RECONCILING MODELS OF DIFFUSION AND INNOVATION:
A THEORY OF THE PRODUCTIVITY DISTRIBUTION AND TECHNOLOGY FRONTIER

Jess Benhabib
Jesse Perla
Christopher Tonetti

Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution
and Technology Frontier
Jess Benhabib, Jesse Perla, and Christopher Tonetti
NBER Working Paper No. 23095
January 2017, Revised October 2019
JEL No. O14,O30,O31,O33,O40

ABSTRACT

We study how innovation and technology diffusion interact to endogenously determine the shape
of the productivity distribution and generate aggregate growth. We model firms that choose to
innovate, adopt technology, or produce with their existing technology. Costly adoption creates a
spread between the best and worst technologies concurrently used to produce similar goods. The
balance of adoption and innovation determines the shape of the distribution; innovation stretches
the distribution, while adoption compresses it. On the balanced growth path, the aggregate growth
rate equals the maximum growth rate of innovators. While innovation drives long-run growth,
changes in the adoption environment can influence growth by affecting innovation incentives,
either directly, through licensing of excludable technologies, or indirectly, via the option value of
adoption.

Jess Benhabib
Department of Economics
New York University
19 West 4th Street, 6th Floor
New York, NY 10012
and NBER
jess.benhabib@nyu.edu

Christopher Tonetti
Graduate School of Business
Stanford University
655 Knight Way
Stanford, CA 94305
and NBER
tonetti@stanford.edu

Jesse Perla
University of British Columbia
Vancouver School of Economics
997 - 1873 East Mall
Vancouver, BC, V6T 1Z1
jesse.perla@ubc.ca

A technical appendix is available at http://www.nber.org/data-appendix/w23095

1 Introduction

This paper studies how the interaction between adoption and innovation determines the shape of the productivity distribution, the expansion of the technology frontier, and the aggregate economic growth rate. Empirical estimates of productivity distributions tend to have a large range, with many low-productivity firms and few high-productivity firms within even very narrowly defined industries and products (Syverson (2011)). The economy is filled with firms that produce similar goods using different technologies, and different firms invest in improving their technologies in different ways. Some firms are innovative, bettering themselves while simultaneously pushing out the frontier by creating technologies that are new to the world. There are, however, many firms that purposefully choose instead to adopt already invented ideas. The main contribution of this paper is to develop a model of aggregate growth that delivers a productivity distribution with an endogenous expanding frontier, range, and tail index as the result of optimal firm behavior.

In the model, innovation pushes out the frontier, creating the technologies that will eventually be adopted, and stretches the distribution. Adoption helps to compress the distribution by keeping the laggards from falling too far behind. Beyond affecting the shape of the distribution, both adoption and innovation affect the aggregate growth rate. Long-run growth is driven by innovation, but that does not necessarily mean that adoption cannot affect long-run growth. Rather, it means that adoption affects growth by affecting the incentives to innovate. Changes in the adoption environment can affect innovation incentives either because innovators may one day become adopters or because adopters may directly pay to license technologies from innovators.

Model Overview and Main Results. We first build a simple model of exogenous innovation and growth to focus on how innovation and adoption jointly determine the shape of the productivity distribution. We then add an innovation decision in which aggregate growth is endogenously driven in the long run by the innovation activity of high-productivity firms. At the core of the model are the costs and benefits of adoption and innovation. Firms are heterogeneous in productivity, and a firm's technology is synonymous with its productivity. Adoption is modeled as paying a cost to instantaneously receive a draw of a new technology; this is a model of adoption because the new productivity is drawn from the existing distribution of technologies currently in use for production. To represent innovation, we model firms as being in either a creative or a stagnant innovation state; when creative, innovation generates geometric growth in productivity at a rate increasing in firm-specific innovation expenditures. A firm's innovation state evolves according to a two-state Markov process. This model of stochastic innovation with finite idiosyncratic growth rates is the key technical feature that delivers many of the desired model properties in a tractable fra-

mework. For example, we want the productivity distribution to have finite support so that there are better technologies to be invented. This is in contrast to producing at some given date using all the knowledge that will ever be known. At each point in time, any firm has the ability to innovate or adopt, and firms optimally choose whether and how to improve their productivity. Since adoption is a function of the distribution of available technologies, the productivity distribution is the aggregate state variable that moves over time, and this movement is driven by firms' adoption and innovation activity.

In equilibrium, there will be low-productivity firms investing in adopting technologies; stagnant firms falling back relative to creative firms; medium-productivity creative firms investing small amounts to grow a bit through innovation; and higher-productivity creative firms investing a lot in innovation to push out the productivity frontier. Easy adoption, in the sense of low cost or high likelihood of adopting a very productive technology, tends to compress the productivity distribution, as the low-productivity firms are not left too far behind. A low cost of innovation tends to spread the distribution, as the high-productivity firms can more easily escape from the pack. Thus, the shape of the distribution, which typically looks like a truncated Pareto with finite support, is determined by the relative efficiency of adoption and innovation. The stochastic innovation state ensures that some firms that have bad luck and are uncreative for a stretch of time fall back relative to adopting and innovating firms, generating a non-degenerate stationary distribution with adoption existing in the long run.

Adoption and innovation are not two completely independent processes, with some firms perpetual adopters and some perpetual innovators. Rather, the ability of all firms to invest in both activities generates general equilibrium interactions between actions. A key spillover between adoption and innovation can be seen in the option value of adoption. For high-productivity firms which are far from being low-productivity adopters, the value of having the option to adopt is small. The lower a firm's productivity, the closer it is to being an adopter and, thus, the higher the option value of adoption. The higher the option value of adoption, the lower the incentive to spend on innovating to grow away from entering the adoption region and exercising that option. Thus, the value of adoption, which is determined by the cost of acquiring a new technology and the probability of adopting a good technology, affects incentives to innovate and aggregate growth. Through this channel, the better adoption is, the more tempting it is to free ride on other firms pushing out the frontier by investing less in innovation. We conclude the paper by exploring how the excludability of technology affects the interplay of adoption and innovation. We model adopters as having to pay a fee to the firm whose technology they adopt. Hence there is an additional direct link between adoption behavior and innovation incentives that affects the shape of the distribution and aggregate growth rates. Through this licensing channel, easier adoption leads to more licensing, which increases investments in innovation and aggregate growth.

1.1 Recent Literature

Our paper is closely related conceptually and technically to the idea diffusion literature including Luttmer (2007), Alvarez et al. (2008), Lucas (2009), Alvarez et al. (2013), Perla and Tonetti (2014), and Lucas and Moll (2014). Buera and Lucas (2018) provides a survey of the literature.

In an early paper on technology diffusion Nelson and Phelps (1966) developed a model that specified a differential equation that determined technology diffusion as a function of the distance between the leader and the follower. Lucas (2009) expanded the domain of idea diffusion, in that it allowed ideas to diffuse not just from a leader to followers, but potentially from and to all agents in the economy. The probability of receiving a particular idea depends on the frequency of that idea in the whole population. In Perla and Tonetti (2014) and Lucas and Moll (2014) agents make a choice to invest in technology diffusion. Thus, the amount of diffusion is no longer exogenous (either as an exogenous function of the distance to the frontier or as an exogenous arrival rate of draws from a source distribution), allowing for the study of incentives, externalities, and welfare improving policies. In this paper, we build on these models of idea diffusion, but allow agents to grow not just through diffusion, but also through innovation.

We contribute to the literature that studies both innovation and technology diffusion. Buera and Oberfield (2019) is a related semi-endogenous growth model of the international diffusion of technology and its connection to trade in goods. The authors combine the process of idea diffusion with innovation, in the spirit of Jovanovic and Rob (1989). They model productivity upgrading according to one joint process that mixes innovation and adoption. In contrast, our paper models these as distinct actions potentially undertaken by different firms. Furthermore, their focus is not on the endogenous determination of the shape of the distribution, since it is given exogenously by the distribution from which innovation increments are drawn.[1] The industry evolution model of Jovanovic and MacDonald (1994) features firms that grow from innovation and imitation, but a firm cannot purposely target its investment to innovation or imitation. Benhabib et al. (2014) explicitly model a choice between investing in innovation and adoption, but in a Nelson-Phelps style model in which agents' decisions do not depend on the entire productivity distribution.

Perhaps the most closely related paper to ours is König et al. (2016) (KLZ). There, as in our paper, firms make an optimal choice between two stochastic processes: one that is related to the existing productivity distribution (imitation) and one that is not (innovation). As in our paper, higher-productivity firms choose to focus more on innovation than imi-

---

[1]Eeckhout and Jovanovic (2002) also model technological spillovers that are a function of the distribution of firm productivity. Acemoglu et al. (2007) also model spillovers across firms in innovation, captured by the number of firms that have attempted to implement a technology before.

tation.[2] Both models produce sustained growth through the interaction of innovation and imitation. A key difference is our analysis of a finite relative technology frontier (i.e., the ratio of the maximum to average productivity). This allows us to study the intensive margin of innovation of frontier firms and how the option value of adoption induces a free-riding incentive and slows growth. We also study excludable technology by modeling licensing, which introduces another mechanism through which adoption affects long-run growth.[3] The importance of the tension between endogenous innovation and imitation is emphasized in König et al. (2018), which uses that perspective to analyze the recent transformation of the Chinese economy.

Acemoglu et al. (2006), Chu et al. (2014), and Stokey (2014, 2017) also explore the relationship between innovation and diffusion from different perspectives.[4] We share a similarity with those papers, as there is an advantage to backwardness in the sense of option value from the ability to adopt. The crucial element that enables the interesting trade-off between innovation and technology diffusion in our model is that the incumbents internalize some of the value from the evolving distribution of technologies, thus distorting their innovation choices. That is, incumbent firms not adopting today realize they may adopt in the future, and they derive positive value from this option to adopt.

One question that arises in diffusion models is where do new ideas come from, and how do those new ideas determine the aggregate growth rate.[5] We model long-run growth that occurs through the interaction of innovation and diffusion. In our model with endogenous innovation, the interaction between innovation and adoption is especially interesting when the distribution has a finite frontier. The rate at which frontier firms invest in discovering new ideas determines the long-run aggregate growth rate. When the frontier is finite, the efficiency of adoption affects long-run aggregate growth in part because frontier firms realize

---

[2]There are a few other differences to consider when comparing to König et al. (2016). KLZ assume no cost of either innovation or imitation and use a limit to firms' absorptive capacity to induce a tradeoff between innovation and adoption, whereas in our paper there is a cost of innovation and a cost of imitation. Furthermore, as the arrival rate of imitation is not immediate in KLZ, they have an asymptotically power-law left tail. Our model has a sweeping barrier, which results from the limit of a rapid imitation rate. Finally, the monopolistic competition and differentiated goods KLZ model provides an economic foundation for a profit function that increases in productivity, which we specify exogenously.

[3]The closest paper to ours in terms of modeling licensing is Shi and Hopenhayn (2017), which provides a search-theoretic framework for analyzing bargaining over technology transfers in an environment with congestion externalities.

[4]An alternative line of literature studies the diffusion of technology from incumbents to entrants, as in Luttmer (2007), Sampson (2015), and Lashkari (2016).

[5]Using an initial distribution with infinite support, such as in Perla and Tonetti (2014) and the baseline model of Lucas and Moll (2014), may provide a good approximation for the contribution of adoption to growth in the medium run, even absent a theory of innovation that pushes out a frontier. See Figure 2 of Perla and Tonetti (2014) and Figure 9 of Buera and Lucas (2018).

they may become adopters in the future.

Staley (2011) and Luttmer (2012a) explore the interaction of innovation and technology adoption in a model in which both are exogenous processes.[6] Their approach allows a formal analysis with a finite number of firms using the well-understood KPP-Fisher equation, which allows for a sharp characterization of model properties. In our paper we explore how adoption and innovation choices respond to economic incentives and generate endogenous growth—in an environment where a finite number of firms would be computationally infeasible.

2 Baseline Model with Exogenous Stochastic Innovation

We first analyze an exogenous growth model to simplify the introduction of the environment and to focus on the economic forces that determine the shape of the stationary normalized productivity distribution. For the exogenously defined innovation process, we choose a specification that leads to a meaningful concept of the technology frontier in a framework amenable to endogenous investment in growth.

The only choice that a firm makes in this version of the model is whether to adopt a new technology or to continue producing with its existing technology. In Section 4, we develop the full version of this model—in which a firm chooses its innovation rate—to study how adoption and innovation activities interact to jointly determine the shape of the productivity distribution and the aggregate growth rate.

2.1 The Baseline Model

Firm Heterogeneity. A continuum of firms produce a homogeneous product and are heterogeneous over their productivity, Z, and innovation ability, $i \in \{\ell, h\}$. For simplicity, firm output equals firm profits equals firm productivity. The measure of firms of productivity less than Z in innovation state i at time t is denoted by $\Phi_i(t, Z)$. The maximum productivity of any firm, $\bar{Z}(t) \equiv \sup \{\text{support} \{\Phi_\ell(t, \cdot)\} \cup \text{support} \{\Phi_h(t, \cdot)\}\}$ is interpreted as the technology frontier. There is a unit measure of firms, so that $\Phi_\ell(t, \bar{Z}(t)) + \Phi_h(t, \bar{Z}(t)) = 1$. At any point in time, the minimum of the support of the distribution will be endogenously determined, denoted by $M_i(t)$, so that $\Phi_i(t, M_i(t)) = 0$. Define the distribution unconditional on type as $\Phi(t, Z) \equiv \Phi_\ell(t, Z) + \Phi_h(t, Z)$.

A firm with productivity Z can choose to continue producing with its existing technology,

---

[6]Luttmer (2012a, 2015a,b) provide careful analysis of the role of hysteresis including the important interaction of the stochastic innovation process with initial conditions and how models with a continuum of agents or states approximate the world that is populated by a finite number of agents.

in which case it would grow stochastically according to the exogenous innovation process, or it can choose to adopt a new technology.

Stochastic Process for Innovation. In the high innovation ability state (h), a firm is innovating and its productivity is growing at an exogenous rate γ (for now, but a choice made by firms in Section 4). In the low innovation ability state (ℓ), it has zero productivity growth from innovation (without loss of generality). Sometimes, firms have good ideas or projects that generate growth, and sometimes firms are just producing using their existing technology. We model innovation according to this continuous-time two-state Markov process because it allows for the existence of balanced growth path equilibria with finite-support productivity distributions in which adoption persists in the long run and growth is driven by the innovation choices of frontier firms. Persistence of the innovation state is modeled primarily for technical reasons related to continuous time rather than being of direct economic interest. Thus, we will perform our numerical exercises calibrating the model with high transition rates. Loosely speaking, shocks to the innovation state are like IID growth rate shocks that avoid continuous-time measurability issues with IID stochastic processes.

The jump intensity from low to high is $\lambda_\ell > 0$ and from high to low is $\lambda_h > 0$. Since the Markov chain has no absorbing states, and there is a strictly positive flow between the states for all Z, the support of the distribution conditional on ℓ or h is the same. With support $\{\Phi(t, \cdot)\} \equiv [M(t), \bar{Z}(t)]$, define the growth rates of the lower and upper bounds as $g(t) \equiv M'(t)/M(t)$ and $g_{\bar{Z}}(t) \equiv \bar{Z}'(t)/\bar{Z}(t)$.

For notational simplicity, define the differential operator $\boldsymbol{\partial}$ such that $\boldsymbol{\partial}_z \equiv \frac{\partial}{\partial z}$ and $\boldsymbol{\partial}_{zz} \equiv \frac{\partial^2}{\partial z^2}$. When a function is univariate, derivatives will be denoted by $M'(z) \equiv \boldsymbol{\partial}_z M(z) \equiv \frac{\mathrm{d}M(z)}{\mathrm{d}z}$.

Adoption and Technology Diffusion. A firm has the option to adopt a new technology by paying a cost. Adoption means switching production practice by changing to a technology that some other firm is using. We model this adoption process as undirected search across firms, as in Perla and Tonetti (2014) and Lucas and Moll (2014).

For simplicity, we model the adopting firm as drawing a new productivity, Z, from the unconditional distribution, $\Phi(t, Z)$, and starting in the low-innovation state, ℓ. That is, adopters cannot innovate immediately after adoption.[7]

The cost of adoption grows as the economy grows, and it is parameterized by ζ. The endogenous scale of the economy can be summarized by the minimum of the support of the productivity distribution $M(t)$. Thus, we model the cost of adoption as $\zeta M(t)$.

---

[7]Model variations on conditional vs. unconditional draws and maintaining or switching innovation types have few quantitative or qualitative consequences, especially given the high switching rates between states used in our numerical examples. The model setup here keeps the formulas cleaner since adopters only show up as a source in the Kolmogorov forward equation for the ℓ-type distribution, which saves on notation.

Leapfrogging to the Frontier. Finally, firms can leapfrog to the frontier of the productivity distribution with arrival rate $\eta > 0$. Innovators jumping to the frontier represents some chance that the innovation process generates a big insight, instead of steady incremental progress. Different from autarkic innovation that generates γ-proportional growth through process improvement, leapfrogging may be viewed as a melding of innovation and diffusion since the jump is a function of the existing productivity distribution. Adopters jumping to the frontier captures that sometimes adopters get lucky and their search for a new technology finds the best one available.

For tractability, we model such a jump as temporarily disruptive to innovation, such that all leapfrogging firms become ℓ-types and must wait for the Markov transition to h before they become innovators again.[8]

Firm Value Functions. Firms discount at rate $r > 0$. Let $V_i(t, Z)$ be the continuation value function for type i—i.e., the value at time t of being an i-type firm and producing with productivity Z.

$$rV_\ell(t, Z) = Z + \underbrace{\lambda_\ell \left(V_h(t, Z) - V_\ell(t, Z)\right)}_{\text{Switch to } h} + \underbrace{\eta(V_\ell(\bar{Z}) - V_\ell(Z))}_{\text{Jump to Frontier}} + \underbrace{\partial_t V_\ell(t, Z)}_{\text{Capital Gains}} \tag{1}$$

$$rV_h(t, Z) = Z + \underbrace{\gamma Z \partial_Z V_h(t, Z)}_{\text{Exogenous Innovation}} + \underbrace{\lambda_h \left(V_\ell(t, Z) - V_h(t, Z)\right)}_{\text{Switch to } \ell} + \eta(V_\ell(\bar{Z}) - V_h(Z)) + \partial_t V_h(t, Z). \tag{2}$$

A firm's continuation value derives from instantaneous production plus capital gains as well as productivity growth if in the high innovation ability state, and it accounts for the intensity of jumps between innovation abilities i and jumps to the frontier.

The value of adopting is the continuation value of an ℓ-type firm having a new productivity drawn from $\Phi(t, Z)$ less the cost of adoption:

$$\text{Net Value of Adoption} = \underbrace{\int_{M(t)}^{\bar{Z}(t)} V_\ell(t, Z) \, d\Phi(t, Z)}_{\text{Gross Adoption Value}} - \underbrace{\zeta M(t)}_{\text{Adoption Cost}}. \tag{3}$$

The Optimal Adoption Policy and the Minimum of the Support of the Productivity Distribution. The optimal firm policy is given by a threshold rule $M_a(t)$ such that all firms with productivity $Z \le M_a(t)$ will choose to adopt and firms with $Z > M_a(t)$ will

---

[8]The assumption that all leapfroggers switch to the ℓ state is purely for analytical convenience and can be changed without introducing qualitative differences. Similarly, the jumps to the frontier could occur exclusively for adopting firms instead of operating firms with no qualitative differences.

not adopt. Since the value of continuing is increasing in Z, and the net value of adopting is independent of Z, the firm's optimal adoption policy takes the form of a reservation productivity rule. While the adoption threshold could conceivably depend on the innovation type i, see Technical Appendix A.4 for a proof showing that ℓ- and h-type firms choose the same threshold, $M_a(t)$, if the net value of adoption is independent of the current innovation type.

As draws are instantaneous, for any $t > 0$ this endogenous $M_a(t)$ becomes the evolving minimum of the $\Phi_i(t, Z)$ distributions, $M(t)$, and in a slight abuse of notation, we will refer to both the minimum of the support and the firm adoption policy as $M(t)$ going forward.[9]

In principle, there may be adopters of either innovation type with productivity in the common adoption region $Z \leq M(t)$. Define $S_i(t) \geq 0$ as the flow of i-type firms entering the adoption region at time t and denote the total flow of adopting firms as $S(t) \equiv S_\ell(t) + S_h(t)$.

The Firm Problem. A firm's decision problem can be described as choosing an optimal stopping time of when to adopt. Equivalently, it can be described as a free boundary problem, choosing the productivity level at which to adopt. Necessary conditions for the free boundary problem include the continuation value functions and, at the endogenously chosen adoption boundary $M(t)$, value matching conditions,

$$\underbrace{V_i(t, M(t))}_{\text{Value at Threshold}} = \underbrace{\int_{M(t)}^{\bar{Z}(t)} V_\ell(t, Z) \mathrm{d}\hat{\Phi}_\ell(t, Z) + \int_{M(t)}^{\bar{Z}(t)} V_h(t, Z) \mathrm{d}\hat{\Phi}_h(t, Z)}_{\text{Gross Adoption Value}} - \underbrace{\zeta M(t)}_{\text{Adoption Cost}}, \qquad (4)$$

and smooth-pasting conditions,

$$\boldsymbol{\partial}_Z V_\ell(t, M(t)) = 0 \quad \text{if } M'(t) > 0 \qquad (5)$$

$$\boldsymbol{\partial}_Z V_h(t, M(t)) = 0 \quad \text{if } M'(t) - \gamma M(t) > 0. \qquad (6)$$

Value matching states that at the optimal adoption reservation productivity, a firm must be indifferent between producing with the reservation productivity and adopting a new productivity. Smooth pasting is a technical requirement that can be interpreted as an intertemporal no-arbitrage condition—necessary only if firms at the boundary are moving backwards relative to the boundary over time.

---

[9]To see why the minimum of the support is the endogenous threshold, consider instantaneous adoption as the limit of the Poisson arrival rate of draw opportunities approaching infinity. In any positive time interval, firms wishing to adopt would gain an acceptable draw with probability approaching 1, so that $Z > M(t)$ almost surely. A heuristic derivation of this limit is given in Technical Appendix A.5.

The Technology Frontier. Given the adoption and innovation processes, if $\bar{Z}(0) < \infty$, then $\bar{Z}(t)$ will remain finite for all t, as it evolves from the innovation of firms in the interval infinitesimally close to $\bar{Z}(t)$. Furthermore, with a continuum of firms the frontier grows at rate γ since there is some firm arbitrarily close to the threshold and growing at rate γ.

Lemma 1 (Growth Rate of the Finite Frontier). *If $\bar{Z}(0) < \infty$, then $\bar{Z}'(t)/\bar{Z}(t) = \gamma$.*

If $\left[\Phi_h(t, \bar{Z}(t)) - \Phi_h(t, \bar{Z}(t) - \epsilon)\right] > 0$ for all $\epsilon > 0$, then there are always firms arbitrarily close to the frontier. If some of them are type h, then they will push the frontier out at rate γ. Given that there are a continuum of firms and that the arrival rate of changes in type i is a memoryless Poisson process, for all finite t there will always be some h-type firms that are arbitrarily close to the frontier and have never jumped to the low state, so the growth rate of the frontier is always γ.

Law of Motion of the Productivity Distribution. The Kolmogorov forward equation (KFE) describes the evolution of the productivity distribution for productivities above the minimum of the support. The KFEs in the CDFs for ℓ- and h-type firms are

$$\boldsymbol{\partial}_t \Phi_\ell(t, Z) = \underbrace{-\lambda_\ell \Phi_\ell(t, Z) + \lambda_h \Phi_h(t, Z)}_{\text{Net Flow from Jumps}} \underbrace{- \eta \Phi_\ell(t, Z) + \eta \mathbb{H}\left(Z - \bar{Z}(t)\right)}_{\text{Leapfroggers}} \tag{7}$$

$$+ \underbrace{(S_\ell(t) + S_h(t))}_{\text{Flow of Adopters}} \underbrace{\Phi(t, Z)}_{\text{Draw} \le Z} - \underbrace{S_\ell(t)}_{\ell\text{-Adopters}}$$

$$\boldsymbol{\partial}_t \Phi_h(t, Z) = \underbrace{-\gamma Z \boldsymbol{\partial}_Z \Phi_h(t, Z)}_{\text{Innovation}} - \lambda_h \Phi_h(t, Z) + \lambda_\ell \Phi_\ell(t, Z) - \eta \Phi_h(t, Z) - S_h(t), \tag{8}$$

where $\mathbb{H}\left(\cdot\right)$ is the Heaviside operator. For each type i, the KFEs keep track of inflows and outflows of firms with a productivity level at or below Z. An i-type firm with productivity less than or equal to Z stops being in the i-distribution at or below Z if it keeps its type and increases its productivity above Z or if it changes its type.

A firm can keep its type and increase its Z in three ways: adoption, innovation, or leapfrogging. An adopting firm has probability $\Phi(t, Z)$ of becoming type ℓ and drawing a productivity less than or equal to Z and the number of firms adopting is $(S_\ell(t) + S_h(t))$. Hence $(S_\ell(t) + S_h(t))\Phi(t, Z)$ is added to the ℓ-distribution. Additionally, the flow of adopters of type i, $S_i(t)$, is subtracted from the corresponding distribution (this term appears conditional on any Z because adoption occurs at the minimum of the support). Intuitively, the adoption reservation productivity acts as an absorbing barrier sweeping through the distribution from below. As it moves forward, it collects adopters at the minimum of the support, removes them from the distribution, and inserts them back into the distribution according to Φ. Recognizing that the jumps in type at intensity λ_i are of measure 0 when calculating the

9

number of firms that cross the boundary in any infinitesimal time period, the flow of adopters comes from the flux across the moving boundary $M(t)$.

The KFE for the h-types has a term that subtracts the firms that grow above Z through innovation: there are $\partial_Z \Phi_h(t, Z)$ number of h-type firms at productivity Z, and because innovation is geometric, they grow above Z at rate γZ.

Firms of productivity Z switching from type i to type i' leave the i-distribution and enter the i'-distribution at rate λ_i. For example, the number of ℓ-type firms with productivity less than or equal to Z is $\Phi_\ell(t, Z)$, and they leave the ℓ-distribution at rate λ_ℓ and enter the h-distribution at the same rate.

Firms jump to the frontier at rate η, so they are subtracted from the CDF. For analytical tractability, all leapfrogging firms become type ℓ, so they are added to the ℓ-distribution at $\bar{Z}(t)$.

The firms are owned by a representative consumer who values the undifferentiated good with log utility and a discount rate $\rho > 0$. If the growth rate on the balanced growth path is g, then the interest rate faced by firms is $r = \rho + g$.

2.2   Normalization, Stationarity, and Balanced Growth Paths

In this paper, we study economies on balanced growth paths (BGPs), in which the distribution is stationary when properly rescaled and aggregate output grows at a constant rate. The economy is characterized by a system of equations defining the firm problem, the laws of motion of the productivity distributions, and consistency conditions that link firm behavior and the evolution of the distributions. To compute BGP equilibria, it is convenient to transform this system to a set of stationary equations. While we could normalize by any variable that grows at the same rate as the economy, it is convenient to normalize variables relative to the endogenous boundary $M(t)$. Define the change of variables, normalized distributions, and normalized value functions as

$$z \equiv \log(Z/M(t)) \tag{9}$$

$$F_i(t, z) = F_i(t, \log(Z/M(t))) \equiv \Phi_i(t, Z) \tag{10}$$

$$v_i(t, z) = v_i(t, \log(Z/M(t))) \equiv \frac{V_i(t, Z)}{M(t)}. \tag{11}$$

The adoption threshold is normalized to $\log(M(t)/M(t)) = 0$, and the relative technology frontier is $\bar{z}(t) \equiv \log(\bar{Z}(t)/M(t))$. See Figure 1 for an illustration of the normalized and unnormalized distributions. Define the normalized unconditional distribution as $F(z) \equiv F_\ell(z) + F_h(z)$, which is a valid CDF (i.e., $F(t) = 0$ and $F(\bar{z}(t)) = 1$).

With the above normalizations, it is possible that the value functions, productivity distributions, and growth rates are stationary—i.e., independent of time.
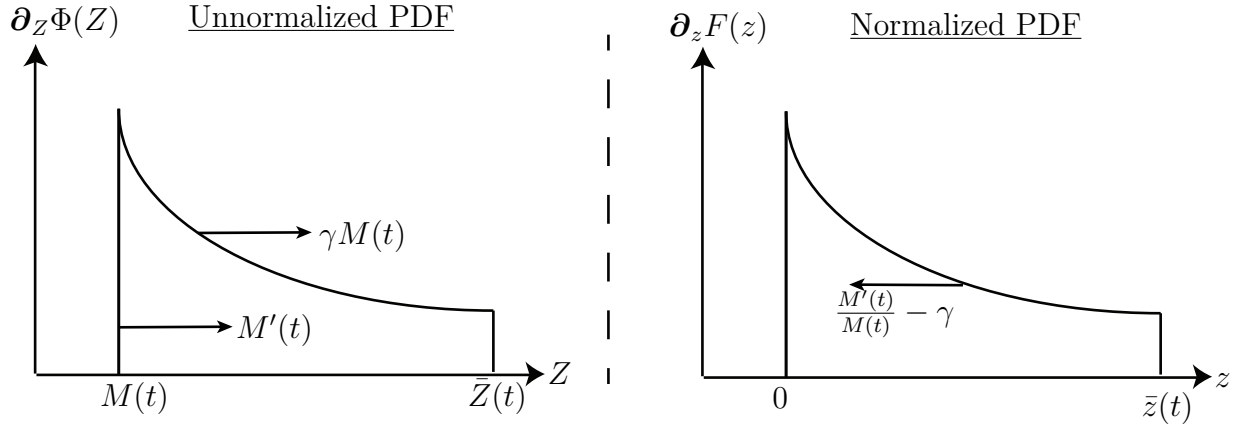
Figure 1: Normalized and Unnormalized Distributions

Summary of Stationary KFEs and Firm Problem. A full derivation of the normalized system is given in Technical Appendix A.1. Here, we summarize the resulting equations that characterize the laws of motion for the normalized productivity distributions and the normalized firm problem. The equations that determine the stationary productivity distributions are:

$$0 = gF_\ell'(z) + \lambda_h F_h(z) - \lambda_\ell F_\ell(z) - \eta F_\ell(z) + \eta \mathbb{H}\,(z - \bar{z}) + SF(z) - S \tag{12}$$

$$0 = \lambda_\ell F_\ell(z) - \lambda_h F_h(z) - \eta F_h(z) \tag{13}$$

$$0 = F_\ell(0) = F_h(0) \tag{14}$$

$$1 = F_\ell(\bar{z}) + F_h(\bar{z}) \tag{15}$$

$$S_\ell = gF_\ell'(0) \text{ if } g > 0 \tag{16}$$

$$S_h = (g - \gamma)\,F_h'(0) \text{ if } g > \gamma. \tag{17}$$

The necessary conditions of the normalized firm problem are:

$$\rho v_\ell(z) = e^z - gv_\ell'(z) + \lambda_\ell(v_h(z) - v_\ell(z))) + \eta(v_\ell(\bar{z}) - v_\ell(z)) \tag{18}$$

$$\rho v_h(z) = e^z + \lambda_h(v_\ell(z) - v_h(z)) + \eta(v_\ell(\bar{z}) - v_h(z)) \tag{19}$$

$$v(0) = \frac{1}{\rho} = \int_0^{\bar{z}} v_\ell(z)\,\mathrm{d}F(z) - \zeta \tag{20}$$

$$v_\ell'(0) = 0 \text{ if } g > 0 \tag{21}$$

$$v_h'(0) = 0 \text{ if } g > \gamma. \tag{22}$$

Given that the two types of firms choose the same adoption threshold, we drop the type index for the value functions at the adoption threshold: $v(0) \equiv v_i(0)$.

Equations (12) to (15) are the stationary KFEs with boundary values. Recall that g is

the growth rate of the minimum of the support and γ is the innovation growth rate. In the normalized setup, firms are moving backwards toward the constant minimum of the support, and their growth rate determines the speed at which they are falling back.

Equations (18) and (19) are the Bellman equations in the continuation region and (20) is the value-matching condition between the continuation and technology adoption regions. The smooth-pasting conditions given in equations (21) and (22) are necessary only if the firms of that particular type are drifting backwards relative to the adoption threshold.

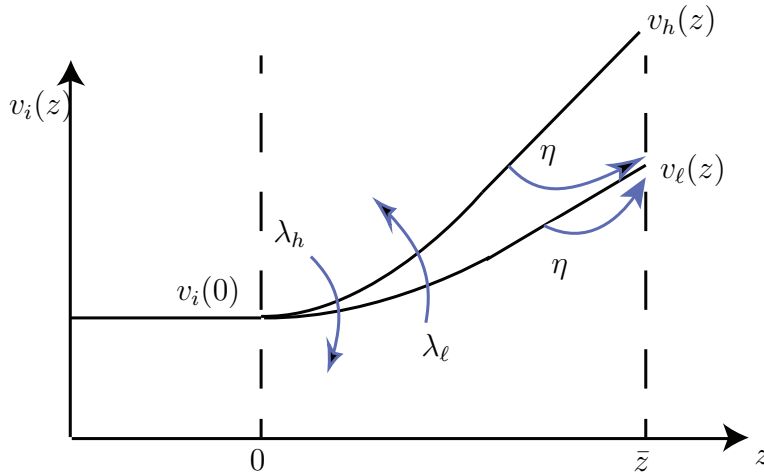See Figure 2 for a visualization of the normalized Bellman equations.

Figure 2: Normalized Stationary Value Functions

Definition 1 (Recursive Competitive Equilibrium with Exogenous Innovation). *A recursive competitive equilibrium with exogenous innovation consists of initial distributions $\Phi_i(0, z)$, adoption reservation productivity functions $M_i(t)$, value functions $V_i(t, z)$, interest rates $r(t)$, and sequences of productivity distributions $\Phi_i(t, z)$ such that the following hold:*

1. *Given $r(t)$ and $\Phi_i(t, z)$, $M_i(t)$ are the optimal adoption reservation productivities, with $V_i(t, z)$ the associated value functions.*

2. *Given $M_i(t)$ and $\Phi_i(t, z)$, $r(t)$ is consistent with the consumer's intertemporal marginal rate of substitution.*

3. *Given $M_i(t)$, $\Phi_i(t, z)$ fulfill the laws of motion in (7) and (8) subject to the initial condition $\Phi_i(0, z)$.*

We restrict our interest to equilibria that are balanced growth paths.

Definition 2 (Balanced Growth Path Equilibrium with Exogenous Innovation). *A balanced growth path equilibrium with exogenous innovation is a recursive competitive equilibrium*

12

such that the growth rate of aggregate output is constant and the normalized productivity distributions are stationary. This is equivalent to requiring that $F_i(t, z) = F_i(z)$ and $g(t) \equiv M'(t)/M(t) = g$ for all t.

Define the growth rate of aggregate output as $g_E(t) \equiv \partial_t \mathbb{E}_t[Z] / \mathbb{E}_t[Z]$. Then Lemma 2 shows that $g = g_E$ is a requirement on a BGP.

Lemma 2 (Growth of the Endogenous Adoption Threshold and Aggregate Output). *On a balanced growth path, the growth rate of the endogenous threshold, $M(t)$, must be the same as the growth rate of aggregate output. That is, $g = g_E$. Furthermore, if $\bar{Z}(0) < \infty$ and $\eta > 0$, then $g = g_{\bar{Z}} = \gamma$.*

Proof. The value-matching condition in equation (20) is normalized to the endogenous threshold, hence adoption has a constant and strictly positive cost. Thus, as $v(z) \geq z$, if the expected value of a draw from the technology distribution were not stationary relative to the adoption cost, then value matching could not hold with equality for all t. For finite frontier economies, the condition $g \leq g_{\bar{Z}}$ must be satisfied; otherwise the minimum of the support of the distribution would eventually be strictly greater than the maximum of the support. See Technical Appendix A.3 for a similar, but more formal, exposition. □

Since $g \leq \gamma$, then on a BGP with exogenous innovation no h-type agents cross the boundary. Thus, S in equation (16) is the flow of ℓ-type agents moving backwards at a relative speed of g across the adoption barrier.
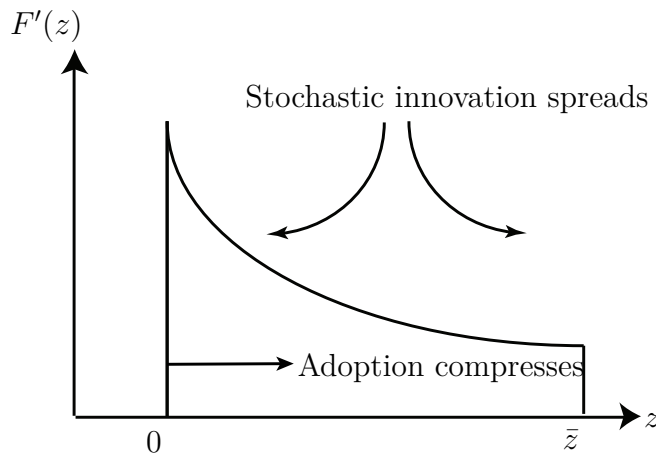

Figure 3: Tension between Stochastic Innovation and Adoption

How Adoption and Innovation Generate a Stationary Normalized Distribution.
Figure 3 provides some intuition on how proportional growth and adoption can create a stationary distribution. Without endogenous adoption, nothing can prevent the proportional

random shocks from spreading out the distribution, driving the variance to infinity.[10] However, when adoption is endogenous, as the distribution spreads, the incentives to adopt a new technology increase, and the adoption decisions of low-productivity agents then act to compress the distribution. In a BGP equilibrium, technology diffusion can balance innovation, thus allowing for a stationary normalized distribution.

The existence of a stationary normalized distribution immediately places restrictions on the relationship between the growth of the frontier and adoption behavior. Recall that g is the growth rate of the endogenous reservation adoption productivity, and thus also the growth rate of the minimum of the support of the unnormalized distribution. As Lemma 2 states, a necessary condition for the existence of a stationary normalized distribution with a finite frontier (i.e., if $\bar{Z}(t) < \infty$ for all t) is that $g = g_E = g_{\bar{Z}} = \gamma$. That is, the minimum of the support, which must grow at the same rate as the aggregate economy, also must grow at the same rate as the frontier. This may make it seem as though the growth rate of the minimum of the support, which is determined by the adopters, determines the long-run growth rate, but it should be interpreted as an equilibrium relationship between adoption and innovation. The flow of adopters endogenously increases or decreases, and adjusts the growth rate $g(t)$ until it is in balance with the growth rate of the frontier, γ.

3 How Adoption and Innovation Interact to Shape the Productivity Distribution

In this section, we compute BGP equilibria for economies with finite-support productivity distributions. There are two main questions that motivate our analysis. First, how do adoption and innovation determine the shape of the productivity distribution? Second, what is the aggregate growth rate, and how is it affected by parameters related to adoption and innovation?

To set the stage, consider Perla and Tonetti (2014), which is essentially a discrete-time version of the economy in this paper with $\eta = \gamma = 0$ and is similar to cases in the literature such as Lucas and Moll (2014), as discussed in Buera and Lucas (2018). In Perla and Tonetti (2014) a BGP equilibrium with strictly positive growth exists only for distributions with power-law tails; the long-run growth rate is a function of the shape of the initial distribution and the cost of adoption. In contrast, here the shape of the stationary distribution is endogenously a function of parameters related to innovation and adoption, while the long-run aggregate growth rate is the maximum rate of growth of innovators and is independent of initial conditions. In this section, the growth rate of innovators is exogenous

---

[10]König et al. (2016) provide a similar intuition in their Proposition 2, which shows the expansion of the distribution in the absence of imitation.

(γ), but Section 4 develops the endogenous growth version of the model in which innovation investment is a choice made by firms.

3.1 The Stationary Finite Support BGP

Define the constants $\hat{\lambda} \equiv \frac{\lambda_\ell}{\eta + \lambda_h}$, $\bar{\lambda} \equiv \frac{\rho + \lambda_\ell + \lambda_h}{\rho + \lambda_h}$, and $\nu = \frac{\rho + \eta}{\gamma} \bar{\lambda}$, and let $\eta > 0$ throughout.

Proposition 1 (Stationary Equilibrium for $\eta > 0$). *Given parameter restrictions, a unique equilibrium with $\bar{z} < \infty$ exists with $g = \gamma$. The stationary distribution is*

$$F_\ell(z) = \frac{F'_\ell(0)}{(F'_\ell(0) - \eta/\gamma)(1 + \hat{\lambda})}(1 - e^{-\alpha z}) \tag{23}$$

$$F_h(z) = \hat{\lambda} F_\ell(z), \tag{24}$$

with

$$\alpha \equiv (1 + \hat{\lambda})(F'_\ell(0) - \eta/\gamma) \tag{25}$$

$$\bar{z} = \frac{\log(\gamma F'_\ell(0)/\eta)}{\alpha}. \tag{26}$$

The equilibrium $F'_\ell(0)$ solves the following implicit equation (substituting for α and $\bar{z}$),

$$\zeta + \frac{1}{\rho} = \frac{\gamma F'_\ell(0)\alpha\bar{\lambda}\left(-\frac{e^{-\nu\bar{z}}\left(-1 + e^{-\alpha\bar{z}}\right)\eta}{\rho\alpha\nu} + \frac{e^{\bar{z}}\eta\left(e^{-\alpha\bar{z}} - 1\right)}{-\alpha\rho} + \frac{-e^{-(\alpha+\nu)\bar{z}} + 1}{\nu(\alpha+\nu)} + \frac{-e^{\bar{z}-\alpha\bar{z}} + 1}{\alpha - 1}\right)}{\gamma(\gamma F'_\ell(0) - \eta)(\nu + 1)}. \tag{27}$$

The firm value functions are

$$v_\ell(z) = \frac{\bar{\lambda}}{\gamma(1 + \nu)}\left(e^z + \frac{1}{\nu}e^{-\nu z} + \frac{\eta}{\rho}\left(e^{\bar{z}} + \frac{1}{\nu}e^{-\nu\bar{z}}\right)\right) \tag{28}$$

$$v_h(z) = \frac{e^z + (\lambda_h - \eta)v_\ell(z) + \eta v_\ell(\bar{z})}{\rho + \lambda_h}. \tag{29}$$

Proof. See Technical Appendix A.2. Parameter restrictions include those such that the $F'_\ell(0)$ that solves equation (27) is larger than η/γ. A proof of uniqueness, showing that $g = \gamma$ is the only equilibrium to exist, is in Technical Appendix A.3. $\square$

Existence, Uniqueness, and Aggregate Growth. If the initial distribution has finite support, then the stationary distribution will have finite support because the maximum growth rate of a firm is finite. The first immediate result is that the equilibrium is unique for any $\eta > 0$. That is, there is a unique stationary distribution that is independent of initial conditions. Furthermore, the aggregate growth rate equals the growth rate of innovators: $g = \gamma$.

Lemma 2 also stated that the aggregate growth rate equals the growth rate of the minimum of the support of the distribution ($g = g_E$), which evolves according to adoption. How could the aggregate growth rate be driven by adoption but also equal the exogenous growth rate of innovation? It is because adoption and innovation are linked in general equilibrium. Adoption is endogenous, and firms invest in adoption to keep up with the frontier. If the frontier were to grow faster, there would be more good new ideas arriving faster, which would make it worthwhile for firms to invest more in adopting those ideas faster. Thus, endogenous adoption is the key that links the minimum of the support of the distribution to the frontier.

Economic Intuition for Firm Behavior that Generates a Stationary Distribution. Leapfrogging to the frontier ($\eta > 0$) plays an important role in delivering a stationary distribution.[11] Leapfrogging prevents the frontier from escaping from the rest of the distribution by ensuring that frontier technologies never have zero probability of being adopted.

The Bellman equations (28) and (29) are the sum of three components: the value of production, the option value of adoption, and the value of jumping to the frontier. That is, in addition to the value of production with the current z modified by time discounting and the probabilities of switching i-type, the value function accounts for changing z through adoption and, in particular, the chance of jumping to the frontier.

Leapfrogging to the frontier by a positive mass of agents can contain the escape in relative productivities by lucky firms that get streaks of long sojourns in the high-growth state h. As they eventually lose their innovative ability and become ℓ-types, they will be overtaken by others that leapfrog to the frontier from within the productivity distribution and replenish it. This leapfrogging/quality ladder process prevents laggards from remaining laggards forever.

Additionally, because the ratio of the frontier to the minimum of the support (the relative frontier, $\bar{z}(t)$) is finite, in the long run, frontier firms still place positive value on the option to adopt. This means that increases in the value of adoption, whether associated with lower costs or higher benefits of adoption, will affect the value of frontier firms. Foreshadowing: In the endogenous growth environment, when γ is a choice, changes in the value of adoption will influence innovation behavior at the frontier, which will affect aggregate growth rates.

Shape. The productivity distribution has an endogenous truncated-tail index, α, that represents the shape of the productivity distribution. Furthermore, there is an additional shape parameter of the distribution: the range of the productivity distribution, given by the max-min ratio $\bar{z}$. Model primitives such as the cost of adoption and the rate of innovation affect both the tail index (α) and how much better the best firm can be relative to the worst

---

[11]If $\eta = 0$, then there is hysteresis and the distribution is stationary only asymptotically. That is, even though $\bar{z}(t) < \infty$ for all t, $\lim_{t \to \infty} \bar{z}(t) = \infty$.

firm in the economy ($\bar{z}$). Because $\bar{z}$ is a finite constant, meaning that the $F_i(z)$ have finite support, α is better interpreted as the shape parameter of a right-truncated power law.

Before moving to the endogenous growth case in the next section, we compute an equilibrium with calibrated parameter values and use comparative statics to illustrate properties of the economy. This analysis will show how the cost of adoption and the innovation growth rate affect the shape of the distribution.

We choose parameters to demonstrate model forces, and explore the relevant region of the state space.[12] The resulting parameterization is $\gamma = 0.02$, $\rho = 0.01$, $\lambda_\ell = 0.533$, $\lambda_h = 1.128$, $\zeta = 25.18$, and $\eta = 0.00098$.

First, as shown in Figure 4, v_ℓ and v_h are very similar because the calibrated λ_i are large, and thus the extra benefit of being in the high state or the relative pain from being in the low state does not last very long. Second, the distributions F_i are power-law shaped, with many low-productivity firms and few high-productivity firms, but they are truncated at the $\bar{z}$ relative frontier. In this calibration, there are also fewer h firms than ℓ firms at all productivities, but those masses are determined largely by the ratio λ_h/λ_ℓ.
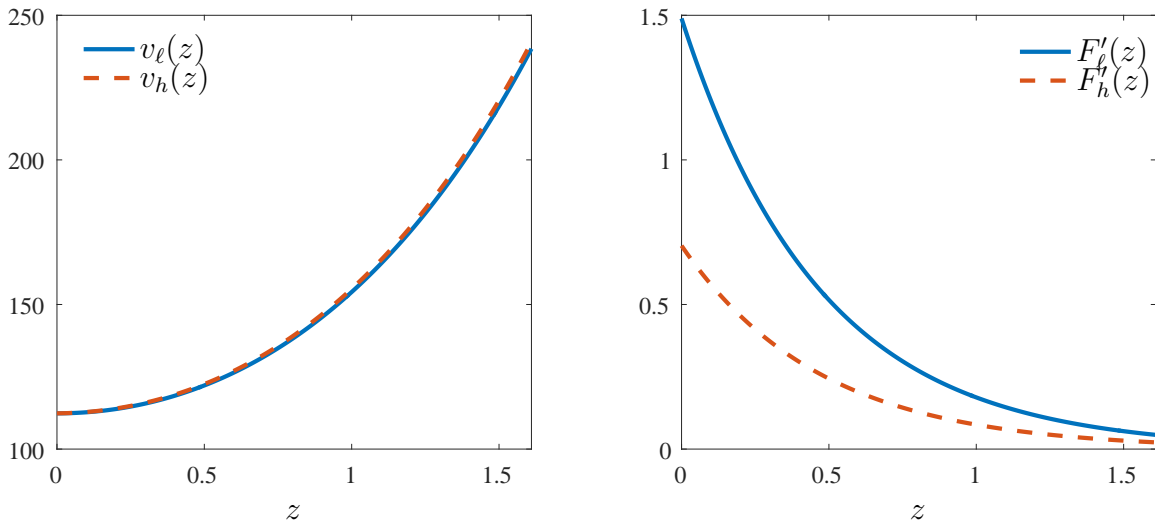

Figure 4: Exogenous $v_i(z)$ and $F_i'(z)$

Comparative statics on how changes in η, γ, ζ, and λ_h affect $\bar{z}$ and α are shown in Figure 5. Easier innovation, in the sense of a higher growth rate for innovators, spreads out the distribution, creating a more distant technology frontier and a thicker tail. Easier leapfrogging, in the sense of a higher probability of jumps to the frontier, also generates thicker tails but generates less of a productivity gap between the best and worst firms. Easier adoption,

---

[12]Details of the simple "calibration" used in our numerical exercises are given in Technical Appendix C. We pick parameters to roughly match firm growth rates, the firm size distribution, the aggregate growth rate, and the risk free interest rate.

captured by lower costs ζ, compresses the distribution, shrinking the relative frontier and thinning the tail.
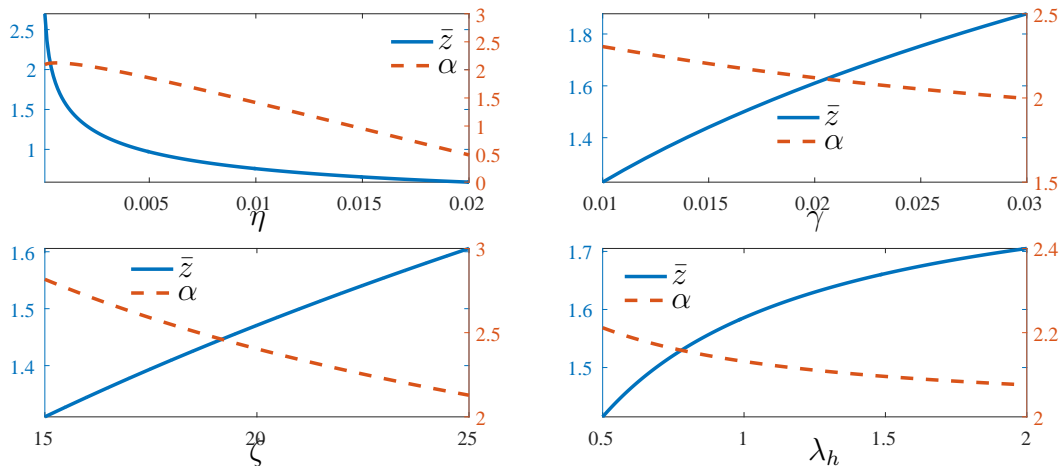

Figure 5: Comparative Statics for Exogenous Innovation

For small η this distribution has a power-law right tail. The abrupt minimum of support in our model comes from the immediacy of technology adoption; with a stochastic arrival as in König et al. (2016) we would also have an asymptoptically power-law left tail.

4 How Adoption and Innovation Interact to Determine the Long-run Aggregate Growth Rate

This section introduces endogenous investments in innovation. In the previous section, the growth rate of the aggregate economy equaled the growth rate of innovation at the finite frontier, $g = \gamma$, but the growth rate of innovation was exogenous. With endogenous innovation, there will be an analogous result ($g = \gamma(\bar{z})$), where both the innovation rate $\gamma(\cdot)$ and the frontier $\bar{z}$ are endogenously (and jointly) determined. Hence we are now in position to discuss how adoption affects the choice of innovation at the frontier, and thus the aggregate growth rate.

4.1 Model with Endogenous Innovation and Excludability

We model firms that can control the drift of their innovation process, as in Atkeson and Burstein (2010) and Stokey (2014).[13] We will focus on two main cases that highlight how adoption activity can affect long-run growth. In both cases, the long-run growth rate is

---

[13]An alternative approach, as taken in Section 6 of König et al. (2016), would be to have the firm choose between adoption and innovation given a capacity.

determined by the growth rate of high-productivity innovators, but changes in the efficiency of adoption affect the innovation behavior of frontier firms.

The first case highlights that a finite support distribution means that there is a positive option value of adoption even for leading firms. When firms decide how much to invest in innovation, they take into account the option value of adoption. On the margin, the more attractive is adoption, the stronger is the temptation to free ride and the weaker is the incentive to innovate.

The second case is that in which ideas are excludable and adopters need to pay a licensing fee to the higher-productivity firms whose technology they want to adopt. In this case, more adopters lead to more licensing fees. When profits from licensing are tied to the quality of the ideas being licensed (e.g., with bargaining over surplus), there can be a positive link between adoption and innovation. In both cases we emphasize how externalities associated with innovation and adoption affect growth.

As in the exogenous growth model of Section 3, in the endogenous growth model the long-run growth rate cannot be larger than the maximum innovation rate of firms. Aggregate growth is therefore driven by innovation (i.e., if the cost of technology diffusion went to 0, the growth rate would still be bounded by γ).

Licensing. Up to now, the firm providing the underlying technology to the adopter was not able to prevent being imitated—i.e., there was no excludability of the technology or intellectual property protection. To bring excludability to this environment with adopters and innovators we model licensing, in which an adopting firm must pay a fee to the technology holder in order to adopt it.

The licensing fee is a fraction of the present discounted value of adopting the technology, paid up front in a lump sum. Firms bargain to determine the size of the licensing fee. The outside option of the adopting firm (i.e., the licensee) is to reject the offer and continue on with its existing technology—i.e., $v(0)$. The outside option of the licensor is simply to reject the offer and gain nothing. Negotiations take the form of Nash bargaining, with a bargaining power parameter $\psi \in (0, 1]$ for the adopting firm.[14]

Proposition 2 (Profits and Value Matching with Licensing). *Given equilibrium innovation policy $\gamma(\cdot)$, growth rate g, and distributions $F_i(\cdot)$, the flow profits are*

$$\pi(z) = e^z + \underbrace{(gF_\ell'(0) + (g - \gamma(0))F_h'(0))}_{\textit{Flow of Licensees}} \underbrace{(1 - \psi)(v_\ell(z) - v(0))}_{\textit{Profits per Licensee}}. \tag{30}$$

---

[14]See Shi and Hopenhayn (2017) for a closely related model of licensing. It provides a richer model of bargaining over technology transfers, including a search-and-matching style congestion. Also related is Luttmer (2015a), which provides a model with assignment between teachers and students decentralized through a price-system.

The value-matching condition for adopting firms is

$$v(0) = \frac{1}{\rho} = \int_0^\infty v_\ell(z) \mathrm{d}F(z) - \frac{\zeta}{\psi}. \tag{31}$$

Proof. See Technical Appendix B.2. □

Full bargaining power to the licensor ($\psi = 1$) nests the baseline case without excludability—i.e., $\pi(z) = e^z$ and the cost of adoption is ζ. The value-matching condition reflects that adopters do not gain the full surplus from the newly adopted technology by increasing the effective cost of search to ζ/ψ.

Thus, from an adopter's perspective, the problems with and without license fees are identical, except for a change in the effective cost of adoption and a modification of the post-adoption continuation value of potentially becoming a licensor in the future. The two environments are quite different for the innovator, however, as license fees provide an extra incentive to innovate.

Endogenous Innovation. A firm in the innovative state can choose its own growth rate $\gamma \geq 0$ subject to a convex cost proportional to its current z. Let $\chi > 0$ be the productivity of its innovation technology and the cost be quadratic in the growth rate γ. h-type firms will choose an optimal innovation rate $\gamma(z)$ by considering the effect of innovation on the profits from production and licensing, given by $\pi(z)$ in equation (30), and the timing of technology diffusion.

With endogenous innovation and licensing, the Bellman equations (18) and (19) become

$$\rho v_\ell(z) = \pi(z) - g v_\ell'(z) + \lambda_\ell(v_h(z) - v_\ell(z)) + \eta(v_\ell(\bar{z}) - v_\ell(z)) \tag{32}$$

$$\rho v_h(z) = \max_{\gamma \geq 0} \left\{ \pi(z) - \underbrace{(g - \gamma)}_{\text{Drift}} v_h'(z) - \underbrace{\frac{1}{\chi} e^z \gamma^2}_{\text{R\&D cost}} + \lambda_h(v_\ell(z) - v_h(z)) + \eta(v_\ell(\bar{z}) - v_h(z)) \right\} \tag{33}$$

Previously, the smooth-pasting condition was not a necessary condition for h-type firms because they never crossed the adoption boundary on the BGP ($g \leq \gamma$). Now, given that $\gamma(0) < g$ is possible, the h-type smooth-pasting condition may be necessary (see Technical Appendix A.4 for more on this). Consequently,

$$v_\ell'(0) = 0 \text{ if } g > 0 \tag{34}$$

$$v_h'(0) = 0 \text{ if } g > \gamma(0). \tag{35}$$

The laws of motion in (12) also need to take into account the state-dependent $\gamma(z)$, and the possibility that h-type firms may cross the lower boundary if $g > \gamma(0)$, resulting in

$$0 = gF'_\ell(z) + \lambda_h F_h(z) - \lambda_\ell F_\ell(z) - \eta F_\ell(z) + \eta \mathbb{H}(z - \bar{z}) + (S_\ell + S_h)F(z) - S_\ell \tag{36}$$

$$0 = (g - \gamma(z))F'_h(z) + \lambda_\ell F_\ell(z) - \lambda_h F_h(z) - \eta F_h(z) - S_h \tag{37}$$

$$S_\ell = gF'_\ell(0) \text{ if } g > 0 \tag{38}$$

$$S_h = (g - \gamma(0))F'_h(0) \text{ if } g > \gamma(0) \tag{39}$$

Summary of Equations and Numerical Methods. The endogenous innovation model is given by the following Bellman equations, value-matching conditions, smooth-pasting conditions, and KFEs: (14), (15) and (32) to (39). With endogenous growth, the need to jointly solve the nonlinear Hamilton–Jacobi–Bellman equations and the Kolmogorov forward equations necessitates numerical methods. The problem takes the form of a set of ODEs with parameters constrained by equilibrium conditions that are themselves functions of the solutions to the ODEs. We compute the equilibrium using a generally applicable numerical technique based on spectral collocation and quadrature, as detailed in the Computational Appendix.[15]

4.2 The Option Value of Adoption Affects Long-run Aggregate Growth

To focus on the first case in which adoption can affect long-run growth rates via the option value, in this section we shut off the second mechanism (licensing). That is, firms' only source of profits is production ($\pi(z) = e^z$) and there is no licensing cost ($\psi = 1$).

Compared to the exogenous innovation case, the key additional necessary equilibrium condition in the endogenous growth model is the first-order condition of the value function equation (33) with respect to $\gamma(z)$, using $\pi(z) = e^z$ in this no-licensing case. The FOC is

$$\gamma(z) = \tfrac{\chi}{2}e^{-z}v'_h(z). \tag{40}$$

With (40), it can be shown that the innovation rate is increasing in productivity—i.e., $\gamma'(z) > 0$. Consider this rate at the adoption boundary $z = 0$ to see that $\gamma(0) = v'_h(0)$. Then, $\gamma(0) = 0$ according to the smooth-pasting equation (35). The intuition is that since

---

[15]The technique uses a simple trick: line up the collocation nodes for the function approximation with those of the quadrature nodes for calculating expectations and equilibrium conditions. After everything is lined up, you can naively stack every equation in the model, including the Bellman equations, KFEs, equilibrium conditions, etc., into a single nonlinear system of equations, and solve without any nested fixed points. In practice, this requires using a high-performance solver and auto-differentiation, but it is easy to implement and reasonably fast.

the firm is right next to the adoption barrier, there is no additional value in increasing its productivity marginally because it will adopt a new technology immediately.

This demonstrates a tradeoff in firms' innovation decisions: Investing more in innovation grows their productivity and increases their profits, but firms with higher productivity are further from the adoption threshold, and thus innovation decreases the option value of adoption. Since the option value of adoption is a larger component of total value for lower-productivity firms, the lower-productivity firms invest less in innovation. Intuitively, for a firm just above the adoption threshold, why invest in innovation to get an incremental improvement when it can save the cost of innovating and, instead, adopt a technology that is discretely better in expectation than the one it is currently using? Of course, the cost of adopting and innovating will jointly determine this adoption threshold.

Proposition 3 (Stationary Equilibrium with Continuous Endogenous Innovation and Bounded Support). *The endogenous innovation choice is such that $\gamma(0) = 0$ and $\gamma(\bar{z}) = g$. A continuum of equilibria exist, parameterized by $\bar{z}$.*

Proof. See Technical Appendix B.1 with numerical methods in the Computational Appendix.
□

In the endogenous innovation case, there are a continuum of equilibria indexed by the frontier $\bar{z}$, each with an associated aggregate growth rate $g(\bar{z})$.

Compared to Stokey (2014), who features a similar innovation process but differs in the treatment of adoption, here the endogenous choice of γ is complicated by the option value of adoption. Different distributions and associated $\bar{z}$ induce different option values and allow for a continuum of self-fulfilling $\gamma(\bar{z})$. That is, a smaller $\bar{z}$ increases the option value of adoption for innovators at the frontier, which is a disincentive to innovate; this leads to less innovation at the frontier, which, consistently, generates a smaller $\bar{z}$. The hysteresis comes from economic forces (rather than technical properties of stochastic processes and initial conditions) and is due to a complementarity between firms' decisions and the shape of the distribution.

Because of this, for any $\eta > 0$, technology adoption leads to differences in the aggregate growth rate by affecting the location of the frontier—and, consequently, the strength of the option value of technology diffusion.

An example of the optimal innovation policy and productivity distributions for an endogenous growth BGP is shown in Figure 6. In addition, Figure 7 plots the growth rate as a function of the frontier. This figure illustrates the intuition that, because of the self-fulfilling balancing of innovation incentives and frontier location, lower values of $\bar{z}$ are associated with lower aggregate growth rates. The smaller the relative frontier, the larger the option value of adoption at the frontier, and the lower the incentive to push out the frontier by innovation.
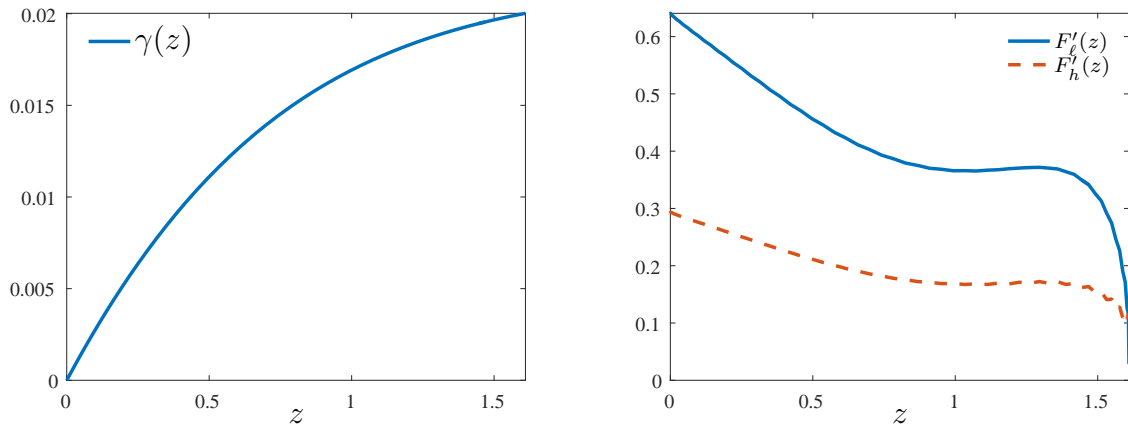
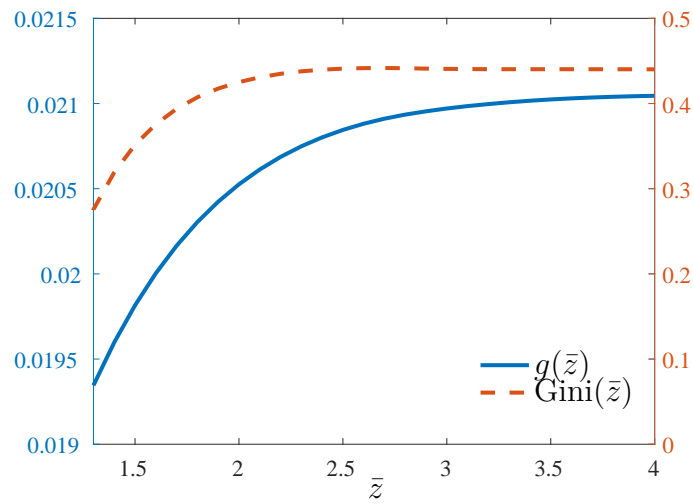Figure 6: Endogenous $\gamma(z)$ and $F_i'(z)$ with $\eta > 0$
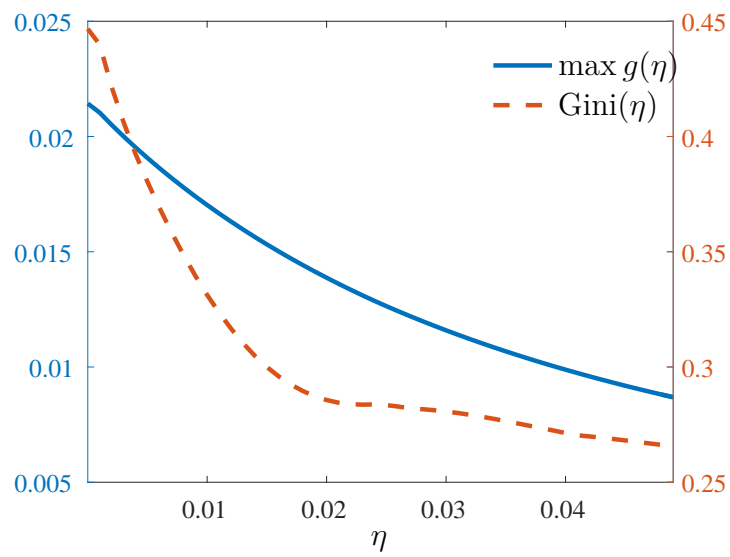

Figure 7: Equilibrium g as a Function of $\bar{z}$

Figure 8: Maximum Equilibrium $g(\eta)$

23

Figure 8 plots the maximum growth rate from the set of equilibrium g as a function of η. The maximum possible growth rate is a decreasing function of η. With more jumps to the frontier, the distribution becomes more compressed, as discussed in the exogenous innovation case of Section 3 and depicted in Figure 5. As the growth rate of the frontier is determined by the innovation decision at $\bar{z}$, the more compressed is the distribution the lower is the innovation rate, for the same reasons that $g(\bar{z})$ is increasing in Figure 7.

Comparison to Models with Exogenous Innovation. The results in this section are in contrast to Perla and Tonetti (2014) and the baseline model in Lucas and Moll (2014), where long-run growth is determined by initial conditions. Those papers can be interpreted as modeling how adoption can generate growth in the medium run.

It is also distinct from models with exogenous innovation and diffusion modeled as geometric Brownian motion, such as Staley (2011) and Luttmer (2012b). In those models, taking the innovation meeting rate to infinity for the large finite agent limit leads to unbounded idiosyncratic growth rates. Here we get the opposite relationship between adoption rates and growth, in part due to adoption and innovation being choices; decreasing the cost of technology diffusion can decrease the growth rate, as it induces more free-riding for innovators. Another key difference in our setup is that the finite-state Markov process and endogenous investment in innovation at convex cost conspire to yield a finite upper bound on firm-level growth rates.

Endogenous Innovation with $\eta \approx 0$. To emphasize that it is the option value of adoption that generates the link between adoption, innovation, and aggregate growth, in this section we briefly study the case where $\bar{z}$ is asymptotically unbounded.

Proposition 4 (Endogenous Innovation with $\eta \approx 0$). *For $\eta \to 0$, an equilibrium exists such that $\lim\limits_{t \to \infty} \bar{z} = \infty$ and the unique growth rate is the solution to the cubic equation*

$$g\left(g^2 + g(2\lambda_h + \lambda_\ell + 3\rho) + 2\rho(\lambda_h + \lambda_\ell + \rho)\right) = \chi(g + \lambda_h + \lambda_\ell + \rho). \tag{41}$$

The endogenous innovation choice is such that $\gamma(0) = 0$ and $\lim\limits_{\bar{z} \to \infty} \gamma(\bar{z}) = g$.

Proof. See Technical Appendix B.1. The numerical method to compute the equilibrium $\gamma(z)$ and the $F_i(z)$ is described in the Computational Appendix. □

As $\eta \to 0$, the number of jumps to the frontier approaches 0 and the model studied in Proposition 3 converges to the model studied in Proposition 4.

Whenever innovation is a choice, the endogenous aggregate growth rate is the growth rate chosen by innovators at the frontier. In contrast to the $\eta > 0$ case for larger η, when $\eta \approx 0$ the aggregate growth rate is independent of the cost of adoption. This is because
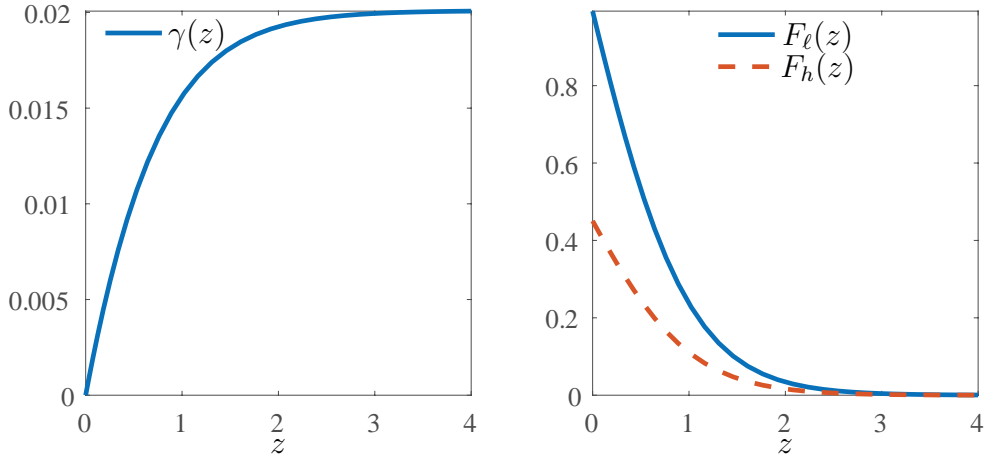
Figure 9: Endogenous $\gamma(z)$, $F_i(z)$, and $\alpha(z)$ with $\eta \approx 0$

innovating firms at the frontier have zero option value from adoption; therefore, changes in the cost or benefits of adoption do not alter their innovation behavior. For intuition, see that the option value term vanishes for large z in equation (29) (i.e., $v_l(z) \propto e^z$ for $z \to \infty$). The right panel of Figure 10 plots comparative statics for a change in the adoption cost when $\eta \approx 0$. As already stated in theory, the figure shows that the aggregate growth rate is invariant to changes in the cost of adoption.
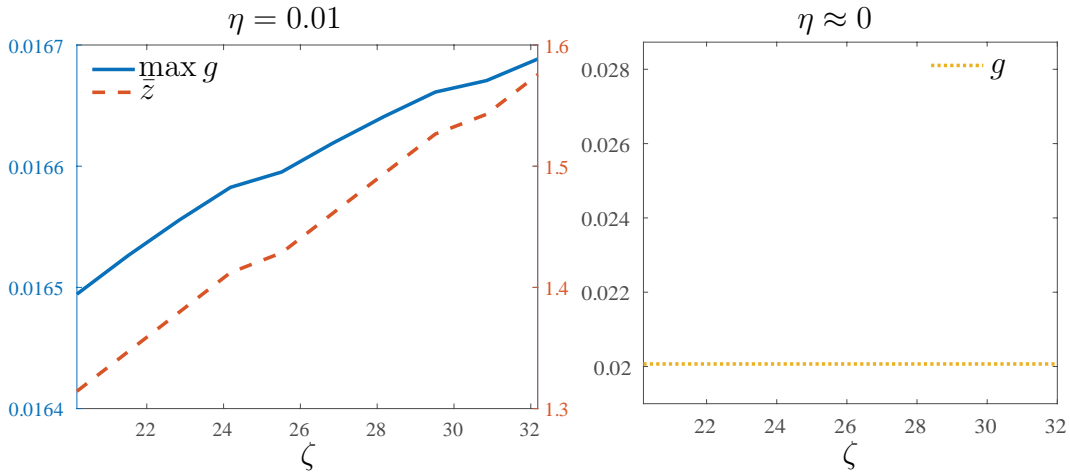

Figure 10: How Aggregate Growth Varies with Adoption Costs with and without Leapfrogging (no licensing $\psi = 1$)

While the growth rates are independent of the adoption costs for $\eta \approx 0$, this is not true in general, since innovation incentives are affected by adoption costs when there is a stationary finite frontier. The left panel of Figure 10 plots comparative statics on the adoption cost ζ in the model with $\eta > 0$. This figure shows that an increase in the adoption cost can increase the growth rate, because it decreases the free-riding incentive for innovators. A welfare interpretation is that a government subsidy to technology adoption financed by lump-sum

taxes would decrease aggregate the growth rate. The effect of the subsidy on aggregate growth would be stronger when frontier firms are closer to the worst firms in the economy. This result holds in the case in which technology is not excludable ($\psi = 1$). We now turn to analyzing the case of licensing, in which a subsidy to adoption can have a very different effect on innovation and aggregate growth.

4.3   Licensing and Partial Excludability

In order to isolate the effect of licensing on aggregate growth, we focus on the $\eta \approx 0$ case, where Figure 10 demonstrates that, absent licensing, changes in the cost of adoption do not affect the aggregate growth rate.

Analysis of Flow Profits.   To study how licensing affects innovation incentives we differentiate the profit function in equation (30):

$$\pi'(z) = e^z + (1 - \psi)gF'(0)v_\ell'(z). \tag{42}$$

On the margin, increasing firm productivity z increases profits for two reasons. The first term e^z is the marginal increase in profits from an increase in production. The second term is the increase in profits from an increase in licensing revenue. Since the value function is increasing in productivity, the second term is positive, so licensing provides a positive incentive to innovate. Furthermore, since the value function is convex, licensing provides stronger incentives to innovate for higher productivity firms. The profits from licensing disappear as $z \to 0$ because the surplus from adopting a technology close to the adoption boundary goes to 0. Consequently, $\pi'(0) = 1$. Finally, with licensing, profits become a function of g. Faster growth means more adopters given a fixed $F(z)$, and $F(z)$ and $v_\ell(z)$ are also themselves functions of g.

Role of Excludability.   Figure 11 plots the aggregate growth rate as a function of the excludability parameter ψ. When excludability is not too strong, the aggregate growth rate is increasing in the degree of excludability (1-ψ) (i.e., growth increases with weaker bargaining power for the adopter). The increase in the aggregate growth rate is due to the added incentive to invest in growing via innovation, as higher-productivity firms gain extra profits from licensing the better technology to adopting firms.

There is, however, a countervailing force that dominates when excludability is already strong. If the licensor's bargaining power is too strong, the incentive to adopt technologies becomes too small. Consequently, fewer firms adopt new technologies, ultimately generating less licensing revenue. Lower licensing revenue decreases the returns to innovation for all firms, including those near the frontier that determine the aggregate growth rate.

To give a sense of the distribution shape, we plot the Gini index. For a wide range of the parameter values, increasing excludability increases innovation activity and generates a more unequal distribution. This shows a trade-off between productivity inequality and aggregate growth rates. This independent positive association between productivity dispersion and the aggregate growth rate operates through innovation activity, compared to the typical link (i.e., Perla and Tonetti (2014)) driven by adoption incentives.
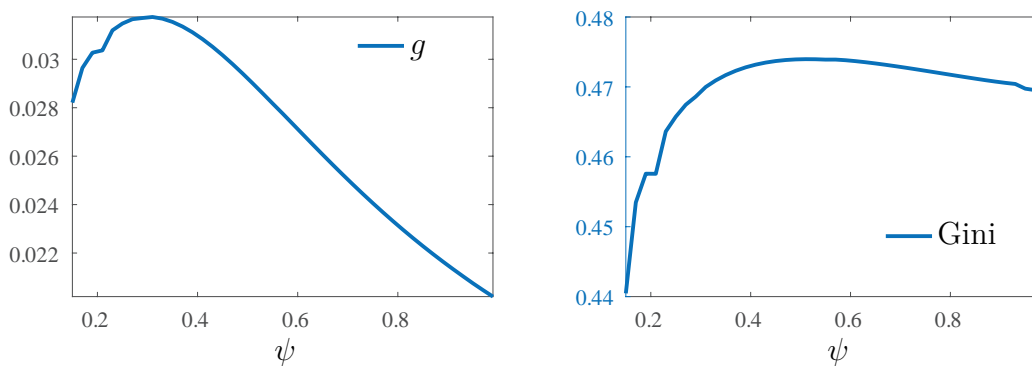

Figure 11: Growth and Distribution Shape under Excludability

Figure 12 presents another perspective on the role of excludability by plotting the aggregate growth rate as a function of the adoption cost for various values of ψ. The model used to generate Figure 12 is identical to the one used to generate the right panel in Figure 10, except ψ is no longer equal to 1. In the absence of excludability (i.e., $\psi = 1$), adoption costs can change the shape of the distribution, but they have no impact on the aggregate growth rate—since the option value of adoption is infinitesimal for the highest-productivity agents making the innovation decision at the frontier. With a strong degree of excludability, however, lower adoption costs drive higher aggregate growth, even in this case with $\eta \approx 0$. While the option value of adoption for firms at the frontier is still infinitesimal, an innovating firm gains extra profits by licensing to adopting firms, and the number of adopting firms increases when adoption costs are lower.

The Gini coefficient decreases modestly but is nearly flat. The reason is that the shape of the distribution near the adoption threshold is impacted by the large mass of agents there, which is generally determined by the innovation decisions of those lower-productivity firms rather than by the frontier innovation rates.

Finally, we consider a variation on the experiment in Figure 10, which showed that without licensing ($\psi = 1$) a decrease in the cost of technology adoption decreased the aggregate growth rate (by strengthening the free-riding incentive for innovators). In Figure 13 we plot the same reduction in adoption costs, but in a model with licensing and with positive option value of adoption at the frontier. Licensing can overturn the negative relationship between adoption costs and aggregate growth rates. Conducting the same exercise with $\psi = 0.5$ yields
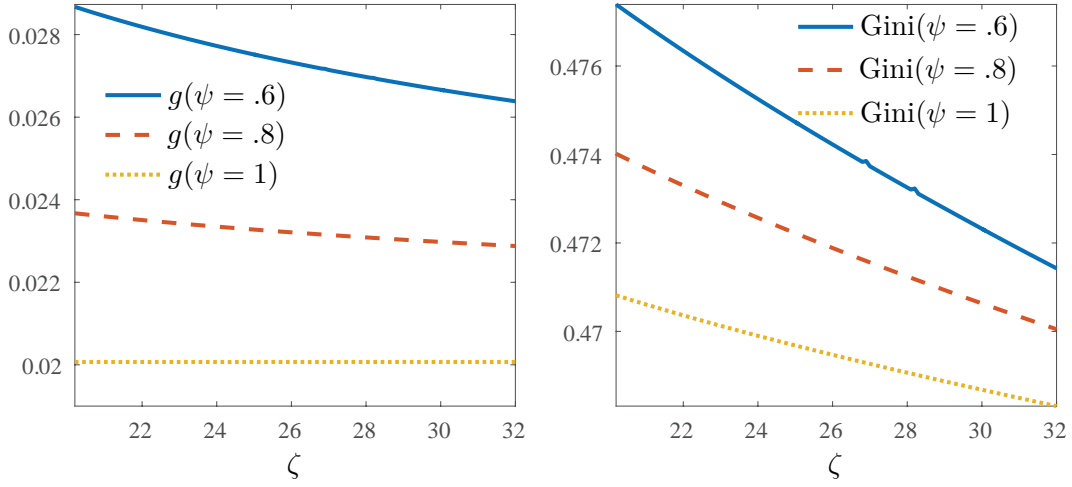
Figure 12: Interaction of Excludability and Adoption Costs

such a reversal, as shown in Figure 13, where subsidizing technology adoption increases the aggregate growth rate.
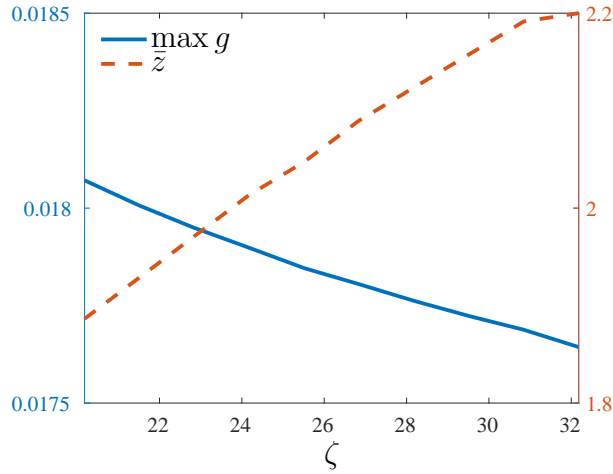

Figure 13: Comparative Statics on Adoption Cost with $\eta = 0.01$ and $\psi = 0.5$

5   Conclusion

This paper develops a theory of the shape of the productivity distribution and of how productivity improves over time, which generates long-run aggregate growth. Firms make choices to invest in adoption and innovation, and the balance of these two activities across firms determines aggregate outcomes. Adoption is a force that compresses the distribution, helping laggards keep up with an expanding frontier. Innovation is a force that stretches the distribution, pushing out the frontier. Balanced growth path equilibria with stationary normalized distributions exist. On a BGP, a lower cost of adoption creates a thinner-tailed

distribution with a smaller distance between best and worst firms. Easier innovation has an opposite effect, generating thicker tails and a larger range in productivity. More firms leapfrogging to the frontier generates thicker tails, but, because it helps laggards keep up with innovators at the frontier, it shrinks the distance from the bottom to the top of the distribution.

In addition to interacting to determine the shape of the distribution, adoption and innovation combine to generate aggregate growth. On a BGP, the long-run aggregate growth rate is the maximum growth rate chosen by innovating firms. If there were no innovation, there would be no growth. In this sense, innovation is the driver of growth. Adoption, however, affects long-run growth by affecting the incentives to innovate. In equilibrium, low-productivity firms choose to adopt and high-productivity firms focus on innovation. Thus, growing through innovation increases the expected time until a firm becomes an adopter. Because innovators may one day become adopters there is an option value of adoption. If this option value is large, e.g., because adoption is low cost, then innovators may be tempted to free ride by investing less in pushing out the frontier, content to fall back to the adoption threshold faster. This complimentary between the distance to the frontier and the incentive to innovate generates multiplicity of BGP. If the initial distribution has a small distance to the frontier, the option value of adoption is high, and the incentives to push out the frontier are low. Thus, the small distance to the frontier is self-reinforcing. In this sense, a subsidy to adoption can reduce aggregate growth.

Adoption, on the other hand, may be a force to increase aggregate growth. When adopting firms must pay a licensing fee to the higher-productivity firm from which they are adopting there is an extra incentive to innovate. Innovation increases productivity, and profits from licensing a high-productivity technology are larger than profits from licensing a mediocre technology. Thus, more adopters induces more innovation, increasing aggregate growth. Furthermore, there is an optimal level of excludability, parameterized by the bargaining power of adopters, that balances the number of adopters and the profits per adopter to maximize growth. Too strong of a bargaining power for adopters limits the innovation incentive provided by the profits from licensing. Too weak of a bargaining power for adopters raises the effective cost of adoption to the point that there are too few adopters.

In sum, the model illustrates how firm choices to adopt and innovate intertwine to generate aggregate effects in equilibrium.

References

ACEMOGLU, D., P. AGHION, C. LELARGE, J. VAN REENEN, AND F. ZILIBOTTI (2007): "Technology, Information, and the Decentralization of the Firm," *The Quarterly Journal of Economics*, 122, 1759–1799.

ACEMOGLU, D., P. AGHION, AND F. ZILIBOTTI (2006): "Distance to Frontier, Selection and Economic Growth," *Journal of the European Economic Association*, 4, 37–74.

ALVAREZ, F. E., F. J. BUERA, AND R. E. LUCAS (2008): "Models of Idea Flows," *NBER WP 14135*, 1–12.

———— (2013): "Idea Flows, Economic Growth, and Trade," Working Paper 19667, National Bureau of Economic Research.

ATKESON, A. AND A. T. BURSTEIN (2010): "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 118, 433–484.

BENHABIB, J., J. PERLA, AND C. TONETTI (2014): "Catch-up and fall-back through innovation and imitation," *Journal of Economic Growth*, 19, 1–35.

BUERA, F. J. AND R. E. LUCAS (2018): "Idea Flows and Economic Growth," *Annual Review of Economics*, 10, 315–345.

BUERA, F. J. AND E. OBERFIELD (2019): "The Global Diffusion of Ideas," *Econometrica (Forthcoming)*.

CHU, A. C., G. COZZI, AND S. GALLI (2014): "Stage-dependent intellectual property rights," *Journal of Development Economics*, 106, 239–249.

EECKHOUT, J. AND B. JOVANOVIC (2002): "Knowledge Spillovers and Inequality," *American Economic Review*, 92, 1290–1307.

JOVANOVIC, B. AND G. M. MACDONALD (1994): "Competitive Diffusion," *Journal of Political Economy*, 102, 24–52.

JOVANOVIC, B. AND R. ROB (1989): "The Growth and Diffusion of Knowledge," *Review of Economic Studies*, 56, 569–582.

KÖNIG, M., J. LORENZ, AND F. ZILIBOTTI (2016): "Innovation vs imitation and the evolution of productivity distributions," *Theoretical Economics*, 1053–1102.

KÖNIG, M., K. STORESLETTEN, Z. SONG, AND F. ZILIBOTTI (2018): "From Imitation to Innovation? Where Is All That Chinese R&D. Going?" Tech. rep., Working Paper.

LASHKARI, D. (2016): "Innovation, Knowledge Diffusion, and Selection," *Working Paper.*

LUCAS, R. E. (2009): "Ideas and Growth," *Economica*, 76, 1–19.

LUCAS, R. E. AND B. MOLL (2014): "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 122, pp. 1–51.

LUTTMER, E. G. (2007): "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, 122, 1103–1144.

———— (2012a): "Eventually, Noise and Imitation Implies Balanced Growth," *Federal Reserve Bank of Minneapolis Working Paper 699.*

———— (2012b): "Technology diffusion and growth," *Journal of Economic Theory*, 147, 602 – 622.

———— (2015a): "An Assignment Model of Knowledge Diffusion and Income Inequality," Tech. Rep. 509, Federal Reserve Bank of Minneapolis.

———— (2015b): "Four Models of Knowledge Diffusion and Growth," Tech. rep., Federal Reserve Bank of Minneapolis.

NELSON, R. R. AND E. S. PHELPS (1966): "Investment in Humans, Technological Diffusion, and Economic Growth," *American Economic Review*, 56, 69–75.

PERLA, J. AND C. TONETTI (2014): "Equilibrium Imitation and Growth," *Journal of Political Economy*, 122, pp. 52–76.

SAMPSON, T. (2015): "Dynamic Selection: An Idea Flows Theory of Entry, Trade and Growth," *The Quarterly Journal of Economics.*

SHI, L. AND H. A. HOPENHAYN (2017): "Knowledge Creation and Diffusion with Limited Appropriation," 2017 Meeting Papers 1501, Society for Economic Dynamics.

STALEY, M. (2011): "Growth and the diffusion of ideas," *Journal of Mathematical Economics*, 47, 470–478.

STOKEY, N. L. (2014): "The Race Between Technology and Human Capital," *Working Paper.*

———— (2017): "Technology, Skill, and Long-Run Growth," *Working Paper.*

SYVERSON, C. (2011): "What Determines Productivity?" *Journal of Economic Literature*, 49, 326–65.