (28) and (29).   ■

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 4. 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 (30) and (31) 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 (30) and (31) become:

If $p > \ln \omega_g$

$$0 = h_g(p) g - \int_{p - \ln \lambda}^p \phi(z_g(\widetilde{p})) h_g(\widetilde{p}) d\widetilde{p} - \int_{p - \ln \kappa}^p \widehat{z}_g(\widetilde{p}) \eta(\widehat{z}_g(\widetilde{p})) h_g(\widetilde{p}) d\widetilde{p}. \tag{46}$$

If $p \leq \ln \omega_g$

$$0 = h_g(p) g - h_g(0) g - \int_{p - \ln \lambda}^p \phi(z_g(\widetilde{p})) h_g(\widetilde{p}) d\widetilde{p} - \int_{p - \ln \kappa}^p \widehat{z}_g(\widetilde{p}) \eta(\widehat{z}_g(\widetilde{p})) h_g(\widetilde{p}) d\widetilde{p}. \tag{47}$$

We also have $h_g(p) = 0$ for all $p \leq 0$. 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

$$gh'_g(p) = \phi(z_g(p)) h_g(p) - \phi(z_g(p - \ln \lambda)) h_g(p - \ln \lambda)$$
$$+ \widehat{z}_g(p) \eta(\widehat{z}_g(p)) h_g(p) - \widehat{z}_g(p - \ln \kappa) \eta(\widehat{z}_g(p - \ln \kappa)) h_g(p - \ln \kappa).$$

We rewrite this equation as

$$gh'_g(p) - (\phi(z_g(p)) + \widehat{z}_g(p)\eta(\widehat{z}_g(p))) h_g(p) = -\phi(z_g(p - \ln\lambda)) h_g(p - \ln\lambda)$$
$$- \widehat{z}_g(p - \ln\kappa)\eta(\widehat{z}_g(p - \ln\kappa)) h_g(p - \ln\kappa).$$

Using the variation of constant formula, this equation yields a unique equation for $0 \le p < \omega_g$ given $h_g(0)$. For $p \ge \omega_g$ 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 χ 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 χ. 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 $\frac{1}{g}\phi\left(z_g^*\right)\frac{\lambda^{\chi}-1}{\chi}+\frac{1}{g}\widehat{z}_g^*\eta\left(\widehat{z}_g^*\right)\frac{\kappa^{\chi}-1}{\chi}=1$ and the functions

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

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

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

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

and

$$\left|\frac{1}{g}\phi\left(z_g\left(p-\ln\lambda\right)\right)-\frac{1}{g}\phi\left(z_g^*\right)\right|<\delta$$

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

We will now state and prove a key lemma. Proposition 4 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*

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

and

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

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

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

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

and

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

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 $h_g\left(p\right)\geq2\overline{B}\left(\delta\right)e^{-\left(\chi-\xi\right)p}$. Consider the infimum of those p, then

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

In the other hand, the equation determining h_g implies

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

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

$$= 2\overline{B}\left(\delta\right)\left(\frac{1}{g}\phi\left(z_g^*\right)+\delta\right)\frac{\lambda^{\chi-\xi}-1}{\chi-\xi}e^{-(\chi-\xi)p} + 2\overline{B}\left(\delta\right)\left(\frac{1}{g}\widehat{z}_g^*\eta\left(\widehat{z}_g^*\right)+\delta\right)\frac{\kappa^{\chi-\xi}-1}{\chi-\xi}e^{-(\chi-\xi)p}$$

$$< 2\overline{B}\left(\delta\right)e^{-(\chi-\xi)p}.$$

This yields a contraction. Therefore

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

Similarly, we can show that

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

∎

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

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

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., $\lim_{g\longrightarrow g^*}\int_0^\infty h_g\left(p\right)e^p dp = \infty$. This is intuitively true using Lemma 4, because $h_{g^*}\left(p\right) \propto e^{-p}$ so $\int_0^\infty h_{g^*}\left(p\right)e^p dp \propto \int_0^\infty 1 dp = \infty$. Unfortunately, this does not work formally because Lemma 4 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*

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

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

$$\phi\left(z_{g^*}\left(\widetilde{p}\right)\right) > \phi\left(z_{g^*}^*\right) - g^*Ce^{-\theta\widetilde{p}}$$
$$\widehat{z}_{g^*}\left(\widetilde{p}\right)\eta\left(\widehat{z}_{g^*}\left(\widetilde{p}\right)\right) > \widehat{z}_{g^*}^*\eta\left(\widehat{z}_{g^*}^*\right) - g^*Ce^{-\theta\widetilde{p}}$$
$$\forall\widetilde{p} \geq p_0.$$

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

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

41

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

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

(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 2.

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 4. $\Phi(g)$ 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 $\omega > 0$, ϵ_g is defined as in (45). 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 (32) can be written using h_g and the change of variable $\widetilde{q} = \epsilon_g e^p$:

$$g' = \frac{(\lambda - 1) \int_{\epsilon_g}^{\infty} \widetilde{q} \phi(z(\widetilde{q})) f_g(\widetilde{q}) d\widetilde{q} + (\kappa - 1) \int_{\epsilon_g}^{\infty} \widetilde{q} \widehat{z}(\widetilde{q}) \eta(\widehat{z}(\widetilde{q})) f_g(\widetilde{q}) d\widetilde{q}}{1 - \epsilon_g f_g(\epsilon_g)(\omega - \epsilon_g)}$$

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

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

$$\epsilon \frac{(\lambda - 1) \int_0^{\infty} e^p \phi(z_{g^* + \Delta}(p)) h_{g^* + \Delta}(p) dp + (\kappa - 1) \int_0^{\infty} e^p \widehat{z}_{g^* + \Delta}(p) \eta(\widehat{z}_{g^* + \Delta}(p)) h_{g^* + \Delta}(p) dp}{1 - h_{g^* + \Delta}(0)(\omega - \epsilon)}$$

$$< g^* + \Delta. \tag{48}$$

We will show that, for each ω such that $0 < \omega < \overline{\varpi}$, 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 ω_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 (48), 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 (36), which themselves converge uniformly. ■

References

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

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

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

Daron Acemoglu, Ufuk Akcigit, Nicholas Bloom, and William Kerr. Innovation, reallocation and growth. MIT Working Paper, 2013.

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

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.

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. *Journal of Political Economy*, 118:433–83, 2010.

Andrew Atkeson and Ariel Burstein. Aggregate implications of innovation policy. UCLA Working Paper, 2013.

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.

Vincenzo Denicolo and Piercarlo Zanchettin. Leadership cycles in a quality-ladder model of endogeneous growth. *The Economic Journal*, 122:618–50, 2012.

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.

Lucia Foster, John Haltiwanger, 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 and product cycles. *Quarterly Journal of Economics*, 106(2):557–86, 1991a.

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

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

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.

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

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.

Pietro F. Peretto. Cost reduction, entry, and the interdependence of market structure and economic growth. *Journal of Monetary Economics*, 43:173–95, 1999.

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.

Paul S. Segerstrom and James M. Zolnierek. The r and d incentives of industry leaders. *International Economic Review*, 40(3):745–65, 1999.

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.