The Growth Dynamics of Innovation, Diffusion, and the Technology Frontier

Jess Benhabib, NYU
Jesse Perla, UBC
Christopher Tonetti, Stanford GSB

March 25, 2016
Draft Version: 193

Active Work in Progress
[Download Newest Version]
[Download Technical Appendix]

Abstract

The recent literature on idea flows studies technology diffusion in isolation, in environments without the generation of new ideas. Without new ideas, growth cannot continue forever if there is a finite technology frontier. In an economy in which firms choose to innovate, adopt technology, or keep producing with their existing technology, we study how innovation and diffusion interact to endogenously determine the productivity distribution with a finite but expanding frontier. There is a tension in the determination of the productivity distribution—innovation tends to stretch the distribution, while diffusion compresses it. Finally, we analyze the degree to which innovation and technology diffusion at the firm level contribute to aggregate economic growth and can lead to hysteresis.

Keywords: Endogenous Growth, Technology Diffusion, Innovation, Imitation, R&D, Technology Frontier

JEL Codes: O14, O30, O31, O33, O40

1 Introduction

This paper studies how adoption and innovation jointly determine the shape of the productivity distribution, the expansion of the technology frontier, and the aggregate economic growth rate. Firms improve their productivity in many ways; some firms do research and development to create new ideas, but there are also many firms modestly improving their production practices over time not through R&D, but rather by adopting already invented ideas. Empirical estimates of productivity distributions tend to have a large range with many low productivity firms and few high productivity firms (Syverson (2011)). The economy is filled with firms of differing production efficiency trying to improve all the time and changes in productivity associated with these innovative and adoptive activities largely determine the shape and movement of the productivity distribution.

Long run per-capita growth is sustained by idea creation that generates productivity growth. Understanding the forces that move the productivity distribution over time is essentially linked to understanding the determinants of aggregate growth, one of the most important welfare-improving forces in human history. Furthermore, the shape of the productivity distribution is itself important to many economic questions across many fields, e.g., in international trade (Eaton and Kortum (2002) and Melitz (2003)), macroeconomics (Hsieh and Klenow (2009)), and industrial organization (Hopenhayn (1992), Foster, Haltiwanger, and Syverson (2008)). Given that adoption is per se of a technology already invented, the existing distribution of technologies available to adopt is related to the existing distribution of technologies in use. That adopted ideas needed to be invented introduces spillovers between the adoption and innovation behavior of all firms.

Given the importance of aggregate growth, there are relatively few foundational models of endogenous growth. Many of the classic papers (Romer (1986, 1990), Segerstrom, Anant, and Dinopoulos (1990), Rivera-Batiz and Romer (1991), Grossman and Helpman (1991, 1993), and Aghion and Howitt (1992)) and their successors study innovation and adoption as in this paper, but often with a different focus. The literature following Romer (1990) highlights the importance of idea creation in generating long run growth, but tends to do so in models with a representative R&D sector not designed to study the productivity distribution, nor how externalities across firms affects innovation activity or how heterogeneous firm data can discipline growth theory. The literature that builds on the framework of Aghion and Howitt (1992), such as Klette and Kortum (2004) or Acemoglu, Akcigit, Bloom, and Kerr (2013), is designed exactly with heterogeneous firms in mind, but they are explicitly models of creative destruction. These are models well suited to represent frontier firms and innovative activity at the top of the distribution.

We propose that different forces are at play for the many firms in the lower tail of the productivity distribution. For these firms, most improvements are likely to be incremental and don't have the winner-take-all aspect of creative destruction. Firms of similar productive ability are able to coexist in the market place due to slight product differentiation better described by monopolistic competition than the limit pricing leader-follower behavior prevalent at the top as modeled by the literature connected to Aghion and Howitt (1992). Thus, we aim to contribute an alternative model of the productivity distribution over the full range of support in as parsimonious and transparent way as possible. In this regard, we build on the idea diffusion literature of Lucas and Moll (2014) and Perla and Tonetti (2014), adding idea creation to the model. Thus, this paper delivers an applied theory to study the key spillovers across firms' productivity-enhancing activity and their macro implications, with the long-term goal of using rich firm dynamics data to modify and further discipline this growth theory.

We first build a simple model of exogenous innovation and growth to focus on how innovation and adoption affect 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 determines 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 a distribution related to the existing distribution of technologies currently in use for production. This is a tractable way to capture the fact that technology adoption must be related to the distribution of technologies already invented and that even adoption is a risky decision for the adopting firm. Since adoption, by definition, depends on the distribution of technologies, and since we want to study the interaction between innovation and adoption, we model the benefits of expenditure on innovation as independent of the productivity distribution. Innovation does not exist in a bubble and firms doing R&D certainly learn from each other, but this way, all interactions between innovation and adoption are isolated to occur directly through firm behavior. Specifically, we model firms as being in either a creative or stagnant innovation state, and when creative, innovation is geometric growth in productivity at a rate increasing in firm-specific innovation expenditure. A firm's innovation state evolves according to a two state Markov process, and this particular stochastic model of innovation is the key technical feature that delivers many of the desired model properties in a tractable framework. For example, we want the productivity distribution to have finite support so that there are better technologies to be invented and not all knowledge that will ever be known was in use for production at time zero. This permits an analysis of idea creation driving long-run growth with adoption potentially affecting growth through the diffusion of newly invented ideas. Given a continuum of firms, modeling stochastic innovation using geometric Brownian motion, as is common in the literature, would generate infinite support instantly, while the finite-state Markov process allows for finite support for all time.

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 R&D to grow fast, create new knowledge, and 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 behind too far. A low cost of innovation tends to spread the distribution, as the high productivity firms can more easily escape from the pack. The stochastic innovation state ensures that some firms who have bad luck and stay uncreative for a stretch of time fall back relative to adopting and innovating firms, generating non-degenerate normalized distributions with adopting activity existing in the long run. Thus, the shape of the distribution, which typically looks like a truncated Pareto with finite support, is determined by the relative ease of adoption and innovation through the differing rates at which high and low productivity firms grow.

Adoption and innovation are not two completely independent processes with some firms perpetual adopters and some perpetual innovators. Rather, that all firms have the ability to invest in both activities generates general equilibrium interactions between actions. The key spillover between adoption and innovation can be seen in the option value of adoption. For high productivity firms far away from being a low productivity adopter, the value of having the option to adopt is small. The lower a firm's productivity, the closer they are to being an adopter and 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. Thus, the value of adoption, which is determined by the probability of adopting a good technology and the cost of acquiring that technology, affects incentives to innovate. In an extension with licensing, if ideas are partially excludable and adopters pay a fee to the firm whose technology they adopt, there is an extra direct link between adoption behavior and innovation incentives affecting the shape of the distribution and aggregate growth rates.

In addition to the baseline model, we introduce a sequence of extensions designed to enrich the model to capture more ways in which innovation and adoption might interact and to relax some of the stark assumptions prevalent in the literature. We introduce a version of quality ladders by including a probability of leap-frogging to the frontier technology. The baseline model has undirected search for a new technology in that a draw is from the unconditional distribution of technologies and there is no action a firm can take to influence the source distribution. In an extension we model "directed" adoption in that firms can obtain a draw from a skewed distribution in which they can increase the probability of adopting better technologies at a cost. In the baseline model firms exist for all time, their output and profits equal their productivity, there is no explicit cost of production, and there is a single market for the common good all firms produce. While this delivers the cleanest framework for analyzing the key forces, the model is extended to include endogenous entry, exogenous exit and firms who hire labor to produce a unique variety sold via monopolistic competition to a CES final good producer. For each extension we examine properties of the BGP productivity distribution, such as the tail index and the ratio of the frontier to the minimum productivity, and whether the equilibrium is unique or if there is hysteresis in the sense that the long run distribution and growth rate depends on initial conditions.

Through the baseline model and extensions we show what types of stochastic processes can generate data consistent with the empirical evidence: balanced growth with a nearly Pareto firm size distribution with finite support, that also has finite support when normalized by the aggregate growth rate. We show which features are necessary to have both innovation and adoption activity exist in the long run and when and how adoption affects the aggregate growth rate. Finally, we also show that assumptions such as infinite initial support are not innocuous, in that the obviously counterfactual infinite support initial condition implies very different important model properties than the finite support initial condition.

1.1 Recent Literature

Related papers in this class of models include Lucas and Moll (2014), Kortum (1997), Luttmer (2007, 2014) and Sampson (2014) which emphasize selection from optimal entry/exit, and Alvarez, Buera, and Lucas (2008, 2013), which emphasizes diffusion as an arrival rate of ideas from the productivity distribution. Buera and Oberfield (2014) is a related semi-endogenous growth model of international diffusion of technology and its connection to trade. Another approach, taken in Jovanovic and Rob (1989), is to add positive spillovers from the diffusion process itself, which can create balanced growth.

While Perla and Tonetti (2014) isolated the role of growth through "catch-up diffusion", in some sense the role of "stochastic diffusion" is isolated in Luttmer (2007). The "catchup diffusion" effect is not present in Luttmer (2007) in the same sense, as the incumbent firms lower in the productivity distribution gain no benefit from growth. However, in our model, "stochastic diffusion" is different from Luttmer (2007), as firms internalize the value of an upgrade rather than being driven into un-profitability and exit from GE effects.

In this paper we are considering process innovation rather than new product innovation. Smaller firms may be especially innovative in coming up with new products as in Klette and Kortum (2004) or Acemoglu, Akcigit, Bloom, and Kerr (2013), but this is not considered in our innovation technology, as all firms have one product and the number of products in equilibrium is kept fixed for simplicity. Other papers emphasizing the role of an endogenous innovation choice include Atkeson and Burstein (2010) and Stokey (2014).

Acemoglu, Aghion, and Zilibotti (2006), König, Lorenz, and Zilibotti (2012), Chu, Cozzi, and Galli (2014), Stokey (2014), and Benhabib, Perla, and Tonetti (2014) also explore the relationship between innovation and diffusion from different perspectives. 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, distorting their innovation choices. We describe this as an "option value of diffusion", where incumbents take into account the possibility of future improvements in their productivity through jumps from technology diffusion. The lower the relative productivity of a firm, the higher the expected benefit of adoption via a jump to a superior technology, and the sooner the expected time to execute the adoption option. Therefore, low productivity firms have high option values of diffusion, while very high productivity firms may have an asymptotically irrelevant contribution from technology diffusion.

This tension between innovation and technology diffusion explored here has a different emphasis in Luttmer (2007, 2012, 2014), where the generator of diffusion is entry/exit in equilibrium, and only new entrants can internalize the benefits of technology diffusion. The main similarity is that in both papers, some firms sample from the existing distribution of productivity. In particular, Luttmer (2007) is interested in the role of technology diffusion through entry, so it is the entrants who gain the benefits of a growing economy. As incumbents pay a fixed cost that grows with the scale of the economy, entry can spur more exit. Therefore negative profits that results in exit leads to entry and to technology diffusion. In our model incumbents, or operating firms, choose when to exploit the incentives to adopt a new technology.[1] The difference between whether incumbents or entrants internalize the value of a growing economy leads to very different implications for technology diffusion.

2 Baseline Model with Exogenous, Stochastic Innovation

2.1 Basic Setup

Consider a discrete two-state Markov process driving the exogenous growth rate of an operating firm. In the high state, the firm is innovating and increasing its productivity deterministically. In the low state, its productivity does not grow through innovation. This captures the concept that some times firms have good ideas or projects that generate growth and some times firms are just producing using their existing technology. This innovation status changes according to a continuous time Markov chain.[2,3]

Firm Heterogeneity and Choices Assume firms producing a homogeneous product are heterogeneous over their productivity, Z, and over their degree of innovation, $i \in \{\ell, h\}$. The mass of firms of productivity less than Z in innovation state i is defined as $\Phi_i(t, Z)$ (i.e., an unnormalized CDF). Define the technology frontier as the maximum productivity, $B(t) \equiv \sup\{\text{support}\{\Phi_i(t, \cdot)\}\} \leq \infty$, and normalize the mass of firms to 1 so that $\Phi_\ell(t, B(t)) + \Phi_h(t, B(t)) = 1$. At any point in time, the minimum of the support of the distribution will be an endogenously determined $M_i(t)$, 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, in which case it would grow stochastically, or it can choose to adopt a new technology instantaneously. As we will show in equilibrium, all firms choose an identical threshold, $M(t)$, above which they will continue operating with their existing technology (i.e., a firm with $Z \leq M(t)$ chooses to adopt a new technology). As draws are instantaneous, this endogenous $M(t)$ becomes the evolving minimum

[1] Mechanically, these differences manifest themselves in the option value of the Bellman equation. In Luttmer (2007), incumbents are only affected negatively by growth and have a zero option value of technology diffusion, whereas in our model incumbents have a positive option value of diffusion as they can always adopt from by taking a draw from the existing distribution.

[2] Luttmer (2011) also emphasizes the need for fast growing firms–driven by differences in the quality of blueprints for size expansion–to account for the size distribution of firms. In his model, firms will stochastically slow down eventually, where here will assume that firms can jump back and forth between the states. In other work, Luttmer (2014) emphasizes the role of a stochastic shock as "experimentation" as distinct from deterministic innovation, and important in the generation of endogenous tail parameters.

[3] Lucas and Moll (2014) provides an extension of their baseline model with the addition of exogenous innovators in the distribution in order to discuss initial finite support.

of the $\Phi_i(t,Z)$ distribution.[4] The cost of adoption scales with the economy, and for simplicity is proportional to the endogenous scale of the economy, $M(t)$.[5]

If a firm adopts a new technology, then it immediately changes its productivity to a draw from a distortion of the $\Phi_i(t,Z)$ distribution.[6] Assume that an adopting firm draws a (i,Z) from distributions $\hat{\Phi}_\ell(t,Z)$ and $\hat{\Phi}_h(t,Z)$, both of which will be determined by the equilibrium $\Phi_\ell(t,Z)$ and $\Phi_h(t,Z)$. Assume that this gives a proper cdf, so $\hat{\Phi}_\ell(t,0) = \hat{\Phi}_h(t,0) = 0$ and $\hat{\Phi}_\ell(t,B(t)) + \hat{\Phi}_h(t,B(t)) = 1$.

Stochastic Process for Innovation 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 (except, perhaps, exactly at an initial condition). Recall that support $\{\Phi(t,\cdot)\} \equiv [M(t), B(t))$. The growth rate of the upper and lower bounds of the support are defined as $g(t) \equiv M'(t)/M(t)$ and $g_B(t) \equiv B'(t)/B(t)$ if $B(t) < \infty$.

Value Functions and the Growth Rate of the Frontier While $i = h$, firms grow at an exogenous innovation rate $\gamma > 0$, and, without loss of generality, do not grow if $i = \ell$.[7] The continuation value functions are $V_i(t,Z)$ and include the drift in a high state as well as the intensity of jumps between i. The firms discount at rate $r > 0$.[8] For the two discrete states, the Bellman equations in the continuation region are,[9]

$$rV_\ell(t,Z) = Z + \underbrace{\lambda_\ell\left(V_h(t,Z) - V_\ell(t,Z)\right)}_{\text{Jump to } h} + \underbrace{\boldsymbol{\partial}_t V_\ell(t,Z)}_{\text{Capital Gains}} \tag{1}$$

$$rV_h(t,Z) = Z + \underbrace{\gamma Z \boldsymbol{\partial}_Z V_h(t,Z)}_{\text{Exogenous Innovation}} + \underbrace{\lambda_h\left(V_\ell(t,Z) - V_h(t,Z)\right)}_{\text{Jump to } \ell} + \boldsymbol{\partial}_t V_h(t,Z) \tag{2}$$

From this process, if $B(0) < \infty$, then $B(t)$ will remain finite for all t, as it evolves from the innovation of firms in the interval infinitesimally close to $B(t)$; that is, $B'(t)/B(t) = \gamma$ if $\Phi_h(t,B(t)) - \Phi_h(t,B(t)-\epsilon) > 0$, for all $\epsilon > 0$. With the continuum of firms and the memoryless

[4] To show that the minimum of support is the endogenous threshold, assume a Poisson arrival rate of draw opportunities approaching infinity. In any positive time interval firms would gain an acceptable draw with probability 1, so that $Z > M(t)$ almost surely. Because of the immediacy of draws, the stationary equilibrium does not depend on whether draws are from the unconditional distribution or are from the distribution conditional on being above the current adoption threshold. This is the same as the small time limit of Perla and Tonetti (2014), which solves both versions of the model. The derivation of the cost function as the limit of the arrival rate of unconditional draws is in Technical Appendix C.1. The online technical appendix is located at [Download Technical Appendix].

[5] In Technical Appendix B, a more elaborate version of this is derived in general equilibrium where ζ is the quantity of labor required for adoption, but it ends up being qualitatively equivalent. An alternative is to have the cost scale with the firm's Z, which introduces a less convenient smooth pasting condition, but remains otherwise tractable. See sec:monopolistic-competition-gbm for endogenization of the number of firms. The results are qualitatively similar to the fixed firm version.

[6] See Technical Appendix C.2 for a proof that the ability for a firm to recall its last productivity doesn't change the equilibrium conditions, and Technical Appendix C.1 for a derivation of this where adoption is not instantaneous.

[7] In Section 3, the growth rate γ will become a control variable for a firm, with the choice subject to a convex cost.

[8] With a representative consumer with CRRA preferences, discount rate ρ, and intertemporal elasticity of substitution Λ, the interest rate is $r = \rho + \Lambda g$, and $r = \rho + g$ for log utility of the representative consumer.

[9] 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 as $v'(z) \equiv \frac{\mathrm{d}v(z)}{\mathrm{d}z}$. Ordering the states as $\{l,h\}$, the infinitesimal generator for this continuous time Markov chain is $\mathbb{Q} = \begin{bmatrix} -\lambda_\ell & \lambda_\ell \\ \lambda_h & -\lambda_h \end{bmatrix}$, with adjoint operator $\mathbb{Q}^*$. The KFE and Bellman equations can be formally derived using these operators and the drift process.

Poisson arrival of changes in i, there will always be some h firms that have not jumped to the low state for any t, so the growth rate of the frontier is always γ.

Technology Diffusion Firms upgrading their technology through adoption receive a new i type and a draw of Z from the productivity distribution. The exact specification typically does not affect the qualitative results, so we will write the process fairly generally and then analyze specific cases. While the draw from $\hat{\Phi}_i(t, Z)$ is left general, we are maintaining a simplification that the gross value of adoption is independent of an agent's current type.

In principle, there may be adopters hitting the adoption threshold with either innovation type. Assume that h and ℓ types have the same adoption threshold $M(t)$, to be proved later. A flow $S_i(t) \geq 0$ of firms cross into the adoption region at time t and choose to adopt a new technology. Denote the total flow of adopting firms as $S(t) \equiv S_\ell(t) + S_h(t)$.

Law of Motion The Kolmogorov Forward Equations (KFEs) in CDFs include the drift and jumps between innovation states,

$$\partial_t \Phi_\ell(t, Z) = \underbrace{-\lambda_\ell \Phi_\ell(t, Z) + \lambda_h \Phi_h(t, Z)}_{\text{Net Flow from Jumps}} + \underbrace{(S_\ell(t) + S_h(t))}_{\text{Flow Adopters}} \underbrace{\hat{\Phi}_\ell(t, Z)}_{\text{Draw} \leq Z} - \underbrace{S_\ell(t)}_{\text{Adopt}} \tag{3}$$

$$\partial_t \Phi_h(t, Z) = \underbrace{-\gamma Z \partial_Z Z \Phi_h(t, Z)}_{\text{Innovation}} - \lambda_h \Phi_h(t, Z) + \lambda_\ell \Phi_\ell(t, Z) + (S_\ell(t) + S_h(t)) \hat{\Phi}_h(t, Z) - S_h(t) \tag{4}$$

Recognizing that the λ_i jumps are of measure 0 when calculating how many firms cross the boundary in any infinitesimal time period, the flow of adopters comes from the flux across the moving $M(t)$ boundary,[10]

$$\tag{5}$$

$$S_\ell(t) \equiv M'(t) \partial_Z \Phi_\ell(t, M(t)) \tag{6}$$

$$S_h(t) \equiv \underbrace{(M'(t) - \gamma M(t))}_{\text{Relative Speed of Boundary}} \underbrace{\partial_Z \Phi_h(t, M(t))}_{\text{PDF at boundary}} \tag{7}$$

Adoption Decision Firms choose thresholds, below which they adopt a new technology through the technology diffusion process. While the threshold could depend on the type i, see Appendix A.2 for a proof that ℓ and h agents choose the same threshold, $M(t)$, if the net value of adoption is independent of the current innovation type. This result is fairly robust, and holds throughout the paper.

Necessary conditions for the optimal stopping problem include value matching and smooth pasting conditions at the endogenously chosen adoption boundary, $M(t)$,

$$\underbrace{V_i(t, M(t))}_{\text{Value at Threshold}} = \underbrace{\int_{M(t)}^{B(t)} V_\ell(t, Z) \mathrm{d}\hat{\Phi}_\ell(t, Z) + \int_{M(t)}^{B(t)} V_h(t, Z) \mathrm{d}\hat{\Phi}_h(t, Z)}_{\text{Gross Adoption Value}} - \underbrace{\zeta M(t)}_{\text{Adoption Cost}} \tag{8}$$

$$\partial_Z V_\ell(t, M(t)) = 0, \quad \text{if } M'(t) > 0 \tag{9}$$

$$\partial_Z V_h(t, M(t)) = 0, \quad \text{if } M'(t) - \gamma M(t) > 0 \tag{10}$$

[10] This is consistent with the solution to the ODEs in (3) and (4) at $Z = M(t)$, and is clear in the normalized (18) and (19).

Consumer Welfare and Discounting The firms are owned by a representative consumer with log utility and a discount rate $\rho > 0$ who values the undifferentiated good. Given the productivity distribution $\Phi(t, Z)$, the welfare is then

$$U(t) = \int_0^\infty e^{-\rho\tau} \log\left(\underbrace{\int_{M(t)}^\infty Z \partial_Z \Phi(t+\tau, Z)\mathrm{d}Z}_{\text{Aggregate Output}}\right) \mathrm{d}\tau \tag{11}$$

As is standard, along a balanced growth path where aggregate output grows at rate g, the discount rate faced by the firm is.[11]

$$r = \rho + g \tag{12}$$

2.2 Normalization and Stationarity

To find a balanced growth path (BGP), it is convenient to transform this system to a stationary set of equations by normalizing variables relative to the endogenous boundary $M(t)$. Define the change of variables, normalized distribution, and normalized value functions as,

$$z \equiv \log(Z/M(t)) \tag{13}$$

$$F_i(t, z) = F_i(t, \log(Z/M(t))) \equiv \Phi_i(t, Z) \tag{14}$$

$$v_i(t, z) = v_i(t, \log(Z/M(t))) \equiv \frac{V_i(t, Z)}{M(t)} \tag{15}$$

The adoption threshold was chosen to be normalized to $\log(M(t)/(M(t))) = 0$, and the relative technology frontier is $\bar{z}(t) \equiv \log(B(t)/M(t)) \leq \infty$. See Figure 1 for a comparison of the normalized and unnormalized distributions. From this, $F_\ell(t, 0) = F_h(t, 0) = 0$ and $F_\ell(t, \bar{z}(t)) + F_h(t, \bar{z}(t)) = 1$. Denote the unconditional, normalized distribution with $F(0) = 0$ and $F(\bar{z}) = 1$ as $F(z) \equiv F_\ell(z) + F_h(z)$.

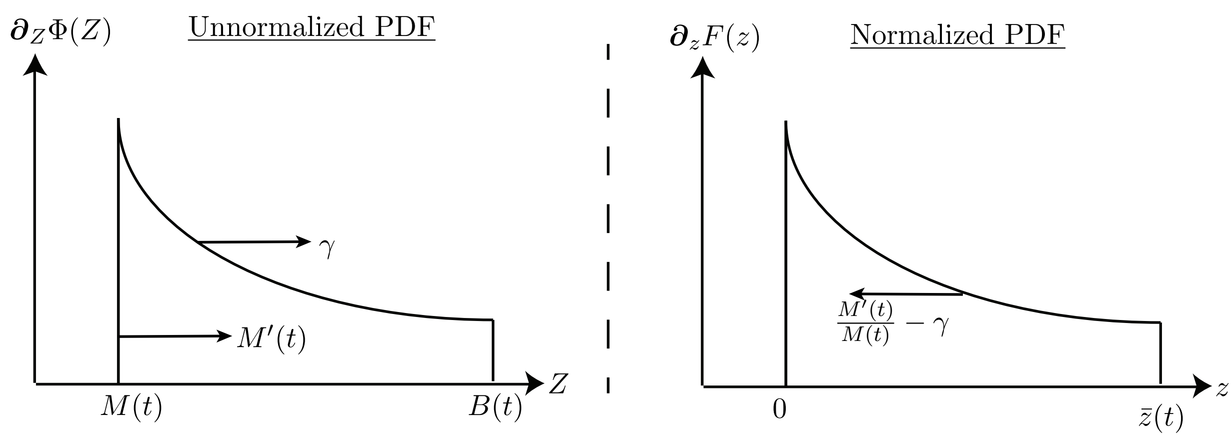

Figure 1: Normalized vs. Unnormalized Distributions

With the above normalizations, the value function, productivity distribution, and growth rates

[11] The appendix uses a general r, and the numerical algorithm in *TA*. **??** is implemented for an arbitrary CRRA utility function. Using log utility here simplifies a few of the parameter restrictions and expressions in the algebra compared to linear utility or general CRRA.

can be stationary and independent of time.[12] When the distribution is not time varying, let $F'(z)$ denote the probability density function.

Summary of Stationary Equilibrium A full derivation of the normalization is done in Appendix A.1, which leads to the following normalized set of stationary equations for the evolution of the distribution,

$$0 = gF_\ell'(z) - \lambda_\ell F_\ell(z) + \lambda_h F_h(z) + (S_\ell + S_h)\hat{F}_\ell(z) - S_\ell \tag{18}$$
$$0 = (g-\gamma)F_h'(z) - \lambda_h F_h(z) + \lambda_\ell F_\ell(z) + (S_\ell + S_h)\hat{F}_h(z) - S_h \tag{19}$$
$$0 = F_\ell(0) = F_h(0) \tag{20}$$
$$1 = F_\ell(\bar{z}) + F_h(\bar{z}) \tag{21}$$
$$S_\ell = gF_\ell'(0) \tag{22}$$
$$S_h = (g-\gamma)\, F_h'(0) \tag{23}$$

As discussed Appendix A.2, the common threshold choice is such that we will drop the index in the rest of the paper, so $v(0) \equiv v_i(0)$. Furthermore, it provides a simple solution for $v(0)$ in terms of the equilibrium growth rate given this convenient normalization. The summary of necessary conditions for the firm's problem are

$$\rho v_\ell(z) = e^z - gv_\ell'(z) + \lambda_\ell\left(v_h(z) - v_\ell(z)\right) \tag{24}$$
$$\rho v_h(z) = e^z - (g-\gamma)v_h'(z) + \lambda_h\left(v_\ell(z) - v_h(z)\right) \tag{25}$$
$$v(0) = \frac{1}{\rho} = \int_0^{\bar{z}} v_\ell(z)\mathrm{d}\hat{F}_\ell(z) + \int_0^{\bar{z}} v_h(z)\mathrm{d}\hat{F}_h(z) - \zeta \tag{26}$$
$$v_\ell'(0) = 0, \text{ if } g > 0 \tag{27}$$
$$v_h'(0) = 0, \text{ if } g > \gamma \tag{28}$$

To interpret: (18) to (21) are the stationary KFE with initial conditions and boundary values. S_ℓ in (22) is the flow of ℓ agents moving backwards at a relative speed of g across the barrier, while S_h in (23) is the flow of h agents moving backwards at the slower relative speed of $g-\gamma$ across the barrier. The $\hat{F}_i(z)$ specification is some function of the equilibrium $F_i(z)$, and will be analyzed further in Sections 2.3 to 2.5.

(24) and (25) are the Bellman Equations in the continuum region, where (26) is the value matching condition between the continuation and technology adoption regions. The smooth pasting conditions in (27) and (28) are only necessary if the firms of a particular i are drifting backwards relative to the adoption threshold. See Figure 2 for a visualization of the normalized Bellman equations.

In the normalized setup, $\bar{z}(t) \equiv \log(B(t)/M(t))$, and a necessary condition for a stationary equilibrium with $\sup\{\bar{z}(t) | \forall t\} < \infty$ is that $g = g_B = \gamma$. This is necessary because if $g < g_B$, then $\bar{z}$ diverges, while if $g > g_B$, the minimum of the support would eventually be strictly greater than the maximum of the support.

[12] An important example is when $\Phi(t,Z)$ is Pareto with minimum of support $M(t)$ and tail parameter α:

$$\Phi(t,Z) = 1 - \left(\frac{M(t)}{Z}\right)^\alpha, \text{ for } M(t) \le Z \tag{16}$$

Then $F(t,z)$ is independent of M and t:

$$F(t,z) = 1 - e^{-\alpha z}, \text{ for } 0 \le z < \infty. \tag{17}$$

This is the cdf of an exponential distribution, with parameter $\alpha > 1$. From a change of variables, if $X \sim Exp(\alpha)$, then $e^X \sim Pareto(1,\alpha)$. Hence, α is the tail index of the unnormalized Pareto distribution for Z.

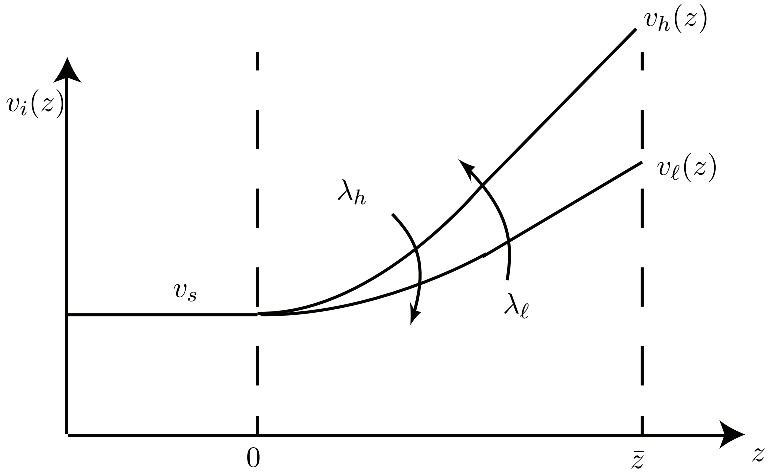

Figure 2: Normalized, Stationary Value Functions

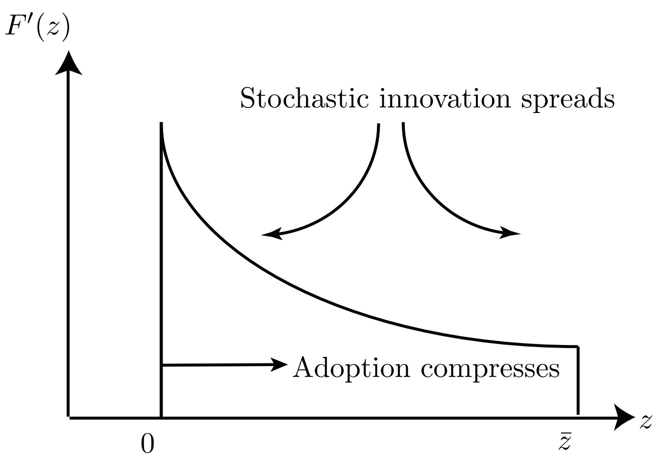

Figure 3: Tension between Stochastic Innovation and Adoption

Terminology for Various Cases of the Normalized Support There are three possibilities for the stationary $\bar{z}$ that we will analyze separately. The first is if $\bar{z} = \infty$, which we will call "infinite support", which happens for any initial condition that starts with $B(0) = \infty$ (i.e., $\sup \text{support}\{F(0, \cdot)\} = \infty$). The second case is when $B(0) < \infty$ (which implies $B(t) < \infty$), but where $\lim_{t\to\infty} \bar{z}(t) = \infty$. We label this case "finite, unbounded support". The final case is when the initial condition has finite support, and $\lim_{t\to\infty} \bar{z}(t) < \infty$, which we will refer to as "finite, bounded support". An important question will be whether the unbounded and infinite support examples have the same stationary equilibrium. It will turn out that this is not the case, suggesting that caution should be used when using an infinite support as an approximation of a finite, but ultimately unbounded, empirical distribution.

An important question is whether a stationary equilibrium with bounded finite support can even exist for a given version of the model. This will be discussed further in Propositions 1 and 2.

2.3 Stationary BGP with Finite Initial Support

In this section we study the stationary distribution, the BGP, when the initial distribution $\Phi(0, Z)$ has finite support. Consider for simplicity that the process of adopting new technologies is disruptive to R&D, so the firm starts in the ℓ type regardless of its former type, and the Z is drawn from a distortion of the unconditional distribution.[13] The distortion, representing the degree of imperfect mobility, is indexed by $\kappa > 0$ where the agent draws its Z from the cdf $\Phi(t, Z)^\kappa$. Note that for higher κ, the probability of a better draw increases. As $\Phi(t, B(t))^\kappa = 1$ and $\Phi(t, M(t))^\kappa = 0$, for all $\kappa > 0$, this is a valid probability distribution. While κ is exogenous here, in Section 3.3, we solve a version of the model with directed technology diffusion where κ is endogenous.

$$\hat{\Phi}_\ell(t, Z) \equiv (\Phi_\ell(t, Z) + \Phi_h(t, Z))^\kappa \tag{29}$$

Normalizing to a stationary draw distribution and then using the definition of the unconditional normalized distribution, $F(z)$, yields adoption distribution

$$\hat{F}_\ell(z) = (F_\ell(z) + F_h(z))^\kappa = F(z)^\kappa \tag{30}$$

We write $F(z)^\kappa$ for the normalized draw process (for which, given the assumptions, all firms end up in the low state). With this, the pdf for the draw distribution is

$$\boldsymbol{\partial}_z(F(z))^\kappa = \kappa F(z)^{\kappa-1}\boldsymbol{\partial}_z F(z) \tag{31}$$

Due to the bounded growth rates of the Markov process, if the support of $\Phi(0, z)$ is finite, then it remains finite as it converges to a stationary distribution. With an exogenous γ and a finite frontier, a necessary requirement for non-degeneracy of $F_i(z)$ is then $g = \gamma$. Hence, in the stationary equilibrium there are no h type agents hitting the adoption threshold, and the smooth pasting condition for h firms is not a necessary condition.

Necessary conditions for a stationary equilibrium with a finite initial frontier are $v_\ell(z)$, $v_h(z)$, $F_\ell(z)$, $F_h(z)$, S—such that (20) to (22), (24) and (27), and

$$\rho v_h(z) = e^z + \lambda_h(v_\ell(z) - v_h(z)) \tag{32}$$

$$v(0) = \frac{1}{\rho} = \int_0^{\bar{z}} v_\ell(z)\mathrm{d}F(z)^\kappa - \zeta \tag{33}$$

$$0 = gF_\ell'(z) + SF(z)^\kappa + \lambda_h F_h(z) - \lambda_\ell F_\ell(z) - S \tag{34}$$

$$0 = \lambda_\ell F_\ell(z) - \lambda_h F_h(z) \tag{35}$$

Define the constants, $\hat{\lambda} \equiv \frac{\lambda_\ell}{\lambda_h}, \bar{\lambda} \equiv \frac{\lambda_\ell}{\rho+\lambda_h} + 1$. The following characterizes the equilibrium,

[13] Unlike the infinite support case in Section 2.3, the equilibrium are not sensitive to the degree of correlation in the draws, and we have simply chosen the most convenient.

Proposition 1 (Stationary Equilibrium with Continuous Draws and Finite, Unbounded Support)**.** *There does not exist an equilibrium with finite and bounded support (for any $\kappa > 0$). There exists a unique equilibrium, with $g = \gamma$ and $\bar{z} \to \infty$.*

In the case of $\kappa = 1$, the unique stationary distribution is,

$$F_\ell(z) = \tfrac{1}{1+\hat{\lambda}} e^{-\alpha z} \tag{36}$$

$$F_h(z) = \hat{\lambda} F_\ell(z), \tag{37}$$

where α is the tail index of the power law distribution:

$$\alpha \equiv (1 + \hat{\lambda}) F_\ell'(0), \tag{38}$$

and $F_\ell'(0)$ is determined by model parameters:

$$\tag{39}$$

$$F_\ell'(0) = \frac{\lambda_h \Big(\zeta(\rho+\gamma)((\rho+\gamma)+\lambda_h+\lambda_\ell) - \sqrt{\zeta((4\gamma+(\rho+\gamma)^2\zeta)(-\gamma+(\rho+\gamma)+\lambda_h)^2+2(-2\gamma+(\gamma-(\rho+\gamma))(\rho+\gamma)\zeta)(\gamma-(\rho+\gamma)-\lambda_h)\lambda_\ell+(\gamma-(\rho+\gamma))^2\zeta\lambda_\ell{}^2)+\zeta}}{\gamma 2\zeta(\lambda_h+\lambda_\ell)(\gamma-(\rho+\gamma)-\lambda_h)} \tag{40}$$

The firm value functions are,

$$v_\ell(z) = \frac{\bar{\lambda}}{\gamma + \rho\bar{\lambda}} e^z + \frac{1}{\rho(\nu+1)} e^{-\nu z} \tag{41}$$

$$v_h(z) = \frac{e^z + \lambda_h v_\ell(z)}{\rho + \lambda_h}, \tag{42}$$

where $\nu > 0$, the rate at which the option value is discounted, is given by

$$\nu \equiv \frac{\rho\bar{\lambda}}{\gamma}. \tag{43}$$

Proof. See Appendix B.1. □

So far, while the distributions of productivities z with initial distributions that have finite support also have finite support for $t < \infty$, stationary distributions all have asymptotically infinite relative support, that is $\bar{z} \to \infty$: the ratio of frontier to lowest productivity goes to infinity. In the ℓ state the growth rate is zero and z stays put, but in the h state the growth rate of z is positive. Given the Markov process for ℓ and h, there will be some agents who hit lucky streaks more than others, escape from the pack, and break away. Given a fixed barrier, the logic is similar to the linear or asymptotically linear Kesten processes bounded away from zero with affine terms, for which, under appropriate conditions, the asymptotic tail index can be explicitly computed in terms of the stationary distribution induced by the Markov process. In our model however, the adoption process introduces non-linear jumps that are not multiplicative in productivities, and which do not permit the simple characterization of the tail index. The endogenous absorbing adoption barrier (which acts like a reflecting barrier with stochastic jumps) complicates the analogy since one might think if the frontier is growing rapidly, the endogenous barrier could also move rapidly to keep up with the frontier. However, the incentives for adoption, which drive the speed of the moving barrier, are driven by the mean draw in productivity. Therefore, if the frontier diverges to infinity but the mean doesn't keep growing at the same rate, the frontier technology will diverge. As we will see in Section 2.4, if we strengthen the adoption process by allowing a positive fraction of adopters to leapfrog to the frontier, the multiplicative jump process generating the escape to infinity in relative productivities may in fact be contained.

This stationary equilibrium is unique and independent of the initial distribution (which will contrast the case of infinite initial support discussed in Proposition 3 which featured hysteresis and a continuum of stationary solutions). Figure 4 provides an example where $\gamma = .02, r = .06, \lambda_\ell = .01, \lambda_h = .03$, and $\zeta = 25$.

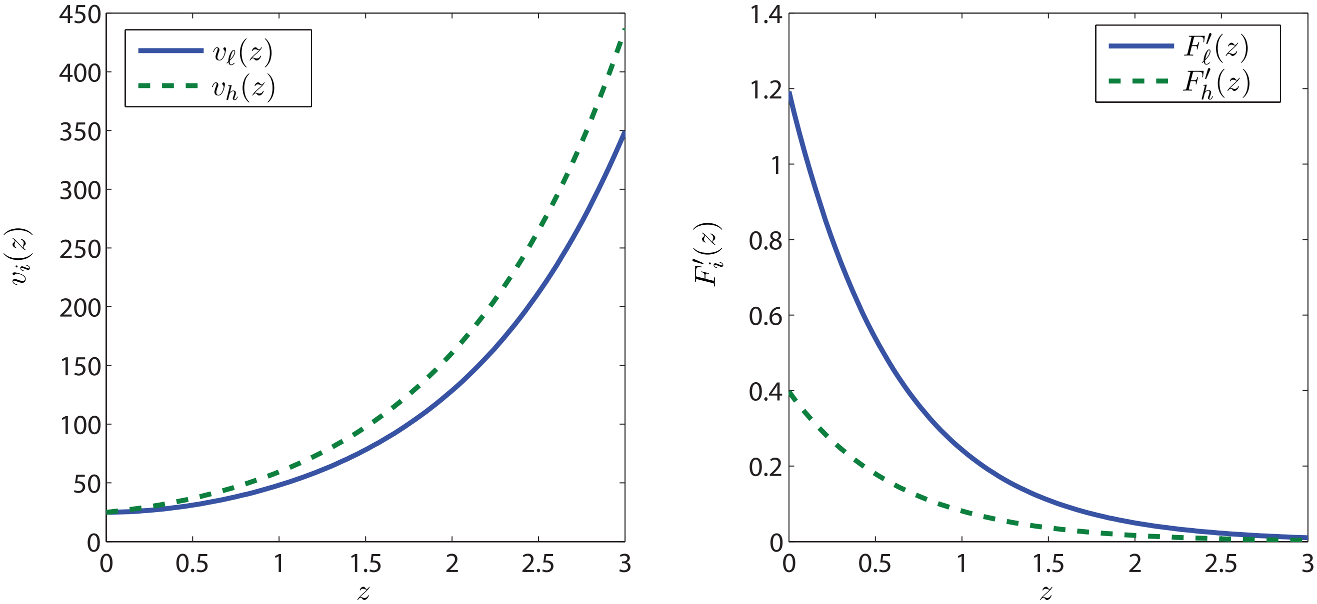

Figure 4: Normalized, Stationary Value Functions and PDFs for the Unbounded Cases

2.4 Stationary BGP with Bounded Relative Support

Proposition 1 shows that while the frontier remains finite for $t < \infty$, the ratio of the frontier productivity to the mean of the distribution is finite, but tends to infinity as $t \to \infty$.[14] This is because a diminishing, yet strictly positive number of firms keep getting lucky and grow at γ forever, but as the mass of agents with extremely high z is thinning out, it doesn't strongly effect the diffusion incentives and adoption probabilities of those with a low z (i.e., the mean expands slower than the frontier). Since there is no leapfrogging of these perpetually lucky firms, this process will continue forever. An alternative way to model within-firm productivity change is to assume that firms can leapfrog to the frontier with some probability. Such leapfrogging is a continuous time version of a quality-ladders model that keeps the frontier bounded.[15] The jumps can occur either for firms adopting through diffusion, or for innovating firms that succesfully leapfrog to the frontier, which can be viewed as positive spillovers from the frontier to innovators.

We model leapfrogging as an innovation that propels firms to the frontier of the productivity

[14]This result is robust to variations in the diffusion specification including assuming that adopting agents draw from the $F_h(z)$ distribution and start with a h type with $\kappa > 1$, which adds the maximum possible incentives to increase diffusion and compress the distribution.

[15]In a model of leapfrogging arrivals and a multiplicative step above the frontier, in continuous time the frontier would become infinite immediately, as in König, Lorenz, and Zilibotti (2012). Alternatively, it could be recast as a step-by-step innovation model in the spirit of Aghion, Akcigit, and Howitt (2013) with the same qualitative results.

distribution.[16] This major innovation with spillovers from the frontier can be achieved by all firms operating their existing technology, i.e., both $i = \ell$ and h types. However, since such an innovation is potentially disruptive, those firms that jump to the frontier become $\ell-$types and must wait for the Markov transition to h before they become innovators again.

To accomodate firms jumping to the frontier, we modify the model presented in Proposition 1 by adding an arrival rate for operating firms of jumps to the frontier, $\eta \geq 0$. See Figure 5 for a visualization of the stationary value functions.[17] To consider the case where it is adopters–rather than just innovating firms–who jump to the frontier, see Section 3.3 where an endogenous jump probability is chosen by adopting firms.

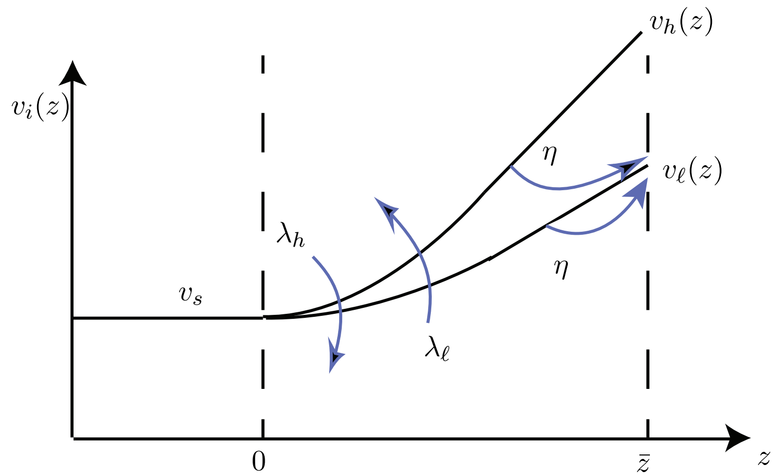

Figure 5: Normalized, Stationary Value Functions with Bounded Support

There *may* be a jump discontinuity in the right continuous cdf at $\bar{z}$. Due to right continuity of the cdf, the mass at the discontinuity $z = \bar{z}$ is:

$$\Delta_\ell = \lim_{\epsilon \to 0} \left(F_\ell(\bar{z}) - F_\ell(\bar{z} - \epsilon)\right) \tag{44}$$

$$\Delta_h = \lim_{\epsilon \to 0} \left(F_h(\bar{z}) - F_h(\bar{z} - \epsilon)\right) \tag{45}$$

Set $\kappa = 1$ for simplicity, and define $\mathbb{H}(z)$ as the Heaviside operator. Along with (20) to (22)

[16] Rather than being the autarkic process improvement of the γ growth, this is leap-frogging and may be viewed as a melding of innovation and diffusion, as the jump is a function of the existing productivity distribution. The intuition here is that while the stochastic, continuous growth of innovators is process improvement, these would be the sorts of innovations that are captured as new patents citing the prior-art. Note that here, unlike quality ladder models, their cannot be a multiplicative jump, or the absolute frontier would diverge to infinity as their would be some agent with an arbitrarily large number of jump arrivals in any positive time period.

[17] The assumption of a jump to the ℓ state at the frontier is only for analytical convenience, and this assumption can be changed with no qualitative differences. If some firms jumped to the h state at the frontier instead, then a right discontinuity in $F_h(z)$ would exist, $\Delta_h > 0$, and more care is necessary in solving the KF and integrating the value matching condition. With this specification, a possible downside is that $v_\ell(\bar{z}) < v_h(\bar{z} - \epsilon)$ for some set of small ϵ, and those firm would rather keep the lower z rather than innovate. Intuitively, the idea with this specification is similar to the notion of negative shocks to productivity due to experimentation, and that innovation can be disruptive to a firm. This simplification helps ensure that the values of jumps to the frontier remains identical for both agents, and hence all types have the same adoption threshold as demonstrated in Appendix A.2. If it is adopters, as nested by the endogenous θ probability in Section 3.3, who jump, this doesn't occur.

and (27), the following characterizes the necessary conditions for a stationary equilibrium,

$$\rho v_\ell(z) = e^z - g v_\ell'(z) + \lambda_\ell(v_h(z) - v_\ell(z))) + \eta(v_\ell(\bar{z}) - v_\ell(z)) \tag{46}$$
$$\rho v_h(z) = e^z + \lambda_h(v_\ell(z) - v_h(z)) + \eta(v_\ell(\bar{z}) - v_h(z)) \tag{47}$$
$$0 = gF_\ell'(z) + \lambda_h F_h(z) - \lambda_\ell F_\ell(z) - \eta F_\ell(z) + \eta \mathbb{H}\left(z - \bar{z}\right) + SF(z) - S \tag{48}$$
$$0 = \lambda_\ell F_\ell(z) - \lambda_h F_h(z) - \eta F_h(z) \tag{49}$$

While the value matching condition itself is identical to (33), the value from (A.29) becomes dependent on the value at the frontier

$$v(0) = \frac{1 + \eta v_\ell(\bar{z})}{\rho + \eta} = \int_0^{\bar{z}} v_\ell(z) \mathrm{d}F(z)^\kappa - \zeta \tag{50}$$

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 $\nu = \frac{\rho + \eta}{\gamma}\bar{\lambda}$. Furthermore, assume that the value of $F_\ell'(0)$ that solves (55) is larger than η/γ.

Proposition 2 (Stationary Equilibrium with a Bounded Frontier)**.** *With the maintained assumptions, a unique equilibrium with $\bar{z} < \infty$ exists with $g = \gamma$ where the stationary distribution is,*

$$F_\ell(z) = \frac{F_\ell'(0)}{(F_\ell'(0) - \eta/\gamma)(1 + \hat{\lambda})}(1 - e^{-\alpha z}) \tag{51}$$
$$F_h(z) = \hat{\lambda} F_\ell(z), \tag{52}$$

where

$$\alpha \equiv (1 + \hat{\lambda})(F_\ell'(0) - \eta/\gamma) \tag{53}$$
$$\bar{z} = \frac{\log(\gamma F_\ell'(0)/\eta)}{\alpha}. \tag{54}$$

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

$$\zeta + \frac{1}{\rho} = \frac{\gamma F_\ell'(0) \alpha \bar{\lambda} \left(-\frac{e^{-\nu \bar{z}}\left(-1 + e^{-\alpha \bar{z}}\right)\eta}{\rho \alpha \nu} + \frac{e^{\bar{z}}\eta\left(e^{-\alpha \bar{z}} - 1\right)}{-\alpha \rho} + \frac{-e^{-(\alpha + \nu)\bar{z}} + 1}{\nu(\alpha + \nu)} + \frac{-e^{\bar{z} - \alpha \bar{z}} + 1}{\alpha - 1} \right)}{\gamma(\gamma F_\ell'(0) - \eta)(\nu + 1)}. \tag{55}$$

The value functions for the firm are,

$$v_\ell(z) = \frac{\bar{\lambda}}{\gamma(1 + \nu)} \left(e^z + \frac{1}{\nu} e^{-\nu z} + \frac{\eta}{\rho} \left(e^{\bar{z}} + \frac{1}{\nu} e^{-\nu \bar{z}} \right) \right) \tag{56}$$
$$v_h(z) = \frac{e^z + (\lambda_h - \eta) v_\ell(z) + \eta v_\ell(\bar{z})}{\rho + \lambda_h}. \tag{57}$$

Proof. See Appendix B.2. □

In the above, note that $F_i(z)$ are continuous, and $\Delta_\ell = \Delta_h = 0$. This is because leapfrogging firms become type ℓ, at which point they immediately fall back in relative terms. As those firms falling back also jump back and forth to the h type, the $F_\ell(z)$ and $F_h(z)$ distribution smoothly mix, ensuring continuity. If such firms were added as h agents, there would be a jump discontinuity in the cdf exactly at $\bar{z}$. The Bellman equations, such as (56), now contain the value of production in perpetuity and the option value for both the current z and the frontier z.

The above α is an empirical "tail index" that can be estimated from a discrete set of data points. An example for $r = .06, \lambda_\ell = 0.01, \lambda_h = 0.03, \zeta = 25, \eta = 0.001, g = \gamma = .02$ is given in

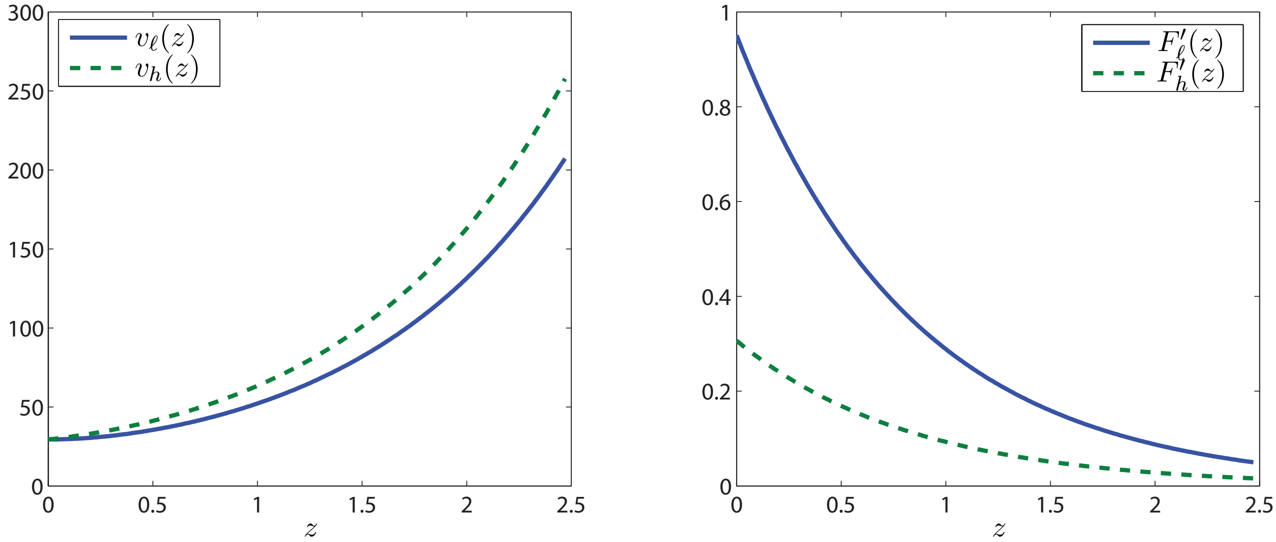

Figure 6: Exogenous $v_i(z)$, and $F'_i(z)$ with a Bounded Frontier

Figure 6. With these parameters, the frontier is $\bar{z} = 2.47$, or converting from logs, the frontier firm is approximately 12 times as efficient as the least productive firm at the adoption threshold. The α computed from this example is 1.2, close to the empirical Zipf's law.

As noted at the end of previous section, leapfrogging to the frontier by a positive mass of agents can contain the escape in relative productivities by lucky firms who get streaks of long sojourns in the high growth group h. Firms which had leapfrogged to the frontier may grow fast until they get a draw that slows them down by putting them back in the ℓ state. They will be overtaken by others who leapfrog to the frontier from anywhere in the productivity distribution and replenish it. This leapfrogging/quality ladder process prevents laggards from remaining as laggards forever. The distribution of relative productivities then remains bounded as the frontier acts as a locomotive in a relay race. Note that this locomotive process is similar to models of technology diffusion where the growth rate of adopters is an increasing function the distance to the frontier, unlike innovators with multiplicative growth in their productivity level (see Benhabib, Perla, and Tonetti (2014)). As adopters fall behind, their growth rate increases to match the growth rate of innovators, so relative productivities remain bounded.

Comparative statics on the $\bar{z}$ and α are shown in Figure 7 for changes in η, γ, ζ, and λ_h. For example, higher growth rates of innovators leads to a more distant technology frontier, but also to thinner tails. Alternatively, a higher cost of adoption leads to a more distant frontier, and thicker tails.

From (54), a relationship can be found that determines the range of the productivity distribution for any particular $F'_\ell(0)$:

$$\bar{z} = \frac{\eta + \lambda_h}{\eta + \lambda_\ell + \lambda_h} \frac{\log(F'_\ell(0)) - \log(\eta/\gamma)}{F'_\ell(0) - \eta/\gamma}. \tag{58}$$

2.5 Stationary BGP with Infinite Support

For completeness, if $\Phi(0, Z)$ has infinite support, $\Phi(t, Z)$ will converge to a stationary distribution as $t \to \infty$. A continuum of stationary distributions, each with its associated aggregate growth rates, are possible from different initial conditions. The intuition for this hysteresis is identical to that discussed in Appendix D and Perla and Tonetti (2014).[18]

[18] Uniqueness of related models with Geometric Brownian Motion is discussed in Luttmer (2012).

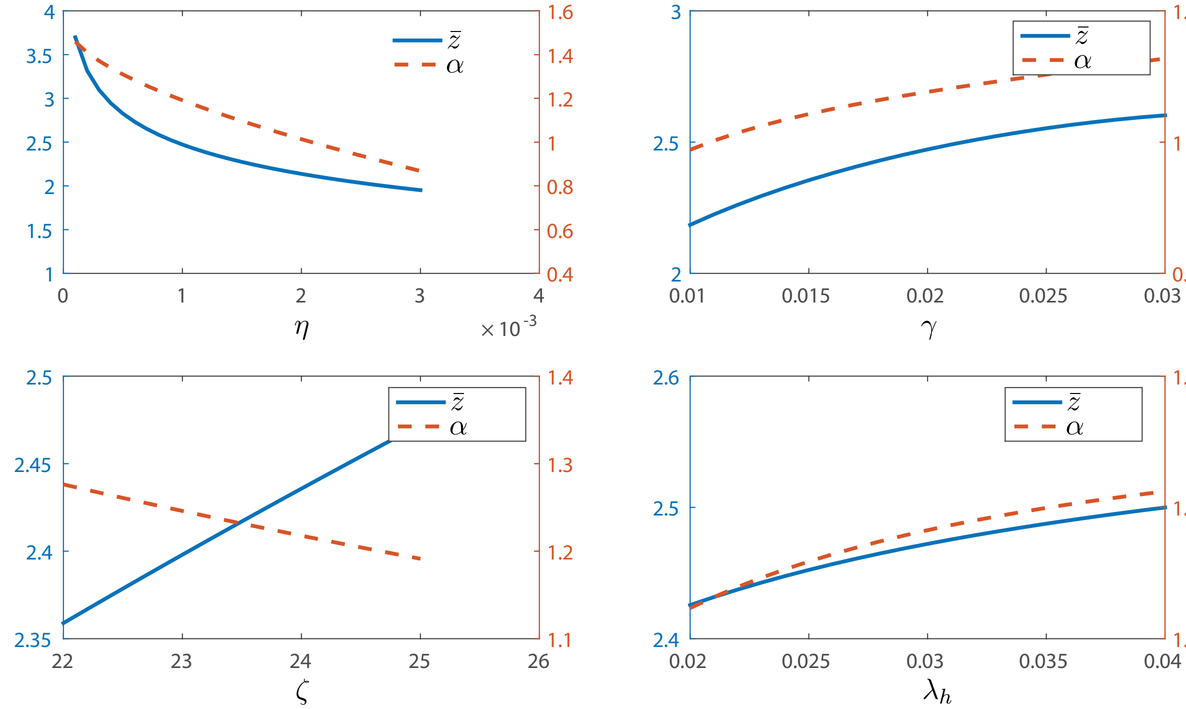

Figure 7: Comparative Statics with a Bounded Frontier

This section introduces an important difference from the setup used in Sections 2.3 and 2.4: the adoption technology will instead have firms copying both the type and productivity of the draw, rather than always starting in the ℓ state.[19] The normalized adoption distributions are then, $\hat{F}_\ell(z) \equiv F_\ell(z)$ and $\hat{F}_h(z) \equiv F_h(z)$, which can be verified to yield a proper CDF: $\hat{F}_\ell(\infty) + \hat{F}_h(\infty) = F_\ell(\infty) + F_h(\infty) = 1$. There can be no jumps to the frontier, or else the problem is undefined (i.e., require $\eta = 0$).

As $B(t) = \infty$ for all t, unlike in the examples with finite support, there is no requirement that $g = \gamma$ to ensure a stationary non-degenerate productivity distribution. However, $g \geq \gamma$ is neccesary to ensure that $S_h \geq 0$. Summarizing the stationary equations along with (21) to (25), (27) and (28),

$$v(0) = \frac{1}{\rho} = \underbrace{\int_0^\infty v_\ell(z)\mathrm{d}F_\ell(z) + \int_0^\infty v_h(z)\mathrm{d}F_h(z)}_{\text{Adopt both } i \text{ and } Z \text{ of draw}} - \zeta \tag{59}$$

$$0 = gF_\ell'(z) - \lambda_\ell F_\ell(z) + \lambda_h F_h(z) + (S_\ell + S_h)F_\ell(z) - S_\ell \tag{60}$$

$$0 = (g-\gamma)F_h'(z) - \lambda_h F_h(z) + \lambda_\ell F_\ell(z) + (S_\ell + S_h)F_h(z) - S_h \tag{61}$$

To characterize the continuum of stationary distributions, parameterize the set of solutions by a scalar α. Define the following as a function of the parameter α with an accompanying growth

[19] While an exactly correlated draw of the type and the productivity is not necessary here, see Technical Appendix C.3 for a proof that independent draws of Z and the innovation type for adopters has only degenerate stationary distributions in equilibrium.

rate, g ,

$$\vec{F}(z) \equiv \begin{bmatrix} F_\ell(z) \\ F_h(z) \end{bmatrix} \qquad v(z) \equiv \begin{bmatrix} v_\ell(z) \\ v_h(z) \end{bmatrix} \tag{62}$$

$$A \equiv \begin{bmatrix} \frac{1}{g} \\ \frac{1}{g-\gamma} \end{bmatrix} \qquad B \equiv \begin{bmatrix} \frac{\rho+\lambda_\ell}{g} & -\frac{\lambda_\ell}{g} \\ -\frac{\lambda_h}{g-\gamma} & \frac{\rho+\lambda_h}{g-\gamma} \end{bmatrix} \tag{63}$$

$$\varphi \equiv \sqrt{(\lambda_h - \alpha\gamma)^2 + 2\lambda_l (\alpha\gamma + \lambda_h) + \lambda_l{}^2} \tag{64}$$

$$C \equiv \begin{bmatrix} \frac{-\alpha\gamma+2\alpha g+\lambda_h-\lambda_l+\varphi}{2g} & \frac{\lambda_h}{g} \\ \frac{\lambda_l}{g-\gamma} & \frac{-\alpha\gamma+2\alpha g-\lambda_h+\lambda_l+\varphi}{2(g-\gamma)} \end{bmatrix} \tag{65}$$

$$D \equiv \begin{bmatrix} \frac{\lambda_h(g(\alpha\gamma-\lambda_h+\varphi-\lambda_l)+2\gamma\lambda_l)}{\gamma g(\alpha\gamma-\lambda_h+\varphi+\lambda_l)} \\ \frac{g(\alpha\gamma-\lambda_h+\varphi-\lambda_l)+2\gamma\lambda_l}{2\gamma(g-\gamma)} \end{bmatrix} \tag{66}$$

Proposition 3 (Stationary Equilibrium with Infinite Support)**.** *There exists a continuum of equilibria parameterized by $\alpha > 1$ for $g(\alpha)$ that satisfies*

$$\frac{1}{\rho} + \zeta = \int_0^\infty \left[\left[(\mathbf{I} + B)^{-1} \left(e^{\mathbf{I}z} + e^{-Bz} B^{-1} \right) A \right]^T e^{-Cz} D \right] \mathrm{d}z \tag{67}$$

and the parameter restrictions given in (B.54) *to* (B.57)*. The stationary distributions and the value functions are given by:*

$$\vec{F}(z) = \left(\mathbf{I} - e^{-Cz} \right) C^{-1} D \tag{68}$$

$$\vec{F}'(z) = e^{-Cz} D \tag{69}$$

$$v(z) = (\mathbf{I} + B)^{-1} \left(e^{\mathbf{I}z} + e^{-Bz} B^{-1} \right) A \tag{70}$$

By construction, α is also the tail index of the unconditional distribution $F(z) \equiv F_\ell(z) + F_h(z)$.

Proof. See Appendix B.3 □

The proof in Appendix B.3 provides (B.54) to (B.57) as a complicated set of parameter restrictions to ensure that $r > g$ for a general CRRA parameter and that the eigenvalues of B and C are positive. Positive eigenvalues of B ensure that value matching is defined, and the option value of diffusion asymptomatically goes to 0 for large z.

See Figures 8 and 9 for an example of infinite support with $\gamma = .01, r = .05, \lambda_\ell = .0004, \lambda_h = .03$, and $\zeta = 6.0$. In this equilibrium, $g = .029, \alpha = 2.5$ and $F_\ell(\infty) = .988$. The relationship between g and α is shown in Figure 9. The growth rate is decreasing as the tail becomes thinner, and the total number of agents in the h state increases.

Note the distinction between Propositions 1 and 3: the stationary equilibrium associated with initially finite vs. initially infinite support are different, even though the initially finite support case of Proposition 1 also has an asymptotically unbounded relative support.

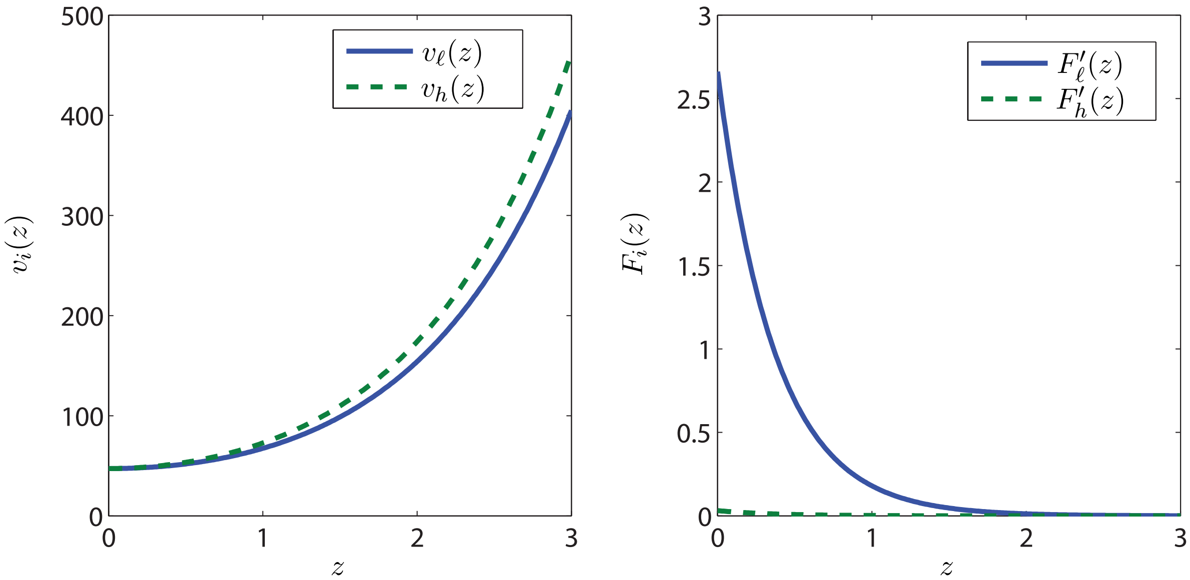

Figure 8: Exogenous $v_i(z)$, and $F'_i(z)$ with a Infinite Frontier

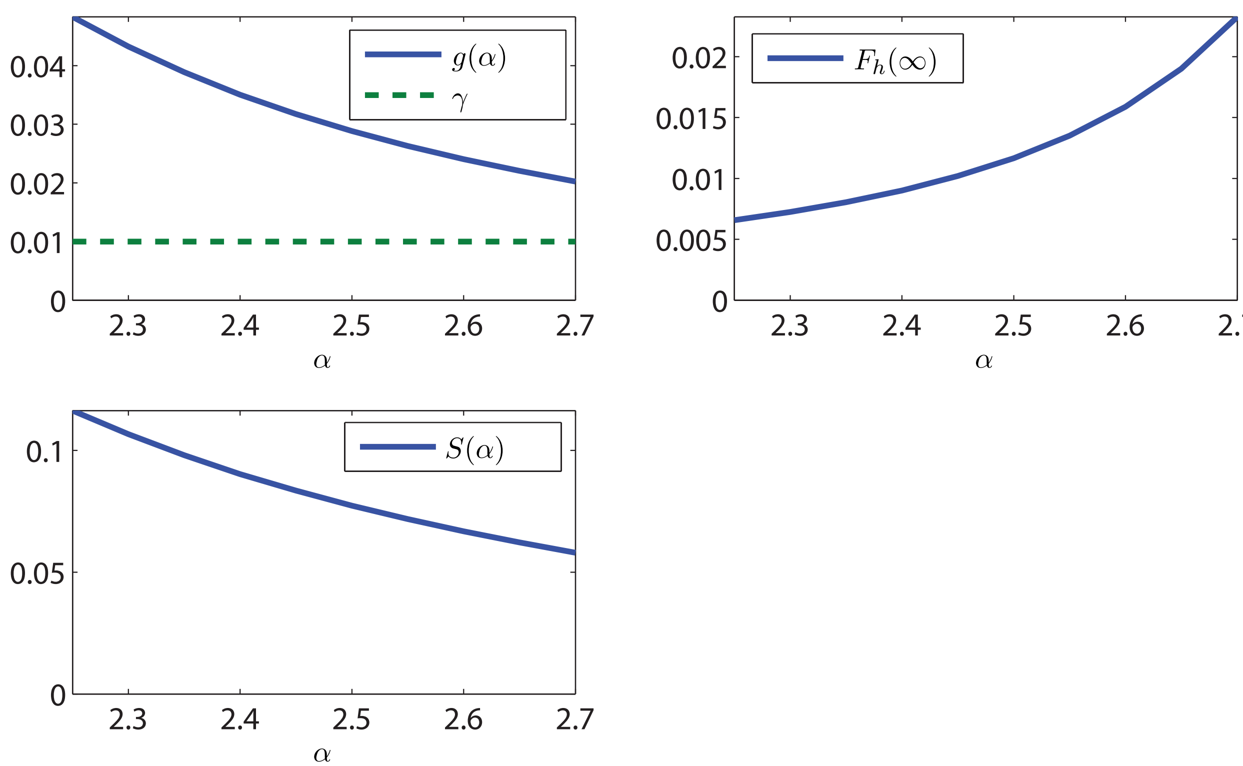

Figure 9: Growth rate as a Function of α with an Infinite Frontier

3 Endogenous, Stochastic Innovation

This section introduces endogenous innovation into the stochastic model with finite support. We assume that firms can control the drift of their innovation process, as in Atkeson and Burstein (2010) and Stokey (2014). In Section 3.2 we first assume that the arrival rate of jumps to the frontier, η, is zero in order to analyze the unbounded case with $\bar{z} \to \infty$. Then in Section 3.3 we move to the model with endogenously chosen innovation rates and jump probabilities and show that the frontier $\bar{z}$ is finite. To better understand the trade-off between innovation and adoption intensity, in Section 3.1 we introduce a simple form of partial excludability where adopters need to pay for part of the technology transfer.

We are modeling the innovation choice with no direct spillovers to ensure that it is as simple as possible, and orthogonal to the technology diffusion process. The interactions are between the tradeoffs in the firm's choices, rather than a coupled innovation and adoption technology.[20]

3.1 Licensing and Partial Excludability

Up to now, the firm providing the underlying technology was not able to prevent the copying—i.e., no excludability or intellectual property protection. To general this, consider partial enforcement where an adopting firm must pay a portion of the PDV of revenue to adopt the new technology. Instead of this being payed as a sequence of residual payment, the transfer will be done up front. There is no cost to the actual transfer of licensing.

Bargaining The timing is that the adopting firm firm pays the adoption cost, but upon the realization of the match, negotiations over licensing commence. These are negotiated with Nash Bargaining, and a bargaining power parameter of $\psi \in (0, 1]$, where $\psi = 1$ represents no excludability. The outside option of the adopting firm (i.e., the licensee) is to reject the offer and maintain their existing technology, i.e., $v(0)$. The outside option of the licensor is simply to reject and gain nothing. From standard Nash bargaining, with a total value of $v_\ell(z)$, let $\hat{v}$ be the proportion for the licensee and $v_\ell(z) - \hat{v}$ be the proportion for the licensor. Then Nash Bargaining problem is,

$$\arg\max_{\hat{v}} \left\{ (\hat{v} - v(0))^{\psi} (v_\ell(z) - \hat{v})^{1-\psi} \right\} \tag{71}$$

Solving for the surplus split, the value for the licensee matching a z firm is,

$$\hat{v}(z) = (1 - \psi)v(0) + \psi v_\ell(z) \tag{72}$$

While the licensor gains,

$$v_\ell(z) - \hat{v}(z) = (1 - \psi)(v_\ell(z) - v(0)) \tag{73}$$

Note from (72), if $\psi = 1$ then the licensee gains the entire value with $\hat{v}(z) = v_\ell(z)$. With risk-neutrality, the value matching condition integrates over surplus split from the possible matches.

Licensing and Flow Profits If both $\psi < 1$ and $\kappa \neq 1$, then the matching probabilities are distorted, and profits become a direct function of $F(z)$. To simplify the analysis and numerical algorithm, we will assume that if $0 < \psi < 1$, then $\kappa = 1$.[21] With this, since there is an equal probability of adopting from every licensor, the flow of licensees for the licensing firm gains is S.

[20] This is in contrast to approaches such as Chor and Lai (2013), where they are interested in the direct interaction with a dependent innovation process, with aggregate spillovers of knowledge.

[21] This assumption shuts down the interaction between κ choices and licensing, but is otherwise innocuous. The technical issue is that the marginal profits in (78) becomes a direct function of $F(z)$, which means that the HJBE and the KFE need to be solved concurrently–increasing the complexity of the numerical algorithms.

If $\gamma(0) = 0$ in equilibrium (as will be verified Section 3.2), then,

$$\pi(z) = e^z + (1-\psi) \underbrace{gF'(0)}_{\text{Licensee Flow}} \underbrace{\left(v_\ell(z) - \frac{1+\eta v_\ell(\bar{z})}{\rho+\eta}\right)}_{\text{Profits per Licensee}} \tag{77}$$

Differentiating and multiplying by e^{-z} we find a function for marginal profits,

$$e^{-z}\pi'(z) = 1 + (1-\psi)gF'(0)e^{-z}v_\ell'(z) \tag{78}$$

Here, if $\psi < 1$, then the profits are now a function g since (1) higher growth leads to more adopters—given a fixed $F(z)$; and (2) higher growth increases the licensor Nash bargaining threat of $v(0)$ by making it relatively more valuable to wait for the economy to grow before accepting a new technology. Given the equilibrium $F(z)$ and $v_\ell(z)$—which are themselves determined by g—these tensions help determine the relative profits from licensing. Also, note that as $z \to 0$, the profits from licensing disappear as the surplus of accepting a draw close to the adoption boundary go to 0. Consequently, $\pi(0) = 1$.

Value Matching Condition Note that as the surplus does not create any state dependence to the adoption costs, the smooth-pasting condition will be identical. However, since adopters may not gain the full surplus from the draw due to licensing costs, the value matching condition changes. Combining (33), (72) and (A.29),[22]

$$v(0) = \frac{1}{\rho} = \int_0^\infty \left[\psi v_\ell(z) + (1-\psi)v(0)\right] \mathrm{d}F(z)^\kappa - \zeta \tag{79}$$

Rearranging, the value matching condition is identical to that of (33) but with a proportional increase in the adoption cost,

$$v(0) = \frac{1}{\rho} = \int_0^\infty v_\ell(z)\mathrm{d}F(z)^\kappa - \frac{\zeta}{\psi} \tag{80}$$

To find the solution for any κ, use a conservation of the total surplus from the value matching condition with draws of $F(z)^\kappa$ distribution, the total surplus flows to the operating firms with distribution $F(z)$, and the flow (73). Then, for some $q(z)$ representing the distortion of the surplus flow to z agents, the flow of adopters licensing the firm's technology is $Sq(z)$, and the conversation of total licensing flows is

$$S\int \left(v_\ell(z) - v(0)\right) \mathrm{d}F(z)^\kappa = S\int \left(v_\ell(z) - v(0)\right) q(z)\mathrm{d}F(z) \tag{74}$$

With (31), we can reverse engineer the distortion as $q(z) \equiv \kappa F(z)^{\kappa-1}$. From (73), we find the flow profits including the direct value of production and the flow licensing given z as

$$\pi(z) \equiv \underbrace{e^z}_{\text{Production}} + \underbrace{\overbrace{\kappa F(z)^{\kappa-1} S}^{\text{Licensee Flow}} \overbrace{(1-\psi)(v_\ell(z) - v(0))}^{\text{Profits per Licensee}}}_{\text{Total Licensing Profits}} \tag{75}$$

The more general form of (78), using definitions (C.1) and (C.13) is then,

$$e^{-z}\pi'(z) \equiv 1 + (1-\psi)\kappa g F'(0)F(z)^{\kappa-1}\left(w_\ell(z) + (\kappa-1)F(z)^{-1}F'(z)e^{-z}\int_0^z w_\ell(\tilde{z})\mathrm{d}\tilde{z}\right) \tag{76}$$

Note that after substitution, the HJBE in $w_\ell(z)$ would then be a system of integro-differential equations, which would significantly increase the complexity of the problem. Hence, we are avoiding this since we don't feel the economics of the κ and ψ interaction are sufficiently interesting. Alternatively an algorithm in $v_\ell(z)$ space might work.

[22]The bargaining is over the ex-post surplus as ζ is a sunk adoption cost independent of the match.

3.2 Continuous Choice with the Finite, Unbounded, Frontier

Assume that, with a convex cost proportional to its current Z, a firm in the innovative state can choose its own growth rate $\gamma \geq 0$. Let $\chi > 0$ be the productivity of their R&D technology, and the cost quadratic in the growth rate γ. Adapting the equations in Section 2.3 after normalizing the innovation cost, then the set of equations is (20) to (23), (27), (28), (77), (77) and (80) with[23]

$$\rho v_\ell(z) = \pi(z) - g v_\ell'(z) + \lambda_\ell(v_h(z) - v_\ell(z)) \tag{84}$$

$$\rho v_h(z) = \max_{\gamma \geq 0} \left\{ \pi(z) - \underbrace{\tfrac{1}{\chi} e^z \gamma^2}_{\text{R\&D cost}} - \underbrace{(g-\gamma)}_{\text{Drift}} v_h'(z) + \lambda_h(v_\ell(z) - v_h(z)) \right\} \tag{85}$$

$$0 = g F_\ell'(z) + S F(z)^\kappa + \lambda_h F_h(z) - \lambda_\ell F_\ell(z) - S_\ell \tag{86}$$

$$0 = \underbrace{(g - \gamma(z))}_{\text{Drift}} F_h'(z) + \lambda_\ell F_\ell(z) - \lambda_h F_h(z) - S_h \tag{87}$$

With this setup, instead of all firms growing at rate γ exogenously, h-type firms are choosing a growth rate γ that is a function of their current productivity level, z. As the choice of γ is increasing in z in equilibrium, agents in the h state will end up crossing the endogenous adoption threshold, as shown in Appendix A.2, and thus the smooth pasting condition for h types is now necessary.

Proposition 4 (Stationary Equilibrium with Continuous Endogenous Innovation and Finite, Unbounded Support)**.** *A unique equilibrium exists with a growth rate as the solution to the cubic equation,*

$$g\left(g^2 + g(2\lambda_h + \lambda_\ell + 3\rho) + 2\rho(\lambda_h + \lambda_\ell + \rho)\right) = \chi(g + \lambda_h + \lambda_\ell + \rho). \tag{88}$$

The value function of the firm solves the following system of non-linear ODEs,

$$\rho v_h(z) = \pi(z) - g v_h'(z) + \tfrac{\chi}{4} e^{-z} v_h'(z)^2 + \lambda_h(v_\ell(z) - v_h(z)) \tag{89}$$

$$\rho v_\ell(z) = \pi(z) - g v_\ell'(z) + \lambda_\ell(v_h(z) - v_\ell(z)) \tag{90}$$

Given a solution to this system, the endogenous innovation choice is such that $\gamma(0) = 0$, $\lim_{z\to\infty} \gamma(z) = g$*, and*

$$\gamma(z) = \tfrac{\chi}{2} e^{-z} v_h'(z). \tag{91}$$

With this $\gamma(z)$*,* $F_i(z)$ *solves the KFEs in* (20)*,* (21)*,* (86) *and* (87)*.*

Proof. See Appendix C. The numerical method to compute the equilibrium is described in Technical **??**.

□

An example with $r = .06, \lambda_\ell = 0.01, \lambda_h = 0.03, \chi = 0.00212, \kappa = 1, \zeta = 25$ is given in Figure 10. The asymptotic growth rate in this example is calibrated to be 2.0%.

[23] To see the derivation for the KFE in (87), assume a stochastic process with a drift $\mu(z)$, then if the pdf is $f(t,z)$, the KFE comes from the adjoint to infinitesimal generator,

$$\boldsymbol{\partial}_t f(t,z) = -\boldsymbol{\partial}_z\left[\mu(z) f(t,z)\right] + \ldots \tag{81}$$

Integrating to get the cdf $F(t,z)$ and either using the fundamental theorem of calculus or interchanging the order of derviatives and integration

$$\boldsymbol{\partial}_t F(t,z) = -\mu(z) f(t,z) + \ldots \tag{82}$$

$$= -\mu(z)\boldsymbol{\partial}_z F(t,z) + \ldots \tag{83}$$

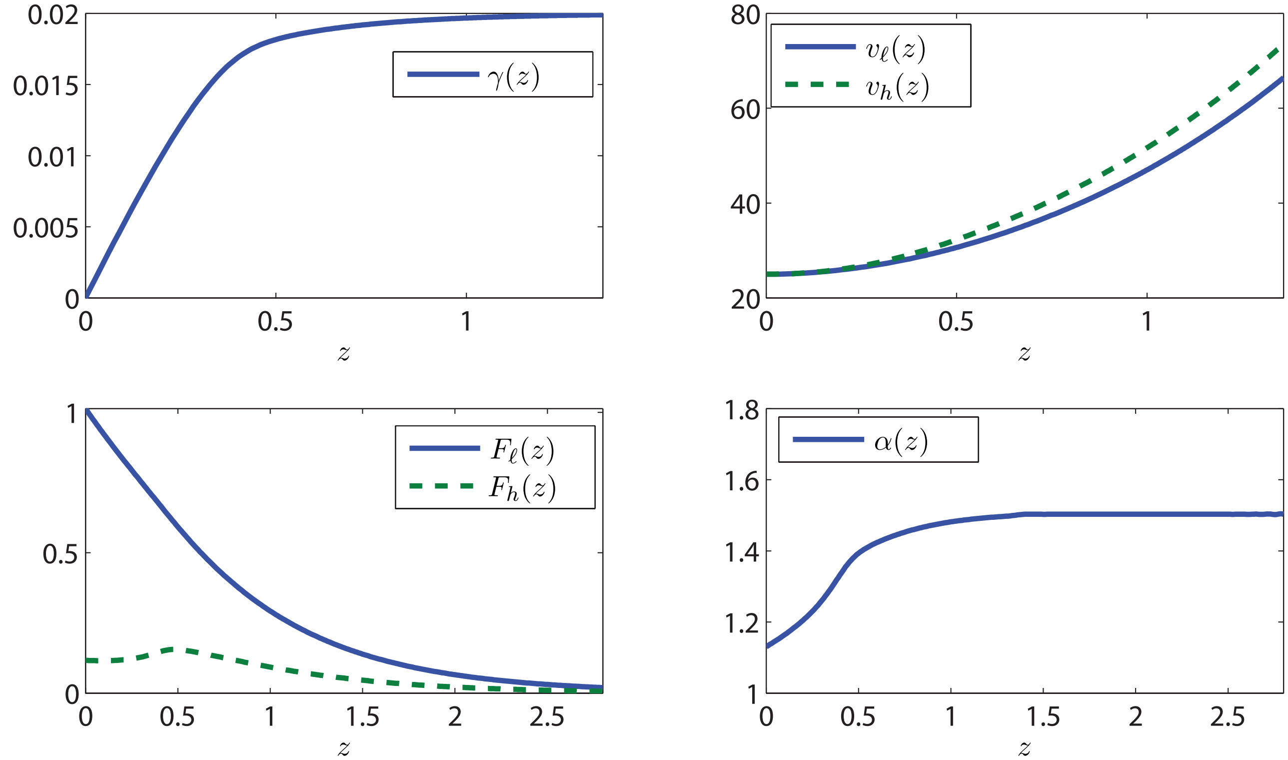

Figure 10: Endogenous $\gamma(z), v_i(z)$, and $F_i(z)$ with an Unbounded Frontier

In order to get a sense of the shape of the unconditional distribution, we define the "local" tail index

$$\alpha(z) \equiv \frac{F'(z)}{1 - F(z)} \tag{92}$$

In the standard log-log plot used for estimating power laws, as in Gabaix (2009), this $\alpha(z)$ would be the slope of the non-linear equation at z. Note that with this definition, the "local" tail index of a Pareto distribution is constant and equal to its true tail index. Furthermore, for any distribution with infinite support, the tail index is $\alpha \equiv \lim_{z\to\infty} \alpha(z)$. Figure 10 plots the local tail coefficient, converging to around 1.5. As the tail index is increasing, this shows that there is more productivity variability for firms with lower relative productivity.[24]

Growth Rates Conditional on Size As the intensity of innovation, $\gamma(z)$, is increasing in z, this would suggest that the larger and more productive firms tend to do the most research. While the economics and model are very different, this is related to Acemoglu, Aghion, and Zilibotti (2006) and Benhabib, Perla, and Tonetti (2014), both of which feature weakly increasing innovation in relative productivity. In our model, the intuition comes from an analysis of the option values, where agents closer to the endogenous adoption threshold have less incentive to invest in incremental productivity enhancement and accordingly decrease their endogenous investment in $\gamma(z)$.

Because $\gamma(z)$ is increasing, the growth rate, conditional on being a high type is also increasing in z. While this may appear to contradict Gibrat's law, and some modern evidence on non-Gibrat's growth as surveyed in Sutton (1997) and modeled in Luttmer (2007) and Arkolakis (2015), consider that 1) technology adoption is a key component growth of small firms but is not measured by γ,

[24]**TODO:** Explain the non-monotonicity in Figure 10

(2) we have left out the role of selection, which is extremely important for reconciling growth rates of small firms, and (3) the growth process is not an iid random walk, but has auto-correlation due to the Markov chain.

The first consideration is that smaller firms at the adoption threshold are growing rapidly (i.e., in fact, conditional on adoption in continuous time, they are growing at an infinite rate). Therefore, the model does have small firms tending to grow faster than larger firms. Here, we have simplified the model to ensure only a single adoption barrier exists and that firms make immediate productivity jumps. With more frictions and heterogeneity leading to a continuum of adoption barriers, the average growth rates might be more empirically plausible.

Second, many models investigating the empirics of Gibrat's law have emphasized that the higher growth rates for small firms is only conditional on survival. As small firms are more likely to exit, it implies that the average growth rate for smaller firms in the sample is higher. As we have purposely shut off exit in our model, this effect is not present. Davis, Haltiwanger, and Schuh (1996) find that when selection and mean reversion in stochastic processes are taken into account, the inverse relationship can disappear. However, Arkolakis (2015) discusses how the inverse relationship between size and growth tends to still exist even after selection, and describes how the Davis, Haltiwanger, and Schuh (1996) adjustment does not apply to random walks.

Finally, using our Markov chain process for growth, there is auto-correlation of growth rates for firms. This is in contrast to simply being a random walk, and hence the Davis, Haltiwanger, and Schuh (1996) results may still apply (where it doesn't in Arkolakis (2015), as discussed).

3.3 Continuous Innovation Choice with a Bounded Frontier

In this section, we combine all of the elements form earlier sections, and add additional features to endogenize all of the innovation and technology diffusion choices.

As discussed in Section 3.2, without leapfrogging $z \to \infty$. To endogenize the choice of leapfrogging, we allow firms that are adopting technology to jump to the frontier with probability $\theta \in [0, 1)$. Furthermore, just as we allow firms to choose their innovation rate γ, subject to a convex cost, we also allow firms to choose the probability of a jump to the frontier with a convex cost. The cost of choosing jump probability θ is $\frac{1}{\vartheta}\theta^2$. When a firm upgrades its technology through adoption, it can discover a state of the art invention and jump to the frontier. Finally, we allow firms to choose the degree of distortion in their draws (e.g., directed search) by choosing $\kappa > 0$ at a cost of $\frac{1}{\varsigma}\kappa^2$. We can adapt Sections 2.4 and 3.2 to include the same endogenous intensive innovation choice of γ, partial-excludability, a choice of θ, and a choice of κ. The system of equations becomes (20) to (25) and (77), with

$$\rho v_\ell(z) = \pi(z) - g v_\ell'(z) + \lambda_\ell(v_h(z) - v_\ell(z)) + \eta(v_\ell(\bar{z}) - v_\ell(z)) \tag{93}$$

$$\rho v_h(z) = \max_{\gamma \geq 0} \left\{ \pi(z) - \tfrac{1}{\chi} e^z \gamma^2 - (g - \gamma) v_h'(z) + \lambda_h(v_\ell(z) - v_h(z)) + \eta(v_\ell(\bar{z}) - v_h(z)) \right\} \tag{94}$$

$$v(0) = \frac{1 + \eta v_\ell(\bar{z})}{\rho + \eta} = \max_{\theta \geq 0,\, \kappa > 0} \left\{ (1 - \theta) \int_0^{\bar{z}} v_\ell(z) \mathrm{d}F(z)^\kappa + \theta v_\ell(\bar{z}) - \frac{1}{\psi} \left(\zeta + \tfrac{1}{\vartheta}\theta^2 + \tfrac{1}{\varsigma}\kappa^2 \right) \right\} \tag{95}$$

$$\begin{aligned} 0 &= g F_\ell'(z) + \lambda_h F_h(z) - \lambda_\ell F_\ell(z) - \eta F_\ell(z) \\ &\quad + (\eta + \theta(S_\ell + S_h)) \mathbb{H}(z - \bar{z}) + (1 - \theta)(S_\ell + S_h) F(z)^\kappa - S_\ell \end{aligned} \tag{96}$$

$$0 = (g - \gamma(z)) F_h'(z) + \lambda_\ell F_\ell(z) - \lambda_h F_h(z) - \eta F_h(z) - S_h \tag{97}$$

Note that when a particular firm is choosing θ or κ, it does not influence the θ or κ chosen by the other agents. Firms will take into account the effects of the aggregate "directed search" choice on $F_i(z)$, and as all adopting firms are a priori identical, each firm will choose the same θ, which will induce a $F_i(z)$ that is consistent with firm beliefs about $F_i(z)$ in equilibrium.

Proposition 5 (Stationary Equilibrium with Continuous Endogenous Innovation and Bounded Support)**.** *A continuum of equilibria exist, parameterized by a* $\bar{z}$*.*
The accompanying $\gamma(z), v_i(z)$*, and* $F_i(z)$ *solve the system of non-linear ODEs,*

$$\rho v_\ell(z) = \pi(z) - g v_\ell'(z) + \lambda_\ell(v_h(z) - v_\ell(z)) + \eta(v_\ell(\bar{z}) - v_\ell(z)) \tag{98}$$

$$\rho v_h(z) = \pi(z) - g v_h'(z) + \tfrac{\chi}{4} e^{-z} v_h'(z)^2 + \lambda_h(v_\ell(z) - v_h(z)) + \eta(v_\ell(\bar{z}) - v_h(z)) \tag{99}$$

$$v_\ell'(0) = v_h'(0) = 0 \tag{100}$$

$$v_\ell(0) = v_h(0) = \frac{1 + \eta v_\ell(\bar{z})}{\rho + \eta} \tag{101}$$

The endogenous innovation choice is such that $\gamma(0) = 0, \gamma(\bar{z}) \equiv g$ *and*

$$\gamma(z) = \tfrac{\chi}{2} e^{-z} v_h'(z) \tag{102}$$

Given the innovation intensity, the distribution solves (95) *to* (97) *where the chosen intensity of jumps to the frontier for adopting agents is,*

$$\theta = \frac{\psi\vartheta}{2} \int_0^{\bar{z}} v_\ell'(z) F(z)^\kappa \mathrm{d}z. \tag{103}$$

The directed technology diffusion choice solves the implicit equation

$$\kappa = \frac{-\varsigma\psi(1-\theta)}{2} \int_0^{\bar{z}} v_\ell'(z) \log(F(z)) F(z)^\kappa \mathrm{d}z \tag{104}$$

Proof. See Appendix C. A numerical method to solve for the continuum of equilibria is described in Technical **??**. □

In the limit, as $\eta \to 0$, the upper bound on g in (88) becomes the limiting case in Proposition 4.

Comparing to Stokey (2014), here the endogenous choice of γ is complicated by the option value. In the case of Proposition 4, the asymptotic γ as $z \to \infty$ becomes unique as the option value disappears, unlike with the bounded $\bar{z}$ in Proposition 5. Hence, a different distribution and $\bar{z}$ induce different growth option values, and give a continuum of self-fulfilling $\gamma(z)$.

Figure 13 plots the maximum growth rate of the set of admissible g as a function of η using (88) and with the same parameters as Figure 10. As $\eta \to 0$, the number of jumps to the frontier approaches 0, and the model in Section 3.3 asymptotically becomes that in Section 3.2. The intuition for a decreasing $\max(g(\eta))$ is that with more jumps to the frontier, the distribution becomes more compressed. As the growth rate of the frontier is determined by the autarkic innovation decision at $\bar{z}$, which takes into account the option value of diffusion, the more compressed the distribution, the lower the innovation rate, as in Figure 12.

3.4 Summary of Hysteresis and Multiplicity

The results of uniqueness of stationary equilibria are summarized in Table 1. Interestingly, hysteresis exists for the two opposites: infinite support with exogenous innovation, and bounded, finite support with endogenous innovation. Figure 13 demonstrates that the unbounded case is the limit of the bounded case (as the maximum at $\eta = 0$ is the g in the unbounded case from (88)), while the differences between Sections 2.3 and 2.5 show that the case of infinite and unbounded support are not identical.

The endogenous choice of technology adoption can give rise to hysteresis, in the sense that the productivity growth rates can depend on the initial distribution of productivities. The fatter the tail of the initial distribution, the richer will be the opportunities to adopt superior technologies,

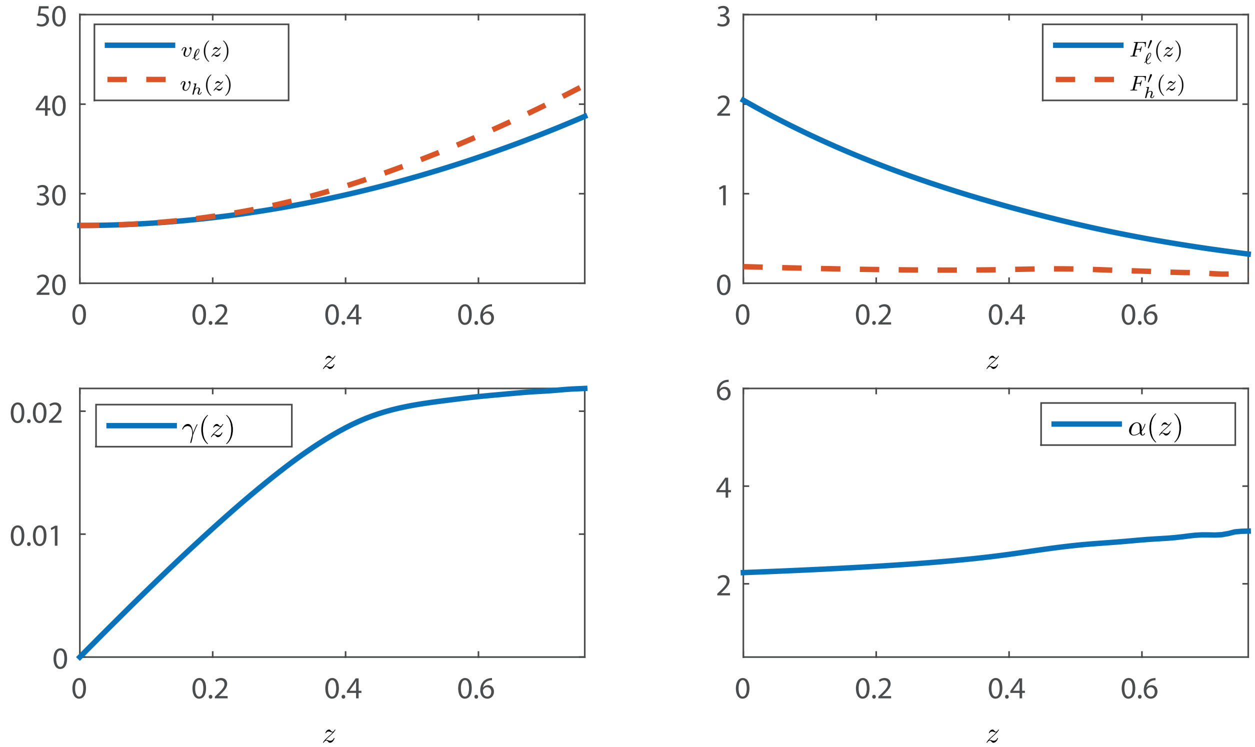

Figure 11: Endogenous $\gamma(z), v_i(z)$, and $F_i(z)$ with an Bounded Frontier

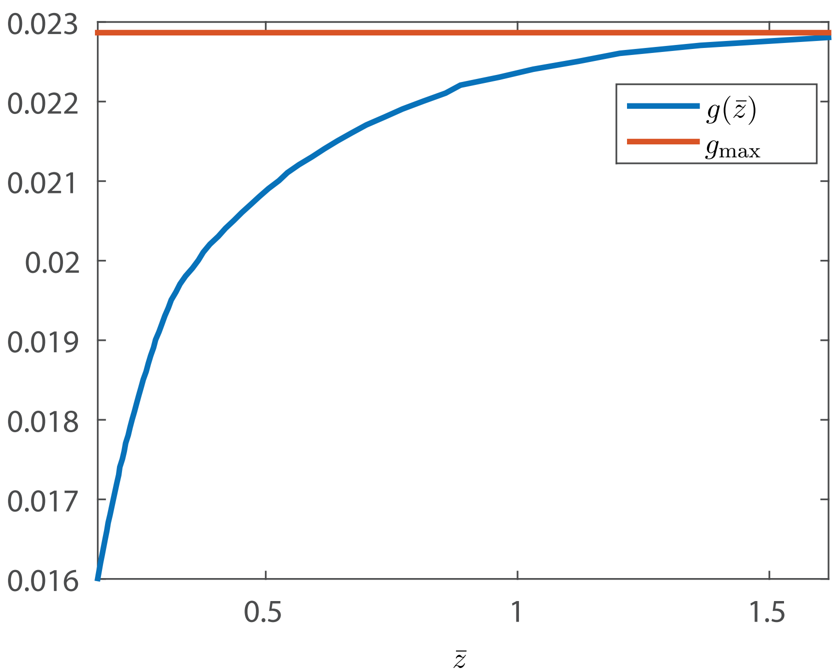

Figure 12: Equilibrium g as a function of $\bar{z}$

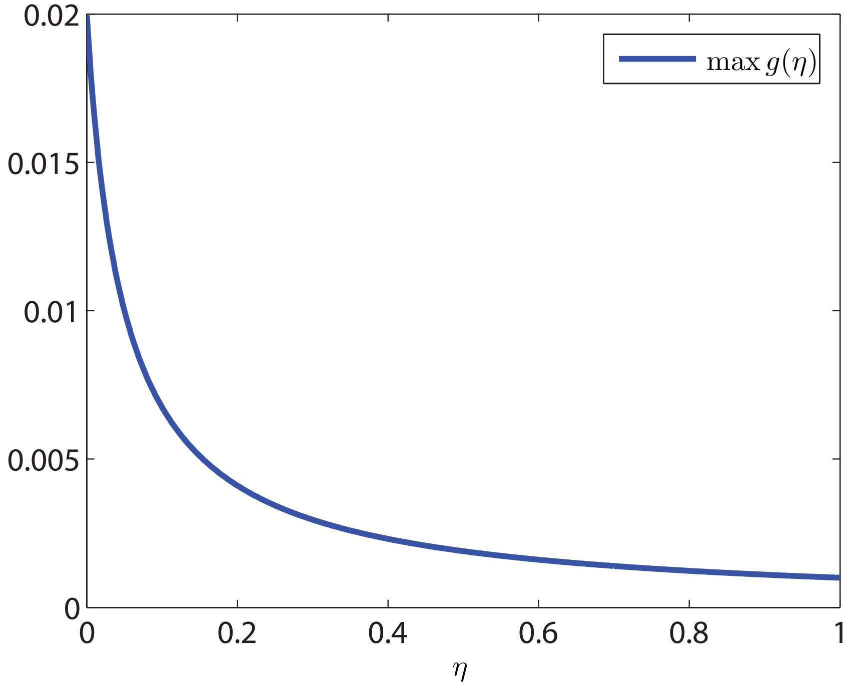

Figure 13: Maximum Equilibrium $g(\eta)$

Support	**Innovation**	
	Exogenous	*Endogenous*
Infinite	Hysteresis	Hysteresis
Unbounded	Unique	Unique
Bounded	Unique	Hysteresis

Table 1: Summary of Hysteresis and Uniqueness

and therefore the overall economy-wide growth rate will be higher. In the limit therefore the stationary distribution may well depend on the initial productivity distribution. With an initial fat-tailed distribution, if the adoption opportunities remain profitable, the limiting growth rate of the economy may forever exceed its growth rate from innovation alone.[25] This is the case for all models considered in Appendix D. It is also the case in Section 2 if the initial distribution has infinite support and states that provide access to adoption opportunities occur randomly, driven by Markov chain (Proposition 3). By contrast, if the initial productivity distribution has finite support, the growth rate of limiting distribution converges to the exogenous innovation rate and is independent of initial conditions (Proposition 1). Asymptotically, diffusion is no longer an independent driver of the economy-wide growth rate, even though some firms below the frontier choose to grow by adopting technology rather than by innovation. This is also the case with leap-frogging, or jumps to the frontier at an exogenous rate: in this case the ratio of the frontier productivity to the mean of the distribution remains finite (Proposition 2). The results do not change if the rate at which access to adoption opportunities can be endogenously chosen at some cost. However if the rate of leapfrogging opportunities can also be chosen, hysteresis re-emerges. The stationary distributions can then be parametrized by $\bar{z}$, the ratio of frontier productivity to the mean of the distribution (Proposition 5).[26] In this case the support and the position of the stationary distribution, as well as the growth rate, are parametrically determined $\bar{z}$.

4 Conclusion

Technology adoption, technological innovation and their interaction contribute to economic growth and to the evolution of the productivity distribution. In the various models that we study, growth rates and the distribution of productivities are endogenous, and they depend on the specification of the innovation and adoption processes, as well as the initial distribution of productivities available for adoption. In particular, whether adoption contributes to long-run growth in addition to innovation can depend on the properties (and tail index) of the initial distribution.[27] The specification of the innovation process (as GBM or a Markov process) can determine whether the asymptotic stationary distribution of relative productivities (the ratio of the frontier to the bottom) has finite support or not. We show in Propositions 1 and 2 that quality ladder type innovations driven by discrete Markov processes, under which a positive fraction of innovators leapfrog to the frontier, guarantee a stationary long run distribution with a finite support. We also study the problem of hysteresis: the possible multiplicity of stationary distributions that depend on initial conditions. Multiple stationary distributions occur in cases where (1) the support of the stationary distribution is infinite and adoption contributes to long-run growth (see Propositions 3, 6 and 7) or (2) when the intensity of the innovation is endogenously chosen with a positive probability of leap-frogging to the finite frontier of the stationary distribution (see Proposition 5 and Section 3.4). The various models of innovation and adoption processes that we have studied describe a rich set of long-run productivity distributions and of growth rates that may be useful for empirical work on the evolution of productivities.

[25] If we define hysteresis strictly as dependence of the limiting distribution on initial conditions, with deterministic growth $g \geq 0$ and for an initial distribution with bounded support, in the limiting distribution growth is driven only by innovation and all diffusion stops, but the position of the limiting distribution depends on the initial distribution.

[26] Note that unlike crefprop:stationary-equilibrium-exogenous-finite-jumps $\bar{z}$ can be jointly solved with $F_l'(0)$, using (54) and (55). $\bar{z}$ Proposition 5 $\bar{z}$ can be chosen parametrically to define the stationary distribution.

[27] See for example Propositions 1 to 3, 6 and 7 where the tail index is denoted as α.

References

Acemoglu, D., P. Aghion, and F. Zilibotti (2006): "Distance to Frontier, Selection and Economic Growth," *Journal of the European Economic Association*, 4(1), 37–74.

Acemoglu, D., U. Akcigit, N. Bloom, and W. R. Kerr (2013): "Innovation, Reallocation and Growth," NBER Working Papers 18993, National Bureau of Economic Research, Inc.

Aghion, P., U. Akcigit, and P. Howitt (2013): "What Do We Learn From Schumpeterian Growth Theory?," NBER Working Papers 18824, National Bureau of Economic Research, Inc.

Aghion, P., and P. Howitt (1992): "A Model of Growth Through Creative Destruction," *Econometrica*, 60(2), 323–51.

Alvarez, F. E., F. J. Buera, and R. E. Lucas (2008): "Models of Idea Flows," *NBER WP 14135*, pp. 1–12.

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

Arkolakis, C. (2015): "A Unified Theory of Firm Selection and Growth," *Quarterly Journal of Economics*, Forthcoming.

Atkeson, A., and A. T. Burstein (2010): "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 118(3), 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), 1–35.

Buera, F. J., and E. Oberfield (2014): "The Global Diffusion of Ideas," *WP*.

Chor, D., and E. L.-C. Lai (2013): "Cumulative Innovation, Growth and Welfare-Improving Patent Policy," CESifo Working Paper Series 4407, CESifo Group Munich.

Chu, A. C., G. Cozzi, and S. Galli (2014): "Stage-dependent intellectual property rights," *Journal of Development Economics*, 106(C), 239–249.

Davis, S. J., J. Haltiwanger, and S. Schuh (1996): " Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts," *Small Business Economics*, 8(4), 297–315.

Eaton, J., and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

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

Gabaix, X. (2009): "Power Laws in Economics and Finance," *Annual Review of Economics*, 1(1), 255–294.

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

——— (1993): *Innovation and Growth in the Global Economy*, vol. 1. The MIT Press.

Hopenhayn, H. A. (1992): "Exit, selection, and the value of firms," *Journal of Economic Dynamics and Control*, 16(3-4), 621–653.

Hsieh, C.-T., and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124(4), 1403–1448.

Jovanovic, B., and R. Rob (1989): "The Growth and Diffusion of Knowledge," *Review of Economic Studies*, 56(4), 569–582.

Klette, T. J., and S. Kortum (2004): "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112(5), 986–1018.

König, M., J. Lorenz, and F. Zilibotti (2012): "Innovation vs imitation and the evolution of productivity distributions," CEPR Discussion Papers 8843, C.E.P.R. Discussion Papers.

Kortum, S. S. (1997): "Research, Patenting, and Technological Change," *Econometrica*, 65(6), 1389–1420.

Lucas, R. E. (2009): "Trade and the Diffusion of the Industrial Revolution," *American Economic Journal: Macroeconomics*, 1(1), 1–25.

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

Luttmer, E. G. (2007): "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, 122(3), 1103–1144.

——— (2011): "On the Mechanics of Firm Growth," *Review of Economic Studies*, 78(3), 1042–1068.

——— (2012): "Eventually, noise and imitation implies balanced growth," (699).

——— (2014): "An Assignment Model of Knowledge Diffusion and Income Inequality," *Federal Reserve Bank of Minneapolis Working Paper*, (Working Paper 715).

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

Perla, J., and C. Tonetti (2014): "Equilibrium Imitation and Growth," *Journal of Political Economy*, 122(1), pp. 52–76.

Perla, J., C. Tonetti, and M. E. Waugh (2015): "Equilibrium Technology Diffusion, Trade, and Growth," Working Paper 20881, National Bureau of Economic Research.

Rivera-Batiz, L. A., and P. M. Romer (1991): "Economic Integration and Endogenous Growth," *The Quarterly Journal of Economics*, 106(2), 531–555.

Romer, P. M. (1986): "Increasing Returns and Long-run Growth," *Journal of Political Economy*, 94(5), 1002–37.

Romer, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98(5), S71–102.

Sampson, T. (2014): "Dynamic Selection: An Idea Flows Theory of Entry, Trade and Growth," *mimeo*.

Segerstrom, P. S., T. C. A. Anant, and E. Dinopoulos (1990): "A Schumpeterian Model of the Product Life Cycle," *American Economic Review*, 80(5), 1077–91.

Staley, M. (2011): "Growth and the diffusion of ideas," *Journal of Mathematical Economics*, 47(4-5), 470–478.

Stokey, N. L. (2009): *The Economics of Inaction*. Princeton Press.

——— (2014): "The Race Between Technology and Human Capital," *WP*.

Sutton, J. (1997): "Gibrat's Legacy," *Journal of Economic Literature*, 35(1), 40–59.

Syverson, C. (2011): "What Determines Productivity?," *Journal of Economic Literature*, 49(2), 326–65.

Appendix A General Proofs

A.1 Normalization

Normalizing the Productivity Distribution Define the normalized distribution of productivity, as the distribution of productivity relative to the endogenous adoption threshold $M(t)$:

$$\Phi_i(t, Z) \equiv F_i(t, \log(Z/M(t))) \tag{A.1}$$

Differentiating to obtain the pdf yields

$$\boldsymbol{\partial}_Z \Phi_i(t, Z) = \frac{1}{Z}\frac{\partial F_i(t, \log(Z/M(t)))}{\partial z} = \frac{1}{Z}\boldsymbol{\partial}_z F_i(t, z) \tag{A.2}$$

Differentiating (A.1) with respect to t and using the chain rule to obtain the transformation of the time derivative

$$\boldsymbol{\partial}_t \Phi_i(t, Z) = \frac{\partial F_i(t, \log(Z/M(t)))}{\partial t} - \frac{M'(t)}{M(t)}\frac{\partial_i F(t, \log(Z/M(t)))}{\partial z} \tag{A.3}$$

Using the definition $g(t) \equiv M'(t)/M(t)$ and the definition of z,

$$\boldsymbol{\partial}_t \Phi_i(t, Z) = \boldsymbol{\partial}_t F_i(t, z) - g(t)\boldsymbol{\partial}_z F_i(t, z) \tag{A.4}$$

Normalizing the Law of Motion Substitute (A.2) and (A.4) into (3) and (4),

$$\begin{aligned}\frac{\partial F_\ell(t, \log(Z/M(t)))}{\partial t} - g(t)\frac{\partial F_\ell(t, \log(Z/M(t)))}{\partial z} &= -\lambda_\ell F_\ell(t, \log(Z/M(t))) + \lambda_h F_h(t, \log(Z/M(t))) \\ &+ (S_\ell(t) + S_h(t))\,\hat{F}_\ell(t, \log(Z/M(t))) - S_\ell(t)\end{aligned} \tag{A.5}$$

$$\begin{aligned}\frac{\partial F_h(t, \log(Z/M(t)))}{\partial t} - g(t)\frac{\partial F_h(t, \log(Z/M(t)))}{\partial z} &= -\lambda_h F_h(t, \log(Z/M(t))) + \lambda_\ell F_\ell(t, \log(Z/M(t))) \\ &- \gamma\frac{Z}{Z}\frac{\partial F_h(t, \log(Z/M(t)))}{\partial z} \\ &+ (S_\ell(t) + S_h(t))\,\hat{F}_h(t, \log(Z/M(t))) - S_h(t)\end{aligned} \tag{A.6}$$

Using the definition of z and reorganizing yields the normalized KFEs,

$$\boldsymbol{\partial}_t F_\ell(t, z) = -\lambda_\ell F_\ell(t, z) + \lambda_h F_h(t, z) + g(t)\boldsymbol{\partial}_z F_\ell(t, z) + S(t)\hat{F}_\ell(t, z) - S_\ell(t) \tag{A.7}$$

$$\boldsymbol{\partial}_t F_h(t, z) = \lambda_\ell F_\ell(t, z) - \lambda_h F_h(t, z) + (g(t) - \gamma)\boldsymbol{\partial}_z F_h(t, z) + S(t)\hat{F}_h(t, z) - S_h(t) \tag{A.8}$$

Take (6) and (7) and substitute from (A.2),

$$S_\ell(t) = g(t)\boldsymbol{\partial}_z F_\ell(t, 0) \tag{A.9}$$

$$S_h(t) = (g(t) - \gamma)\boldsymbol{\partial}_z F_h(t, 0) \tag{A.10}$$

Normalizing the Value Function Define the normalized value of the firm as,

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

Rarrange and differentiating (A.11) with respect to t

$$\boldsymbol{\partial}_t V_i(t,Z) = M'(t)v_i(t,\log(Z/M(t))) - M'(t)\frac{\partial v_i(t,\log(Z/M(t))}{\partial z} + M(t)\frac{\partial v_i(t,\log(Z/M(t))}{\partial t} \tag{A.12}$$

Divide by $M(t)$ and use the definition of $g(t)$

$$\frac{1}{M(t)}\boldsymbol{\partial}_t V_i(t,Z) = g(t)v_i(t,z) - g(t)\boldsymbol{\partial}_z v_i(t,z) + \boldsymbol{\partial}_t v_i(t,z) \tag{A.13}$$

Differentiating (A.11) with respect to Z and rearranging

$$\frac{1}{M(t)}\boldsymbol{\partial}_Z V_i(t,Z) = \frac{1}{Z}\boldsymbol{\partial}_z v_i(t,z) \tag{A.14}$$

Divide (2) by $M(t)$ and then substitute from (A.13) and (A.14),

$$r\frac{1}{M(t)}V_h(t,Z) = \frac{Z}{M(t)} + \gamma\frac{M(t)}{M(t)}\frac{Z}{Z}\boldsymbol{\partial}_z v_h(t,z) + g(t)v_h(t,z) - g(t)\boldsymbol{\partial}_z v_h(t,z) + \lambda_h(v_\ell(t,z) - v_h(t,z)) + \boldsymbol{\partial}_t v_h(t,z) \tag{A.15}$$

Use (A.11) and the definition of z and rearrange,

$$(r - g(t))v_h(t,z) = e^z + (\gamma - g(t))\boldsymbol{\partial}_z v_h(t,z) + \boldsymbol{\partial}_t v_h(t,z) \tag{A.16}$$

Similarly, for (1)

$$(r - g(t))v_\ell(t,z) = e^z - g(t)\boldsymbol{\partial}_z v_\ell(t,z) + \lambda_\ell(v_h(t,z) - v_\ell(t,z)) + \boldsymbol{\partial}_t v_\ell(t,z) \tag{A.17}$$

Optimal Stopping Conditions Divide the value matching condition in (8) by $M(t)$,

$$\frac{V_i(t,M(t))}{M(t)} = \int_{M(t)}^{B(t)} \frac{V_\ell(t,Z)}{M(t)}\boldsymbol{\partial}_Z\hat{\Phi}_\ell(t,Z)\mathrm{d}Z + \int_{M(t)}^{B(t)} \frac{V_h(t,Z)}{M(t)}\boldsymbol{\partial}_Z\hat{\Phi}_h(t,Z)\mathrm{d}Z - \frac{M(t)}{M(t)}\zeta \tag{A.18}$$

Use the substitutions in (A.2) and (A.11), and a change of variable $z = \log(Z/M(t))$ in the integral, which implies $\mathrm{d}z = \frac{1}{Z}\mathrm{d}Z$. Note that the bounds of integration change to $[\log(M(t)/M(t)), \log(B(t)/M(t))] = [0, \bar{z}(t)]$

$$v_i(t,0) = \int_0^{\bar{z}(t)} v_\ell(t,z)\mathrm{d}\hat{F}_\ell(t,z) + \int_0^{\bar{z}(t)} v_h(t,z)\mathrm{d}\hat{F}_h(t,z) - \zeta \tag{A.19}$$

Evaluate (A.14) at $Z = M(t)$, and substitute this into (10) to give the smooth pasting condition,[28]

$$\boldsymbol{\partial}_z v_i(t,0) = 0 \tag{A.20}$$

[28] As a model variation, if the cost is proportional to Z, then the only change to the above conditions is that the smooth pasting condition becomes $\boldsymbol{\partial}_z v(t,0) = -\zeta$. This cost formulation has the potentially unappealing feature that the value is not monotone in Z, as firms close to the adoption threshold would rather have a lower Z to decrease the adoption cost for the same benefit.

A.2 Common Adoption Threshold for All Idiosyncratic States

Proof. This section proves under what general conditions heterogeneous firms will choose the same adoption threshold.

Allow for some discrete type i, and augment the state of the firm with an additional state x (which could be a vector or a scalar). Assume that there is some control u which controls the infinitesimal generator $\mathbb{Q}_u$ of the Markov process on type i and potentially x. Also assume that the agent can control the growth rate $\hat{\gamma}$ at some cost. The feasibility set of the controls is $(u, \hat{\gamma}) \in U(t, i, z, x)$ The cost of the controls for adoption and innovation have several requirements for this general property to hold: (a) The net value of searching, $v_s(t)$ is identical for all types i, productivities z, and additional state x, (b) The minimum of the cost function is 0 and in the interior of the feasibility set: $\min_{(\hat{\gamma},u)\in U(t,i,z,x)} c(t, z, \hat{\gamma}, i, x, u) = 0$, for all t, x, i and (c) the value of a jump to the frontier, $\bar{v}(t)$, is identical for all agent states (e.g. $\bar{v}(t) = v_\ell(t, \bar{z}(t)) = v_h(t, \bar{z}(t))$).[29]. Let the flow profits be a potentially type dependent $\pi_i(t, z, x)$, but require that $\pi(t, 0) \equiv \pi_i(t, 0, x)$ is identical for all i and x. Then the normalization of the firm's problem gives the following set of necessary conditions,

$$(r - g(t))v_i(t, z, x) = \max_{(\hat{\gamma},u)\in U(\cdot)} \Big\{\pi_i(t, z, x) - c(t, z, \hat{\gamma}, i, x, u) + (\hat{\gamma} - g)\frac{\partial v_i(t, z, x)}{\partial z} + \frac{\partial v_i(t, z, x)}{\partial t} + \mathbf{e_i} \cdot \mathbb{Q}_u \cdot v(t, z, x) + \eta(\bar{v}(t) - v_i(t, z, x))\Big\} \tag{A.21}$$

$$v_i(t, \underline{z}(t, i, x), x) = v_s(t) \tag{A.22}$$

$$\frac{\partial v_i(t, \underline{z}(t, i, x), x)}{\partial z} = 0 \tag{A.23}$$

Where $\underline{z}(t, i, x)$ is the normalized search threshold for type i and additional state x. To prove that these must be identical, we will assume that $\underline{z}(t, i, x) = 0$ for all types and additional states, and show that this leads to identical necessary optimal stopping conditions. Evaluating at $z = 0$,

$$v_i(t, 0, x) = v_s(t) \tag{A.24}$$

$$\frac{\partial v_i(t, 0, x)}{\partial z} = 0 \tag{A.25}$$

Note that (A.24) and (A.25) are identical for any i and x. Substitute (A.24) and (A.25) into (A.21)

$$(r - g(t))v_s(t) = \max_{(\hat{\gamma},u)\in U(\cdot)} \left\{\pi(t, 0) - c(t, z, \hat{\gamma}, i, x, u) + \mathbf{e_i} \cdot \mathbb{Q}_u \cdot v_s(t) + \eta(\bar{v}(t) - v_s(t)) + v_s'(t)\right\} \tag{A.26}$$

Since in order to be a valid intensity matrix, all rows in $\mathbb{Q}_u$ add to 0 for any u, the last term is 0 for any i or control u,

$$(r - g(t))v_s(t) = \max_{(\hat{\gamma},u)\in U(\cdot)} \left\{\pi(t, 0) - c(t, z, \hat{\gamma}, i, x, u) + v_s'(t) + \eta(\bar{v}(t) - v_s(t))\right\} \tag{A.27}$$

The optimal choice for any i or x is to minimize the costs of the $\hat{\gamma}$ and u choices. Given our assumption that the cost at the minimum is 0 and is interior,

$$(r - g(t))v_s(t) = \pi(t, 0) + v_s'(t) + \eta(\bar{v}(t) - v_s(t)) \text{ for all } i \tag{A.28}$$

Therefore, the necessary conditions for optimal stopping are identical for all i, x, z, confirming our guess. Furthermore, (A.28) provides an ODE for $v_s(t)$ based on aggregate $g(t)$ and $\bar{v}(t)$ changes.

[29] Without this requirement, firms may have differing incentives to "wait around" for arrival rates of jumps at the adoption threshold. A slightly weaker requirement is if the arrival rates and value are identical only at the threshold: $\eta(t, 0, \cdot)$ and $\bar{v}(t, 0, \cdot)$ are idiosyncratic states.

Solving this in a stationary environment gives an expression for v_s in terms of equilibrium g, $\bar{v}$ and the common $\pi(0)$,

$$v(0) \equiv v_s = \frac{\pi(0) + \eta\bar{v}}{r - g + \eta} \tag{A.29}$$

Furthermore, note that from (75), $\pi(0) = 1$ for all variations of the presented model

□

Appendix B Exogenous Markov Innovation

B.1 Stationary BGP with a Finite, Unbounded Technology Frontier

Proof of Proposition 1. Define the following to simplify notation,

$$\alpha \equiv (1 + \hat{\lambda})\frac{S}{g} \tag{B.1}$$

$$\hat{\lambda} \equiv \frac{\lambda_\ell}{\lambda_h} \tag{B.2}$$

$$\bar{\lambda} \equiv \frac{\lambda_\ell}{r - g + \lambda_h} + 1 \tag{B.3}$$

$$\nu = \frac{(r - g)\bar{\lambda}}{g} \tag{B.4}$$

See Technical Appendix C.4 for a proof that there are no bounded, finite equilibria for any $\kappa > 0$. For the solution to the $\kappa = 1$ case, take (35) and solve for $F_h(z)$

$$F_h(z) = \hat{\lambda}F_\ell(z) \tag{B.5}$$

Substitute into (34)

$$S = gF_\ell'(z) + \left(\hat{\lambda} + 1\right) SF_\ell(z) \tag{B.6}$$

Solve this as an ODE in $F_\ell(z)$, subject to the $F_\ell(0) = 0$ boundary condition in (20)

$$F_\ell(z) = \tfrac{1}{1+\hat{\lambda}}e^{-\alpha z} \tag{B.7}$$

We can check that if $\alpha > 0$ the right boundary conditions hold

$$\lim_{z\to\infty} \left(F_\ell(z) + F_h(z)\right) = 1 \tag{B.8}$$

Differentiating (B.7),

$$F_\ell'(z) = \tfrac{\alpha}{1+\hat{\lambda}}e^{-\alpha z} \tag{B.9}$$

With (B.5), the pdf for the unconditional distribution, $F(z)$,

$$F'(z) = \alpha e^{-\alpha z} \tag{B.10}$$

Solve (32) for $v_h(z)$

$$v_h(z) = \frac{e^z + \lambda_h v_\ell(z)}{r - g + \lambda_h} \tag{B.11}$$

Substituting into (18) gives the following ODE in $v_\ell(z)$

$$(r-g)v_\ell(z) = e^z + \lambda_h \hat{\lambda}\left(-v_\ell(z) + \frac{e^z + \lambda_h v_\ell(z)}{r-g+\lambda_h}\right) - g v_\ell'(z) \tag{B.12}$$

Using the constant definitions and simplifying

$$(r-g)v_\ell(z) = e^z - \frac{g v_\ell'(z)}{\bar{\lambda}} \tag{B.13}$$

Solve this ODE subject to the smooth pasting condition in (27) and simplify,

$$v_\ell(z) = \frac{\bar{\lambda}}{g+(r-g)\bar{\lambda}} e^z + \frac{1}{(r-g)(\nu+1)} e^{-z\nu} \tag{B.14}$$

Using the definitions of the constants and (B.14)

$$v_\ell(0) = \tfrac{1}{r-g} \tag{B.15}$$

Substitute (B.10), (B.14) and (B.15) into the value matching condition in (33) and simplify

$$\frac{1}{r-g} = \int_0^\infty \left[\frac{e^{z\left(\bar{\lambda}-\alpha-\frac{r\bar{\lambda}}{g}\right)}\alpha g}{(g-r)\left(-r\bar{\lambda}+g\left(\bar{\lambda}-1\right)\right)} + \frac{e^{z-z\alpha}\alpha\bar{\lambda}}{g+r\bar{\lambda}-g\bar{\lambda}}\right] \mathrm{d}z - \zeta \tag{B.16}$$

Evaluating the integral,

$$\zeta = \frac{\alpha\left(-r\bar{\lambda}+g\left(\bar{\lambda}-\alpha+1\right)\right)}{(g-r)\left(r\bar{\lambda}+g\left(\alpha-\bar{\lambda}\right)\right)(\alpha-1)} - \frac{1}{(r-g)(\nu+1)} - \frac{\bar{\lambda}}{g+r\bar{\lambda}-g\bar{\lambda}} \tag{B.17}$$

Substitute for α gives an implicit equation in S

$$0 = \zeta + \frac{g\left(\frac{1}{r-g} + \frac{\bar{\lambda}}{S-g+S\hat{\lambda}} - \frac{\bar{\lambda}}{S-g\bar{\lambda}+r\bar{\lambda}+S\hat{\lambda}}\right)}{-r\bar{\lambda}+g\left(\bar{\lambda}-1\right)} + \frac{1}{(r-g)(\nu+1)} \tag{B.18}$$

As $g=\gamma$ in equilibrium, only S is unknown. This equation is a quadratic in S, and can be analytically in terms of model parameters as,

$$S = \frac{\lambda_h\left(\zeta r(r+\lambda_h+\lambda_\ell) - \sqrt{\zeta((4g+r^2\zeta)(-g+r+\lambda_h)^2+2(-2g+(g-r)r\zeta)(g-r-\lambda_h)\lambda_\ell+(g-r)^2\zeta\lambda_\ell{}^2)}+\zeta g^2 2+\zeta(-g)(3r+2\lambda_h+\lambda_\ell)\right)}{2\zeta(\lambda_h+\lambda_\ell)(g-r-\lambda_h)} \tag{B.19}$$

From this S, α can be calculated through (B.1) and the the rest of the equilibrium follows. ☐

B.2 Bounded Support

Proof of Proposition 2. Define the following to simplify notation,

$$\alpha \equiv (1+\hat{\lambda})\frac{S-\eta}{g} \tag{B.20}$$

$$\hat{\lambda} \equiv \frac{\lambda_\ell}{\eta+\lambda_h} \tag{B.21}$$

$$\bar{\lambda} \equiv \frac{r-g+\lambda_\ell+\lambda_h}{r-g+\lambda_h} \tag{B.22}$$

$$\nu = \frac{r-g+\eta}{g}\bar{\lambda} \tag{B.23}$$

Solve for $F_h(z)$ in (49),

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

Substitute this back into (48) to get an ODE in F_ℓ

$$0 = gF_\ell'(z) + (S-\eta)(1+\hat{\lambda})F_\ell(z) + \eta H(z-\bar{z}) - S \tag{B.25}$$

Solve this ODE with the boundary condition $F_\ell(0) = 0$

$$F_\ell(z) = \begin{cases} \frac{S}{(S-\eta)(1+\hat{\lambda})}(1-e^{-\alpha z}) & 0 \le z < \bar{z} \\ \frac{S}{(S-\eta)(1+\hat{\lambda})}(1-e^{-\alpha \bar{z}}) & z = \bar{z} \end{cases} \tag{B.26}$$

This function is continuous at $z = \bar{z}$, and therefore so is $F_h(z)$. The unconditional distribution is,

$$F(z) = (1+\hat{\lambda})F_\ell(\bar{z}) \tag{B.27}$$

$$= \frac{S}{S-\eta}(1-e^{-\alpha z}) \tag{B.28}$$

Using the boundary condition that $F(\bar{z}) = 1$, and solving for $\bar{z}$ with the assumption that $S > \eta$,

$$\bar{z} = \frac{\log(S/\eta)}{\alpha} \tag{B.29}$$

The pdf of the unconditional distribution is,

$$F'(z) = \frac{\alpha S}{S-\eta} e^{-\alpha z} \tag{B.30}$$

$$\tag{B.31}$$

To solve for the value solve (47) for $v_h(z)$,

$$v_h(z) = \frac{e^z + (\lambda_h - \eta)v_\ell(z) + \eta v_\ell(\bar{z})}{r - g + \lambda_h} \tag{B.32}$$

Substitute into (46) and simplify

$$(r-g+\eta)v_\ell(z) = e^z + \eta v_\ell(\bar{z}) - \frac{g}{\bar{\lambda}} v_\ell'(z) \tag{B.33}$$

Solving (27) and (B.33) and simplifying,

$$v_\ell(z) = \frac{\bar{\lambda}}{g + (r+\eta-g)\bar{\lambda}} e^z + \frac{\eta}{r-g+\eta} v_\ell(\bar{z}) + \frac{1}{(r+\eta-g)(\nu+1)} e^{-z\nu} \tag{B.34}$$

Evaluate (B.34) at $\bar{z}$ and solve for $v_\ell(\bar{z})$,

$$v_\ell(\bar{z}) = \left(-\frac{\eta}{g-r} + 1\right)\left(\frac{e^{\bar{z}}\bar{\lambda}}{g + (\eta+r-g)\bar{\lambda}} + \frac{e^{-\nu\bar{z}}}{(\eta+r-g)(\nu+1)}\right) \tag{B.35}$$

Subtitute (B.35) into (B.34) to find an expression for $v_\ell(z)$

$$v_\ell(z) = \frac{\bar{\lambda}}{g(1+\nu)}\left(e^z + \frac{1}{\nu}e^{-\nu z} + \frac{\eta}{r-g}\left(e^{\bar{z}} + \frac{1}{\nu}e^{-\nu\bar{z}}\right)\right) \tag{B.36}$$

Substitute (B.30) and (B.36) into the value matching condition in (33) and evaluate the integral,

$$\zeta + \frac{1}{r-g} = \frac{S\alpha\bar{\lambda}\left(-\frac{e^{-\nu\bar{z}}\left(-1+e^{-\alpha\bar{z}}\right)\eta}{(-g+r)\alpha\nu} + \frac{e^{\bar{z}}\eta\left(e^{-\alpha\bar{z}}-1\right)}{\alpha(g-r)} + \frac{-e^{-(\alpha+\nu)\bar{z}}+1}{\nu(\alpha+\nu)} + \frac{-e^{\bar{z}-\alpha\bar{z}}+1}{\alpha-1}\right)}{g(S-\eta)(\nu+1)} \tag{B.37}$$

To find an implicit equation for the equilibrium S, take (B.37) and substitute for α and $\bar{z}$ from (B.20) and (B.29)

$$\zeta + \frac{1}{r-g} = \frac{S\bar{\lambda}\left(\hat{\lambda}+1\right)\left(\frac{-\left(\frac{S}{\eta}\right)^{-1-\frac{g\nu}{(S-\eta)\left(1+\hat{\lambda}\right)}}+1}{\nu\left(\nu+\frac{(S-\eta)\left(\hat{\lambda}+1\right)}{g}\right)} + g\left(\frac{1}{-g+(S-\eta)\left(\hat{\lambda}+1\right)} + \frac{\eta\left(-\frac{\left(\frac{S}{\eta}\right)^{-\frac{g\nu}{(S-\eta)\left(1+\hat{\lambda}\right)}}}{\nu} + \frac{\left(\frac{S}{\eta}\right)^{\frac{g}{(S-\eta)\left(\hat{\lambda}+1\right)}}\left(\eta+r-S+\hat{\lambda}(\eta+r-g-S)\right)}{-g+(S-\eta)\left(\hat{\lambda}+1\right)}\right)}{S(g-r)\left(\hat{\lambda}+1\right)}\right)\right)}{g^2(\nu+1)} \tag{B.38}$$

□

B.3 Stationary Stochastic Innovation Equilibrium with Infinite Support

Proof of Proposition 3. Define $\mathbf{0}, \mathbf{1}, \mathbf{I}$ as a vector of 0, 1, and the identity matrix and the following:

$$A \equiv \begin{bmatrix} \frac{1}{g} \\ \frac{1}{g-\gamma} \end{bmatrix} \qquad B \equiv \begin{bmatrix} \frac{r+\lambda_\ell-g}{g} & -\frac{\lambda_\ell}{g} \\ -\frac{\lambda_h}{g-\gamma} & \frac{r+\lambda_h-g}{g-\gamma} \end{bmatrix} \tag{B.39}$$

$$C \equiv \begin{bmatrix} \frac{gF_\ell'(0)+(g-\gamma)F_h'(0)-\lambda_\ell}{g} & \frac{\lambda_h}{g} \\ \frac{\lambda_\ell}{g-\gamma} & \frac{gF_\ell'(0)+(g-\gamma)F_h'(0)-\lambda_h}{g-\gamma} \end{bmatrix} \qquad D \equiv \begin{bmatrix} F_\ell'(0) \\ F_h'(0) \end{bmatrix} \tag{B.40}$$

$$\vec{F}(z) \equiv \begin{bmatrix} F_\ell(z) \\ F_h(z) \end{bmatrix} \qquad v(z) \equiv \begin{bmatrix} v_\ell(z) \\ v_h(z) \end{bmatrix} \tag{B.41}$$

Then the equilibrium conditions can be written as a linear set of ODEs:

$$v'(z) = Ae^z - Bv(z) \tag{B.42}$$

$$v'(0) = \mathbf{0} \tag{B.43}$$

$$\vec{F}'(z) = -C\vec{F}(z) + D \tag{B.44}$$

$$\vec{F}(0) = \mathbf{0} \tag{B.45}$$

$$\vec{F}(\infty) \cdot \mathbf{1} = 1 \tag{B.46}$$

$$v_\ell(0) = v_h(0) = \int_0^\infty \left(v(z)^T \cdot \vec{F}'(z)\right) \mathrm{d}z - \zeta \tag{B.47}$$

Solve these as a set of matrix ODEs, where e^{Az} is a matrix exponential. Start with (B.42) and (B.43) to get,[30]

[30] The equation $\vec{F}'(z) = AF[z] + b$ subject to $\vec{F}(0) = \mathbf{0}$ has the solution,

$$\vec{F}(z) = \left(e^{Az} - \mathbf{I}\right) A^{-1} b \tag{B.48}$$

The derivation of these results uses that $\int_0^T e^{tA}\mathrm{d}t = A^{-1}\left(e^{TA} - \mathbf{I}\right)$. With appropriate conditions on eigenvalues, this implies that $\int_0^\infty e^{tA}\mathrm{d}t = -A^{-1}$
Equations of the form,$v'(z) = Ae^z - B \cdot v(z)$ with the initial condition $v'(0) = \mathbf{0}$ have the solution,

$$v(z) = (I+B)^{-1}\left(e^{Iz} + e^{-Bz}B^{-1}\right) A \tag{B.49}$$

This derivation exploits commutativity, as both e^{Bz} and $(I+B)^{-1}$ can be expanded as power series of B.

$$v(z) = (I+B)^{-1}\left(e^{Iz} + e^{-Bz}B^{-1}\right)A \tag{B.50}$$

Evaluate at $z = 0$,

$$v(0) = B^{-1}A = \begin{bmatrix}1/(r-g) & 1/(r-g)\end{bmatrix} \tag{B.51}$$

Then (B.44) and (B.45) gives

$$\vec{F}(z) = \left(\mathbf{I} - e^{-Cz}\right)C^{-1}D \tag{B.52}$$

Take the derivative,

$$\vec{F}'(z) = e^{-Cz}D \tag{B.53}$$

For (B.50) and (B.52) to be well defined as $z \to \infty$, we have to impose parameter restrictions that constrain the growth rate g so that the eigenvalues of B and C are positive or have positive real parts. S_l and S_h are defined in equations (22) and (23) in terms of $F_l'(0)$ and $F_h'(0)$, C and B will have roots with positive real parts iff their determinant and their trace are strictly positive. For C it is straightforward to compute that the conditions for a positive trace and determinant are

$$S_l + S_h > \frac{(g-\gamma)\lambda_l + g\lambda_h}{(g-\gamma)+g} \tag{B.54}$$

$$S_h + S_l > \lambda_h + \lambda_l \tag{B.55}$$

and for B the corresponding conditions are

$$r > g > \gamma \tag{B.56}$$

$$r - g + \lambda_h + \lambda_l > 0 \tag{B.57}$$

With these conditions imposed, we can proceed to characterize the solutions to the value functions and the stationary distribution.

Evaluate (B.53) at $z = 0$,

$$\vec{F}'(0) = D \tag{B.58}$$

Take the limit of (B.52)

$$\vec{F}(\infty) = C^{-1}D \tag{B.59}$$

And (B.46) becomes

$$1 = C^{-1}D \cdot \mathbf{1} \tag{B.60}$$

We can check that, by construction, with the C and D defined by (B.40), (B.60) is fulfilled for any $F_\ell'(0), F_h'(0), \lambda_\ell$, and λ_h.

For $\vec{F}'(z)$ to define a valid pdf it is necessary for $\vec{F}'(z) > 0$ for all z. It can be shown, that for $C > 0$, the only $D > 0$ fulfilling this requirement is one proportional to the eigenvector associated with the dominant eigenvalue of C.[31] The unique constant of proportionality is determined by (B.60). The two eigenvectors of C fulfilling this proportionality are,

$$\nu_i \equiv \begin{bmatrix} -\frac{F_h'(0)\left(\gamma(g-\gamma)F_h'(0)\pm\sqrt{2(g-\gamma)\lambda_l\left(\gamma\left(g\left(F_h'(0)+F_l'(0)\right)-\gamma F_h'(0)\right)+g\lambda_h\right)+\left(\gamma\left(g\left(F_h'(0)+F_l'(0)\right)-\gamma F_h'(0)\right)-g\lambda_h\right)^2+(g-\gamma)^2\lambda_l{}^2}+\gamma g F_l'(0)-g\lambda \right.}{2g\lambda_l} \\ F_h'(0) \end{bmatrix} \tag{B.61}$$

Denote ν as the eigenvector with both positive elements–which is associated with the dominant eigenvalue–then as discussed above, $D \propto \nu$. Using (B.40) and (B.61), and noting the eigenvector has already been normalized to match the second parameter.

$$D = \nu \tag{B.62}$$

The 2nd coordinate already holds with equality by construction, for the first coordinate equating (B.40) and (B.61). Equating the first parameter and choosing the positive eigenvector,

$$F_\ell'(0) = -\frac{F_h'(0)\left(\gamma(g-\gamma)F_h'(0)-\sqrt{2(g-\gamma)\lambda_l\left(\gamma\left(g\left(F_h'(0)+F_l'(0)\right)-\gamma F_h'(0)\right)+g\lambda_h\right)+\left(\gamma\left(g\left(F_h'(0)+F_l'(0)\right)-\gamma F_h'(0)\right)-g\lambda_h\right)^2+(g-\gamma)^2\lambda_l{}^2}+\gamma g F_l'(0)-g\lambda_h+g\lambda_l-\gamma\lambda_l\right)}{2g\lambda_l} \tag{B.63}$$

Solving this equation for $F_\ell'(0)$ and choosing the positive root

$$F_\ell'(0) = \frac{F_h'(0)\lambda_h}{\gamma F_h'(0) + \lambda_\ell} \tag{B.64}$$

We can check that with the C and D defined by (B.40), (B.60) is fulfilled by construction. The value matching in (B.47) becomes,

$$\frac{1}{r-g} + \zeta = \int_0^\infty \left[\left[(I+B)^{-1}\left(e^{Iz} + e^{-Bz}B^{-1}\right)A\right]^T e^{-Cz} D\right] \mathrm{d}z \tag{B.65}$$

Note that if B has positive eigenvalues, then $\lim_{z\to\infty} v\left(z\right) = (1+B)^{-1}\left(e^z\right)A$. Therefore, as long as C has a minimal eigenvalue (defined here as α), strictly greater than one, the integral is defined.

The tail index of the unconditional distribution, $F(z) \equiv F_\ell(z)+F_h(z)$ can be calculated from the C matrix in (B.53). As sums of power law variables inherit the smallest tail index, the endogenous power law tail is minimum eigenvalue of C. After the substitution for $F_\ell'(0)$ from above, the smallest eigenvalue of C is

$$\alpha \equiv \frac{\left((g-\gamma)F_h'(0)-\lambda_l\right)\left(\gamma(g-\gamma)F_h'(0)+g\left(\lambda_h+\lambda_l\right)-\gamma\lambda_l\right)}{g(g-\gamma)\left(\gamma F_h'(0)+\lambda_l\right)} \tag{B.66}$$

Solving (B.66) for $F_h'(0)$ as a function of α,

$$F_h'(0) = \frac{g\left(\alpha\gamma-\lambda_h+\sqrt{\left(\lambda_h-\alpha\gamma\right)^2+2\lambda_l\left(\alpha\gamma+\lambda_h\right)+\lambda_l{}^2}-\lambda_l\right)+2\gamma\lambda_l}{2\gamma(g-\gamma)} \tag{B.67}$$

[31] Since $C > 0$ and irreducible (in this case off diagonals not zero), then by Perron-Frobenius it has a simple dominant real root α and an associated eigenvector $\nu > 0$. Hence, as $\vec{F}(0) = 0, F_\ell(\infty) + F_h(\infty) = 1$, and $\vec{F}'(z) > 0$, we have a valid pdf. This uniqueness of the ν solution only holds if the other eigenvector of C has a positive and negative coordinate, which always holds in our model.

Substituting for $F_i'(0)$ into C and D gives a function in terms of g and α,

$$C = \begin{bmatrix} \frac{-\alpha\gamma+2\alpha g+\lambda_h+\sqrt{(\lambda_h-\alpha\gamma)^2+2\lambda_l(\alpha\gamma+\lambda_h)+\lambda_l{}^2}-\lambda_l}{2g} & \frac{\lambda_h}{g} \\ \frac{\lambda_l}{g-\gamma} & \frac{-\alpha\gamma+2\alpha g-\lambda_h+\sqrt{(\lambda_h-\alpha\gamma)^2+2\lambda_l(\alpha\gamma+\lambda_h)+\lambda_l{}^2}+\lambda_l}{2(g-\gamma)} \end{bmatrix} \tag{B.68}$$

$$D = \begin{bmatrix} \frac{\lambda_h\left(g\left(\alpha\gamma-\lambda_h+\sqrt{(\lambda_h-\alpha\gamma)^2+2\lambda_l(\alpha\gamma+\lambda_h)+\lambda_l{}^2}-\lambda_l\right)+2\gamma\lambda_l\right)}{\gamma g\left(\alpha\gamma-\lambda_h+\sqrt{(\lambda_h-\alpha\gamma)^2+2\lambda_l(\alpha\gamma+\lambda_h)+\lambda_l{}^2}+\lambda_l\right)} \\ \frac{g\left(\alpha\gamma-\lambda_h+\sqrt{(\lambda_h-\alpha\gamma)^2+2\lambda_l(\alpha\gamma+\lambda_h)+\lambda_l{}^2}-\lambda_l\right)+2\gamma\lambda_l}{2\gamma(g-\gamma)} \end{bmatrix} \tag{B.69}$$

As in the example with Geometric Brownian Motion, there are multiple stationary equilibria. While both $F_i'(0)$ could conceivably parameterize a set of solutions for each g, they are constrained by the eigenvector proportionality condition, which ensures that the manifold of solutions is 1 dimensional.

□

Comparing (D.31) with (68) shows that the positivity of the tail index α is now equivalent to C having positive eigenvalues. For the decomposition of the option value, comparing (D.32) with (70) shows that positive eigenvalues of B ensure the option values in the vector $v(z)$ converges to 0 as z increases.

In Propositions 6 and 7 we characterized the stationary distributions in terms of the tail index of the initial distribution of productivities, given in Technical Appendix (D.19) explicitly by $\alpha = \kappa F'(0)$, a scalar. In this section, the stationary distribution is a vector $\vec{F}(z)$ solving (60) and (61) , a system of linear ODEs. If we define the unconditional distribution $F(z) \equiv F_\ell(z) + F_h(z)$, and if both $F_\ell(z)$ and $F_h(z)$ are power laws, any mixture of these distributions inherits the smallest (i.e.thickest) tail parameter (as discussed in Gabaix (2009)). Since there are now two dimensions of heterogeneity, the tail index, α, is defined as that of the unconditional distribution, $F(z)$. The ODE solution for the vector $\vec{F}(z)$ given in Proposition 3 by (68) will depend on the roots of C (both positive, see Appendix B.3). The smallest root of C, representing the slower rate of decay for both elements of $F(z)$, is the tail index α by the construction of (65).

Note that in Propositions 6 and 7 and TA.Proposition 1, the tail index is determined by the single initial condition $F'(0) > 0$, a scalar. In Proposition 3 the initial condition $F'(0)$ is a vector, so in principle this raises the possibility that the continuum of stationary equilibria could be two dimensional, parametrized by $F_\ell'(0) > 0$ and by $F_h'(0) > 0$. However as shown in Appendix B.3 this is not possible since the only initial condition that ensures that $F_\ell(z)$ and $F_h(z)$ remain positive and satisfy (18) and (19) is exactly the eigenvector of C corresponding to its dominant (Frobenius) eigenvalue. Since the eigenvector is determined only up to a multiplicative constant, the continuum of stationary distributions is therefore one dimensional. We use the smallest eigenvalue of C, defined as the tail index α, to solve for $F_h'(0)$, which then determines $F_\ell'(0)$ from the eigenvector restriction. This then allows us to obtain the expressions (65) and (66) in terms of parameters, α and g. Then value matching, (B.47) and (B.65), gives us expression (67) to define g in terms α, so we end up with a continuum of stationary equilibria parametrized by α.

Appendix C Endogenous Markov Innovation

Proof of Propositions 4 and 5 . Note that Section 3.3 nests Section 3.2 when $\eta = 0$

Nested Derivation of Stationary HBJE To create a stationary solution for the value function define a change of variables,[32]

$$w_i(z) \equiv e^{-z} v_i'(z) \tag{C.1}$$

From (27) and (28),

$$w_\ell(0) = w_h(0) = 0 \tag{C.2}$$

Differentiate (C.1) and reorganize ,

$$e^{-z} v_i''(z) = w_i'(z) + w_i(z) \tag{C.3}$$

Assuming an interior solution, take the first order necessary condition of the Hamilton-Jacobi-Bellman equation in (94), and reorganize

$$\gamma(z) = \tfrac{\chi}{2} e^{-z} v_h'(z) \tag{C.4}$$

Substitute this back into (94) to get a non-linear ODE,

$$(r-g)v_h(z) = \pi(z) - g v_h'(z) + \tfrac{\chi}{4} e^{-z} v_h'(z)^2 + \lambda_h (v_\ell(z) - v_h(z)) + \eta \left(v_\ell(\bar{z}) - v_h(z)\right) \tag{C.5}$$

Differentiate (93),

$$(r-g)v_\ell'(z) = \pi'(z) - g v_\ell''(z) + \lambda_\ell (v_h'(z) - v_\ell'(z)) - \eta v_\ell'(z) \tag{C.6}$$

As before, for simplicity, assume that if $\psi < 1$, then $\kappa = 1$. Multiply (C.6) by e^{-z} and use (78), (C.1) and (C.3).[33]

$$(r + \lambda_\ell + \eta - (1-\psi) g F'(0)) w_\ell(z) = 1 - g w_\ell'(z) + \lambda_\ell w_h(z) \tag{C.7}$$

Note that using (C.3),

$$e^{-z} \boldsymbol{\partial}_z \left(e^{-z} v_h'(z)^2\right) = 2 e^{-z} v_h''(z) e^{-z} v_h'(z) - \left(e^{-z} v_h'(z)\right)^2 \tag{C.8}$$

$$= 2 w_h(z) w_h'(z) + w_h(z)^2 \tag{C.9}$$

Differentiate (C.5), multiply by e^{-z}, and use (78), (C.1), (C.3) and (C.9)

$$(r + \lambda_h + \eta) w_h(z) = 1 - \left(g - \tfrac{\chi}{2} w_h(z)\right) w_h'(z) + \left(\lambda_h + (1-\psi) g F'(0)\right) w_\ell(z) + \tfrac{\chi}{4} w_h(z)^2 \tag{C.10}$$

From (C.4),

$$\gamma(z) = \tfrac{\chi}{2} w_h(z) \tag{C.11}$$

$$g \equiv \tfrac{\chi}{2} w_h(\bar{z}) \tag{C.12}$$

Define the integrated marginal utility from 0 to z as,

$$\hat{w}_i(z) \equiv \int_0^z v_i'(\hat{z}) \mathrm{d}\hat{z} = \int_0^z e^{\hat{z}} w_i(\hat{z}) \mathrm{d}\hat{z} \tag{C.13}$$

Integrate (C.1) with the initial value from (95) and use (C.13) to get,

$$v_i(z) = v(0) + \hat{w}_i(z) \tag{C.14}$$

Substitute (77) and (C.14) into (A.29) and rearrange to get an expression for $v(0)$ in terms of $\hat{w}_\ell$ and intrinsics,

$$v(0) = \frac{1 + \eta v_\ell(\bar{z})}{r - g + \eta} = \frac{1 + \eta \hat{w}_\ell(\bar{z})}{r - g} \tag{C.15}$$

[32] An alternative approach to solving the HJBE would be to use up/downwind methods of **?**.

[33] The general form of $\kappa \neq 1$ and $\psi < 1$ is covered in (76). We are avoiding this due to numerical difficulties rather than anything intrinsic in the model.

Value Matching with Endogenous choice of θ and κ: A change to $w_i(z)$ space will also be useful for simplifying integrals. Note that,[34]

$$\int_0^{\bar{z}} v_\ell(z)\mathrm{d}F(z)^\kappa = v(0) + \int_0^{\bar{z}} e^z w_\ell(z)\left(1 - F(z)^\kappa\right)\mathrm{d}z \tag{C.17}$$

And expanding when $\bar{z} < \infty$,

$$\int_0^{\bar{z}} v_\ell(z)\mathrm{d}F(z)^\kappa = v_\ell(\bar{z}) - \int_0^{\bar{z}} e^z w_\ell(z)F(z)^\kappa \mathrm{d}z \tag{C.18}$$

Take the value matching condition for the choice of the idiosyncratic $\hat{\theta}$ and $\hat{\kappa}$ given equilibrium θ and κ choices of the other firms.

$$v(0) = \max_{\hat{\theta}\geq 0,\, \hat{\kappa}>0}\left\{(1-\hat{\theta})\int_0^{\bar{z}} v_\ell(z)\mathrm{d}F(z)^{\hat{\kappa}} + \hat{\theta}v_\ell(\bar{z}) - \frac{1}{\psi}\left(\zeta + \tfrac{1}{\vartheta}\hat{\theta}^2 + \tfrac{1}{\varsigma}\hat{\kappa}^2\right)\right\} \tag{C.19}$$

Use (C.14) and (C.17)

$$v(0) = \max_{\hat{\theta}\geq 0,\, \hat{\kappa}>0}\left\{(1-\hat{\theta})\left(v(0) + \int_0^{\bar{z}} e^z w_\ell(z)(1-F(z)^{\hat{\kappa}})\mathrm{d}z\right) + \hat{\theta}\left(v(0) + \hat{w}_\ell(\bar{z})\right) - \frac{1}{\psi}\left(\zeta + \tfrac{1}{\vartheta}\hat{\theta}^2 + \tfrac{1}{\varsigma}\hat{\kappa}^2\right)\right\} \tag{C.20}$$

Simplify,

$$0 = \max_{\hat{\theta}\geq 0,\, \hat{\kappa}>0}\left\{(1-\hat{\theta})\int_0^{\bar{z}} e^z w_\ell(z)(1-F(z)^{\hat{\kappa}})\mathrm{d}z + \hat{\theta}\hat{w}_\ell(\bar{z}) - \frac{1}{\psi}\left(\zeta + \tfrac{1}{\vartheta}\hat{\theta}^2 + \tfrac{1}{\varsigma}\hat{\kappa}^2\right)\right\} \tag{C.21}$$

In the case of $\theta = 0$ this simplifies to,

$$0 = \max_{\hat{\kappa}>0}\left\{\int_0^{\bar{z}} e^z w_\ell(z)(1-F(z)^{\hat{\kappa}})\mathrm{d}z - \frac{1}{\psi}\left(\zeta + \tfrac{1}{\varsigma}\hat{\kappa}^2\right)\right\} \tag{C.22}$$

In the case of $\theta > 0$, (C.21) simplifies to,

$$0 = \max_{\hat{\theta}\geq 0,\, \hat{\kappa}>0}\left\{\hat{w}_\ell(\bar{z}) - (1-\hat{\theta})\int_0^{\bar{z}} e^z w_\ell(z)F(z)^{\hat{\kappa}}\mathrm{d}z - \frac{1}{\psi}\left(\zeta + \tfrac{1}{\vartheta}\hat{\theta}^2 + \tfrac{1}{\varsigma}\hat{\kappa}^2\right)\right\} \tag{C.23}$$

Crucially, if the firm chooses a $\hat{\theta} \neq \theta$, they are infinitesimal and have no influence on the value or equilibrium distributions. Taking the first order condition of (C.23) with respect to $\hat{\theta}$ and then letting $\hat{\theta} = \theta$ in equilibrium gives,

$$\theta = \frac{\psi\vartheta}{2}\int_0^{\bar{z}} e^z w_\ell(z)F(z)^\kappa \mathrm{d}z \tag{C.24}$$

Take the first order condition of (C.21) and equate $\hat{\kappa} = \kappa$ in the economy. Note that $F(z)^\kappa = \exp(\kappa \log F(z))$, so that $\boldsymbol{\partial}_\kappa F(z)^\kappa = \log(F(z))F(z)^\kappa$ and assume conditions to differentiate under the integral

$$\kappa = \frac{-\varsigma\psi(1-\theta)}{2}\int_0^{\bar{z}} e^z w_\ell(z)\log(F(z))F(z)^\kappa \mathrm{d}z \tag{C.25}$$

This is an implicit equation in κ. Note that as $0 < F(z) < 1$, $\log(F(z)) < 0$, so the sign of this term is correct to ensure a positive κ.

[34] These come out of using integration by parts on the calculation of the expectation. For example, if $F(z)$ is the CDF for a random variable Z with minimum and maximum support $\underline{z}$ and $\bar{z}$, then the following holds for any reasonable $h(z)$,

$$\mathbb{E}\left[h(Z)\right] = \int_{\underline{z}}^{\bar{z}} h'(z)(1-F(z))\mathrm{d}z + h(\underline{z}) \tag{C.16}$$

KFE and Value Matching From (96), for $z < \bar{z}$ the KFE is,

$$0 = gF_\ell'(z) + \lambda_h F_h(z) - \lambda_\ell F_\ell(z) - \eta F_\ell(z) + (1-\theta)(S_\ell + S_h)F(z)^\kappa - S_\ell,\ z < \bar{z} \tag{C.26}$$

In the limit as $z \to \bar{z}$, we know both $\gamma(z)$ and $F(z)$ are continuous. Assume that $\lim_{z\to\bar{z}} F_h'(z) < \infty$ and use $g - \gamma(\bar{z}) = 0$ in (21) to (23) and (97) to get the system of equations,[35]

$$0 = (\lambda_h + \eta)F_h(\bar{z}) - \lambda_\ell F_\ell(\bar{z}) + gF_h'(0) \tag{C.27}$$
$$1 = F_\ell(\bar{z}) + F_h(\bar{z}) \tag{C.28}$$

Solving gives a boundary condition for $F(\bar{z})$ for a given $F'(0)$

$$F_\ell(\bar{z}) = \frac{1}{\lambda_\ell + \lambda_h + \eta}\left(gF_h'(0) + \eta + \lambda_h\right) \tag{C.29}$$
$$F_h(\bar{z}) = \frac{1}{\lambda_\ell + \lambda_h + \eta}\left(-gF_h'(0) + \lambda_\ell\right) \tag{C.30}$$

From (C.30) , for $F_h(\bar{z}) < 1$, it must be that $F_h'(0) < \lambda_\ell/g$, which provides a bound for possible guesses.[36]

Upper bound on g: For the unbounded case where $\eta = \theta = 0$, and $\bar{z} \to \infty$, we can check the asymptotic value comes from (C.1)

$$\lim_{z\to\infty} w_i(z) = c_i \tag{C.32}$$

To find an upper bound on g, note that as $w_i(z)$ is increasing, the maximum growth rate is as $\bar{z} \to \infty$. In the limit, $\lim_{z\to\infty} w_i'(z) = 0$ as $w_i(z)$ have been constructed to be stationary. Furthermore, note that the maximum g from (C.12) is,

$$g = \lim_{\bar{z}\to\infty} \tfrac{\chi}{2} w_h(\bar{z}) = \tfrac{\chi}{2} c_h \tag{C.33}$$

Therefore, looking at the asymptotic limit of (C.7) and (C.10),

$$(r + \lambda_\ell + \eta - (1-\psi)\tfrac{\chi}{2}c_h F'(0))c_\ell = 1 + \lambda_\ell c_h \tag{C.34}$$
$$(r + \lambda_h + \eta)c_h = 1 + \left(\lambda_h + (1-\psi)\tfrac{\chi}{2}c_h F'(0)\right)c_\ell + \tfrac{\chi}{4}c_h^2 \tag{C.35}$$

Given a $F'(0)$, (C.34) and (C.35) provide a quadratic system of equations c_l and c_h—and ultimately g through (C.33). While analytically tractable given an $F'(0)$, this quadratic has a complicated solution—except if $\psi = 0$. For that case, define

$$\bar{\lambda} \equiv \frac{r + \eta + \lambda_\ell + \lambda_h}{r + \eta + \lambda_\ell} \tag{C.36}$$

Then, an upper bound on the growth rate with $\psi = 1$ and $\eta > 0$ is

$$g < \bar{\lambda}(r+\eta)\left[1 - \sqrt{1 - \frac{\chi}{\bar{\lambda}(r+\eta)^2}}\right] \tag{C.37}$$

[35]The KFE for h in (96) is co-linear with (C.26) at $\bar{z}$, and hence wouldn't provide an additional equation.

[36]Using (C.29) gives the same equation due to collinearity. An alternative approach is to rearrange (C.30) and get the $F_h'(0)$ given the particular $F_h(\bar{z})$ guess,

$$F_h'(0) = \frac{\lambda_\ell - (\eta + \lambda_h + \lambda_\ell)F_h(\bar{z})}{g} \tag{C.31}$$

where if $\eta = 0$, the unique solution is,

$$g = \bar{\lambda} r \left[1 - \sqrt{1 - \frac{\chi}{\bar{\lambda} r^2}}\right] \tag{C.38}$$

where a necessary condition for an interior equilibrium is

$$r > \sqrt{\frac{\chi}{\bar{\lambda}}} \tag{C.39}$$

□

Summarizing the full set of equations to solve for $F_i(z)$ and $w_i(z)$ from (20) to (23), (97), (C.2), (C.7), (C.10) to (C.12), (C.14), (C.15), (C.22) to (C.26) and (C.36).

Appendix D Exogenous Geometric Brownian Innovation

D.1 Model Summary

We begin by describing features of the baseline model of catchup and stochastic diffusion in the absence of a finite technology frontier, similar to Perla, Tonetti, and Waugh (2015) and Perla and Tonetti (2014). To investigate the role of stochastic innovation, this model nests a stochastic, exogenous innovation process modelled as Geometric Brownian Motion (GBM).[37,38]

Comparing to Section 2, firms are now only heterogeneous over their productivity, with cdf $\Phi(t, Z)$. On a BGP, this is equivalent to having the firm pay an upgrade cost as a fraction of the Z they draw.

Diffusion and Evolution of the Distribution If a firm adopts a new technology, then it immediately changes its productivity to a draw from the distribution $\Phi(t, Z)$, potentially distorted. The degree of imperfect mobility is indexed by $\kappa > 0$ where the agent draws its Z from the cdf $\Phi(t, Z)^\kappa$. Note that for higher κ, the probability of a better draw increases. As $\Phi(t, B(t))^\kappa = 1$ and $\Phi(t, M(t))^\kappa = 0$, for all $\kappa > 0$, this is a valid probability distribution. As before, in equilibrium, all firms choose an identical threshold, $M(t)$, above which they will continue operating with their existing technology.

A flow $S(t) \geq 0$ of firms cross into the adoption region at time t and choose to adopt a new technology. For the case in which innovation is driven by GBM, with a drift of γ and variance σ, the Kolmogorov Forward Equation (KFE) below (in cdfs) is[39]

$$\boldsymbol{\partial}_t \Phi(t, Z) = \underbrace{-(\gamma - \sigma^2/2) Z\, \boldsymbol{\partial}_Z \Phi(t, Z)}_{\text{Deterministic Drift}} + \underbrace{\frac{\sigma^2}{2} Z^2\, \boldsymbol{\partial}_{ZZ} \Phi(t, Z)}_{\text{Brownian Motion}} + \underbrace{S(t)\Phi(t, Z)^\kappa - S(t)}_{\text{Firm draws - Adopters}}, \quad \text{for } M(t) \leq Z \leq B(t) \tag{D.1}$$

[37] Staley (2011) adds exogenous geometric Brownian motion to an economy with a Lucas (2009) technology diffusion model, and investigates the evolution of the productivity distribution and growth rates.

[38] For a version of the model using monopolistic competition and associated equilibrium conditions, see Appendix D.6 and Technical Appendix B.

[39] To derive from the more common KFE written in pdfs, use the adjoint of the infinitesimal generator of GBM,

$$\frac{\partial \phi(t, Z)}{\partial t} = -\frac{\partial}{\partial Z}\left((\mu + \upsilon^2/2) Z \phi(t, Z)\right) + \frac{\partial}{\partial Z^2}\left(\frac{\upsilon^2}{2} Z^2 \phi(t, Z)\right) + \dots$$

Integrate this with respect to Z to convert into cdf $\Phi(t, Z)$, take the first derivative of the 3rd term, and then rearrange to find,

$$\frac{\partial \Phi(t, Z)}{\partial t} = (\upsilon^2/2 - \mu) Z \frac{\partial \Phi(t, Z)}{\partial Z} + \frac{\upsilon^2}{2} Z^2 \frac{\partial^2 \Phi(t, Z)}{\partial Z^2} + \dots$$

If $\sigma > 0$, then $B(t) = \infty$ immediately. Otherwise, if $\sigma = 0$ and $B(t) < \infty$, then the frontier grows at rate $B'(t)/B(t) = \gamma$.

D.2 Firm's Problem

The firm maximizes the present discounted value of profits, discounting at rate $r > 0$, where Z evolves following a GBM. The firm chooses the productivity threshold $M(t)$, below which they choose to adopt a new technology. A firm's productivity may hit $M(t)$ due to a sequence of bad relative shocks or because the $M(t)$ barrier is overtaking their Z. [40]

Assuming continuity of $\Phi(0, Z)$, then the necessary conditions for an equilibrium, $\Phi(t, Z)$ and $M(t)$, are,

$$rV(t,Z) = Z + (\gamma + \sigma^2/2)Z\,\boldsymbol{\partial}_Z V(t,Z) + \frac{\sigma^2}{2}Z^2\,\boldsymbol{\partial}_{ZZ}V(t,Z) + \boldsymbol{\partial}_t V(t,Z) \tag{D.2}$$

$$V(t,M(t)) = \int_{M(t)}^{B(t)} V(t,\hat{Z})\mathrm{d}\Phi(t,\hat{Z})^\kappa - \zeta M(t) \tag{D.3}$$

$$\boldsymbol{\partial}_Z V(t,M(t)) = 0, \quad \text{if } S(t) > 0 \tag{D.4}$$

$$\boldsymbol{\partial}_t \Phi(t,Z) = -(\gamma - \sigma^2/2)Z\,\boldsymbol{\partial}_Z\Phi(t,Z) + \frac{\sigma^2}{2}Z^2\,\boldsymbol{\partial}_{ZZ}\Phi(t,Z) + S(t)\Phi(t,Z) - S(t) \tag{D.5}$$

$$\Phi(t,M(t)) = 0 \tag{D.6}$$

$$\Phi(t,B(t)) = 1 \tag{D.7}$$

$$B'(t)/B(t) = \gamma, \text{ if } \sigma > 0 \tag{D.8}$$

where equation (D.2) is the Bellman Equation in the continuation region, and equations (D.3) and (D.4) are the value matching and smooth pasting conditions. While the value matching condition always holds, the smooth pasting condition is only necessary if there is negative drift relative to the boundary $M(t)$. Equations (D.5) to (D.7) are the Kolmogorov forward equation with the appropriate boundary conditions. Equation (D.8) is the deterministic growth of the boundary, which is simply the growth rate of frontier agents as $M(t) < B(t)$ in equilibrium.

D.3 Normalization and Stationarity

Following the normalization of Appendix A.1, leads to the following normalized set of equations. Given an initial condition $F(0, z)$, the dynamics of $v(t, z)$, $F(t, z)$, $g(t) \geq 0$, and $S(t) \geq 0$, must satisfy

$$(r - g(t))v(t,z) = e^z + (\gamma - g(t))\boldsymbol{\partial}_z v(t,z) + \tfrac{\sigma^2}{2}\boldsymbol{\partial}_{zz}v(t,z) + \boldsymbol{\partial}_t v(t,z) \tag{D.9}$$

$$v(t,0) = \int_0^\infty v(t,z)\mathrm{d}F(t,z)^\kappa - \zeta \tag{D.10}$$

$$\boldsymbol{\partial}_z v(t,0) = 0 \tag{D.11}$$

$$0 = (g(t) - \gamma)\boldsymbol{\partial}_z F(t,z) + \tfrac{\sigma^2}{2}\boldsymbol{\partial}_{zz}F(t,z) + S(t)F(t,z)^\kappa - S(t) \tag{D.12}$$

$$F(t,0) = 0 \tag{D.13}$$

$$F(t,B(t)) = 1 \tag{D.14}$$

$$S(t) = (g(t) - \gamma)\boldsymbol{\partial}_z F(t,0) + \tfrac{\sigma^2}{2}\boldsymbol{\partial}_{zz}F(t,0) \tag{D.15}$$

[40] The sequential formulation and connection to a recursive optimal stopping of a deterministic process is given on page 110-112 of Stokey (2009).

In stationary form, these become $v(z)$, $F(z)$, $g \geq 0$, $S > 0$, and $0 < \bar{z} \leq \infty$ such that,

$$(r-g)v(z) = e^z + (\gamma - g)v'(z) + \frac{\sigma^2}{2}v''(z) \tag{D.16}$$
$$v(0) = \int_0^\infty v(z)\mathrm{d}F(z)^\kappa - \zeta \tag{D.17}$$
$$v'(0) = 0 \tag{D.18}$$
$$0 = (g-\gamma)F'(z) + \frac{\sigma^2}{2}F''(z) + SF(z) - S \tag{D.19}$$
$$F(0) = 0 \tag{D.20}$$
$$F(\infty) = 1 \tag{D.21}$$
$$S = (g-\gamma)F'(0) + \frac{\sigma^2}{2}F''(0) \tag{D.22}$$

The value matching condition in (D.17) can also be written, using (C.16), as

$$\zeta = \int_0^\infty v'(z)\left(1 - F(z)^\kappa\right)\mathrm{d}z \tag{D.23}$$

To interpret (D.23), in equilibrium, as the firm is already about to gain $v(0)$ costlessly, it is indifferent between production and adoption only if the sum of all marginal values over the counter-cdf of draws is identical to the cost of adoption. (D.22) can be understood as the flux crossing the endogenous barrier, where in normalized terms the barrier is moving at rate $g-\gamma$ and collecting the infinitesimal mass at the boundary, i.e. the pdf F'(0). Additionally, there is a Brownian diffusion term where a σ dependent flow of agents are moving back purely randomly.

D.4 Deterministic Balanced Growth Path

We begin by analyzing the deterministic balanced growth path, in which $\sigma = 0$, and innovation is common and constant for all firms. First we assume that firms are adopting technologies from the unconditional distribution by setting $\kappa = 1$. Proposition 6 characterizes the balanced growth path equilibrium.

Proposition 6 (Deterministic Equilibrium with Pareto Initial Condition and $\kappa = 1$)**.** *If* $\Phi(0, Z) = 1 - \left(\frac{M_0}{Z}\right)^\alpha$, $\alpha > 1$, *and* $r > \gamma + (\zeta\alpha(\alpha-1))^{-1} > 0$, *then*

$$g = \frac{1 - \zeta(\alpha-1)(r-\alpha\gamma)}{\zeta(\alpha-1)^2}, \tag{D.24}$$
$$v(z) = \frac{1}{r-\gamma}e^z + \frac{1}{\nu(r-\gamma)}e^{-\nu z}, \tag{D.25}$$

where,

$$\nu \equiv \frac{r-g}{g-\gamma} > 0, \tag{D.26}$$

and the stationary distribution in logs is

$$F(z) = 1 - e^{-\alpha z}. \tag{D.27}$$

Proof. See Technical Appendix D.1. □

The first term of equation (D.25) is the value of production in perpetuity. This would be the value of the firm if it did not have the option of adopting a better technology. The second term of equation (D.25) is the *option value of technology diffusion.* It is decreasing in z since the optimal time to adopt is increasing with better relative technologies. The exponent, ν in (D.26) determines the rate at which the option value is discounted. More discounting of the future, or slower growth rates, lead to a more rapid drop-off of this option value.

For $\kappa > 0$, the draws are distorted and the stationary distribution is a non-Pareto power-law.

Definition 1 (Power Law Distribution)**.** *A distribution $\Phi(Z)$ is defined as a power-law, or equivalently is fat-tailed, if there exists an $\alpha > 0$ such that for large Z, the counter-cdf is asymptotically Pareto $1 - \Phi(Z) \approx Z^{-\alpha}$. Under the change of variables $z \equiv \log(Z)$ with $F(z) \equiv \Phi(e^z)$, the distribution $\Phi(Z)$ is a power-law if the counter-cdf is asymptotically exponential $1 - F(z) \approx e^{-\alpha z}$.*

Pareto distributions trivially fulfill the requirements of a Power Law. See Technical Appendix D.4.1 for more formal definition based on the theory of regularly varying functions, and Technical Appendix D.2 for more on the tail index α. We say a distribution is thin-tailed if there does not exist any $\alpha > 0$ such that the definition can hold.

D.5 Stochastic Exogenous Innovation

If innovation is stochastic and driven by GBM, even with a finite $F(0, z)$ initial condition, the support of a stationary $F(z)$ must be $[0, \infty)$. With a continuum of agents, Brownian motion instantaneously increases the support of the distribution.

When geometric random shocks are added, the stationary solutions will endogenously become power-law distributions, as discussed with generality in Gabaix (2009). Figure 3 provides some intuition on how these forces can create a stationary distribution with technology diffusion. Stochastic innovation spreads out the distribution and in the absence of endogenous adoption this would prevent the existence of a stationary distribution.[41] However, as the distribution spreads, the incentives to adopt a new technology increase, and this in turn acts to compress the distribution. In equilibrium, technology diffusion occurs with certainty because otherwise the returns to adopt a new technology become infinite in relative terms.

To see this intuitively, consider the alternative where there are geometric stochastic shocks for operating firms, but no firm chooses to adopt new technologies. A distribution generated by a random walk has a growing variance, and its support is unbounded unless there is adoption or death, even if the drift is zero. But when firms are choosing whether to adopt a new technology or not, the increasing variance of the distribution implies the returns to adoption go to infinity, overcoming any finite adoption cost. The firms at the lower end of the productivity distribution would choose to adopt, and the spread of the distribution would be contained.

Proposition 7 (Equilibrium with Geometric Brownian Motion Innovations)**.** *A continuum of equilibria parameterized by α exist satisfying*

$$\alpha > \frac{1}{2}\left(1 + \sqrt{\frac{4 + \zeta(r - \gamma - \sigma^2/2)}{\zeta(r - \gamma - \sigma^2/2)}}\right) \tag{D.28}$$

and

$$0 > (\alpha - 1)^2\alpha\zeta^2\sigma^4 - 2(\alpha - 1)\zeta\sigma^2\left((\alpha - 3)\alpha + \left(\alpha^2 - 1\right)\zeta(r - \gamma)\right) + 4((\alpha - 1)\zeta(r - \gamma) - 1)^2. \tag{D.29}$$

[41] Without endogenous adoption there is no "absorbing" or "reflecting" barrier and geometric random shocks lead to a diverging variance in the KFE.

For a given α, the growth rate is

$$g = \underbrace{\gamma}_{Innovation} + \underbrace{\frac{1-(\alpha-1)\zeta\,(r-\gamma)}{(\alpha-1)^2\zeta}}_{Catch\text{-}up\ Diffusion} + \underbrace{\frac{\sigma^2}{2}\frac{\alpha\left(\alpha(\alpha-1)\left(r-\gamma-\frac{\sigma^2}{2}\right)\zeta-2\right)+1}{(\alpha-1)\left((\alpha-1)\left(r-\gamma-\frac{\sigma^2}{2}\right)\zeta-1\right)}}_{Stochastic\ Diffusion}. \tag{D.30}$$

Furthermore the stationary distribution and value function are

$$F(z) = 1 - e^{-\alpha z} \tag{D.31}$$

$$v(z) = \frac{1}{r-\gamma-\sigma^2/2}\left(e^z + \frac{1}{\nu}e^{-\nu z}\right), \tag{D.32}$$

where,

$$\nu = \frac{\gamma-g}{\sigma^2} + \sqrt{\left(\frac{g-\gamma}{\sigma^2}\right)^2 + \frac{r-g}{\sigma^2/2}}. \tag{D.33}$$

Proof. See Technical Appendix A.2. □

The continuum of equilibria in this case is similar to that discussed in Luttmer (2007, 2012, 2014). While there exist multiple stationary equilibria, the uniqueness of a stationary equilibrium given a particular initial condition in a similar model is discussed in Luttmer (2014). This corresponds to hysteresis (i.e., dependence on initial condition $\Phi(0, Z)$): a unique path exists given a particular set of parameters and initial conditions. Furthermore, given α, the flow of adopters, S, satisfies

$$\alpha = \frac{g-\gamma}{\sigma^2} - \sqrt{\left(\frac{g-\gamma}{\sigma^2}\right)^2 - \frac{S}{\sigma^2/2}}, \tag{D.34}$$

that is,

$$S = \frac{\sigma^2}{2}\left(\alpha - \frac{g-\gamma}{\sigma^2}\right)^2 + \left(\frac{g-\gamma}{\sigma^2}\right)^2. \tag{D.35}$$

The tail index in (D.34) is of a very similar form to that in Luttmer (2014) Proposition 1, where our endogenous flow of adopters, S, is related to Luttmer's exogenous arrival rate of learning opportunities.

The growth rate and value function as $\sigma \to 0$ is identical to that in Proposition 6. Hence, in the decomposition of growth rates of (D.30), the first term is the "catch-up diffusion" caused by the same incentives as in the deterministic case, while the second is the "stochastic diffusion" caused by negative (unlucky) shocks to firms close to the adoption threshold. While some firms will receive positive shocks and be lucky when near the threshold, half of the drift-adjusted Brownian motion ends up crossing the threshold.

Another place where the Brownian motion can be seen to influence the equilibrium is in (D.33). This is a stochastic version of the ν in (D.26). Higher variance decreases the expected hitting time at which productivity reaches the normalized zero threshold, and hence increases the value of technology diffusion. Note that, due to risk neutrality, the variance is constructed to have no direct effect on the expected continuation profits, just on the expected length of time the firm will operate its existing technology.

To decompose the contributions to growth, define the growth rate with no stochastic shocks to productivity as $g_c(\alpha)$ as in (D.36). Since the contributions to growth rates from stochastic diffusion (i.e., unlucky experimentation pushing firms below the boundary in relative terms) have been removed, this can be interpreted as "catchup diffusion".

$$g_c(\alpha) \equiv \lim_{\sigma \to 0} g(\alpha; \sigma) \tag{D.36}$$

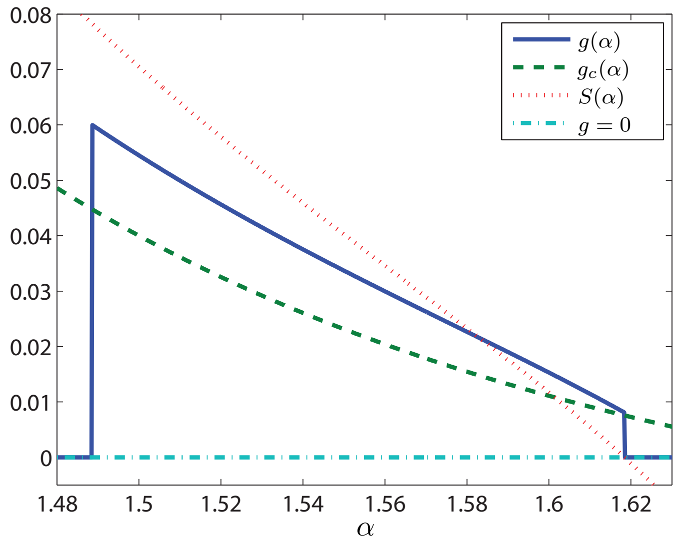

Figure 14: Growth rate and Growth from Catchup Diffusion as a Function of α

Figure 14 shows a plot of $g(\alpha)$ from equation (D.30) and $g_c(\alpha)$ from (D.36), for parameter values: $r = 0.06, \zeta = 25, \sigma = .1$, and $\gamma = 0$. The range of admissible α from the intersection of the sets in (D.28) and (D.29) is relatively tight, between about 1.49 and 1.62. As discussed before, a stationary equilibrium with strictly positive Brownian term and no equilibrium technology diffusion is not possible, so the $\alpha \simeq 1.62$ is likely the stationary distribution for a large number of initial conditions with relatively thin tails. Equilibria where $\alpha < 1.49$ here are simply not defined as the option value explodes.[42]

The minimum growth rate of around 1% occurs at the maximum α, and is strictly positive. This occurs at the point where the contribution of "stochastic diffusion", $g - g_c$, is zero. Otherwise, the contribution stochastic diffusion is strictly positive in the admissible range

D.6 Monopolistic Competition, Adoption Costs in Labor, and Free Entry

In most of this paper, we concentrate on a simple formulation with linear profits, adoption costs scaling with $M(t)$ as the economy grows, and an exogenous number of firms. Richer models of general equilibrium with employment, labor market clearing, endogenous number of varieties, and downward sloping demand are more directly comparable to Luttmer (2007) and provide qualitatively very similar results.[43]

This section nests a standard model of monopolistic competition with free entry into the model with GBM, where all costs are paid in labor at the market wage.

[42] In determining whether these α are empirically plausible, consider the crude adjustment to tail indices based on for markups discussed in (D.47).

[43] See Perla, Tonetti, and Waugh (2015) for a related derivation of the deterministic equilibrium with richer cost functions and international trade.

Consumers and Final Goods Assume a standard setup with a representative consumer, a competitive final goods sector, and a monopolistically competitive intermediate sector. The consumer gains flow utility from consumption of final goods with intertemporal elasticity of substitution, $\Lambda > 0$. Future utility is discounted at a rate $\rho > 0$. The consumer purchases the final consumption goods purchased by supplying 1 unit of labor inelastically at real wage $W(t)$ and gaining profits from a perfectively diversified portfolio of firms.

A competitive final goods sector produces a good with elasticity of substitution $\varpi > 1$ between all available intermediate varieties, given prices and real final revenue. Assume there are $N(t)$ varieties produced.

Firm's Static Choice Monopolistically competitive firms have a productivity Z, with a normalized distribution of $\Phi(t, Z)$ (i.e., there are $N(t)\Phi(t, Z)$ total firms with productivity below Z). Given a $N(t)$, an exogenous death rate, and entry decisions, the $\Phi(t, z)$ will evolve accordingly to a nearly identical form as that of Appendix D.

These production technology uses labor hired at the market wage $W(t)$, with constant returns to scale and productivity Z. Following standard algebra, the static pricing decision is a constant markup over marginal cost, $\varpi/(\varpi - 1)$, with period profits $\Pi(t, Z) \propto Z^{\varpi-1}$.

Firm's Dynamic Choice Assume that firm's exit at some exogenous $\delta \geq 0$ rate. The firm's discount future consumption using the IES of the consumer. On a BGP with constant consumption growth, g, the interest rate is $r \equiv \delta + \rho + \Lambda g$.

Productivity of operating firms evolves according to the exogenous GBM in Appendix D. Adopting a new technology requires hiring ζ units of labor at the market wage $W(t)$. Otherwise, as shown in Technical Appendix B, the problem is nearly identical to that of Appendix D.

Free Entry and Market Clearing The only major addition to this model is a free entry condition to determine $N(t)$. If $\delta > 0$, then on a BGP there will be gross entry of δN firms. Otherwise, if $\delta = 0$, then there will be no entry or exit, but an endogenously determined $N(t)$ as is standard in monopolistically competitive models with free entry.[44]

Assume that hiring $\theta > \zeta$ units of labor at the market wage $W(t)$, firms can enter and draw a Z according to the same procedure as adopting incumbents. The free entry condition becomes, $\theta W(t) = \mathbb{E}\left[V(t, Z)\right]$, with the expected draw of Z from the equilibrium distribution. As the value of adopting a new technology is $V(t, M(t))$, the free entry condition is related to the value matching condition of the dynamic problem through $(\theta - \zeta)\, W(t) = V(t, M(t))$.

The labor market clearing condition distributes the inelastic supply of labor between variable production, adoption, and entry.
Define the following,

$$\tilde{\pi} \equiv \frac{1 + \alpha - \varpi}{\alpha(\varpi - 1)}(1/N - \zeta\alpha\left(g - \gamma - \alpha\frac{\sigma^2}{2}\right) - \delta\theta) \tag{D.37}$$

$$\nu \equiv \frac{4\tilde{\pi}\alpha - 2(\alpha - 1)\varpi\left(\tilde{\pi} + \alpha\zeta(g - \gamma)\right) - 2\tilde{\pi} + (\alpha - 1)\alpha\zeta(\varpi - 1)^2\sigma^2 + 2(\alpha - 1)\alpha\zeta(-\gamma + 2g - r)}{(\alpha - 1)\zeta\left(2(-\gamma\varpi + \gamma + g(\varpi - 2) + r) - (\varpi - 1)^2\sigma^2\right) - 2\tilde{\pi}} \tag{D.38}$$

$$a = \frac{\tilde{\pi}}{r - g - (\varpi - 1)(\gamma - g + (\varpi - 1)\sigma^2/2)} \tag{D.39}$$

[44] While this determines endogenizing the number of varieties and can handle exogenous exit rates, this does not have fixed costs and does not nest a model of exit selection. To understand the orthogonal impact of endogenous selection in a model of exit, see Luttmer (2007).

Proposition 8 (Monopolistic Competition with Free Enttry on a BGP)**.** *There exist a continuum of equilibria parameterized by* α *where,*

$$F(z) = 1 - e^{-\alpha z}. \tag{D.40}$$

The tail parameter of the underlying productivity distribution is α*, while the tail parameter of the profit and firm size distributions, is given by*

$$\hat{\alpha} \equiv (\varpi - 1)\alpha. \tag{D.41}$$

Then, given the definitions for $\tilde{\pi}$ *and* ν*, the equilibrium* $\{g, N\}$ *is a solution to the following system of two equations,*

$$0 = -g + \frac{2\tilde{\pi}(\alpha-1)(\varpi-1)\sigma^2}{(\alpha-1)\zeta\left(2(-\gamma\varpi+\gamma+g(\varpi-2)+r)-(\varpi-1)^2\sigma^2\right)-2\tilde{\pi}} + \alpha\sigma^2 + \gamma + \sqrt{(g-\gamma)^2+2\sigma^2(r-g)} \tag{D.42}$$

$$\theta - \zeta = \frac{\tilde{\pi}(\nu+\varpi-1)}{\nu(-\gamma\varpi+\gamma+g(\varpi-2)+r)-\nu(\varpi-1)^2\sigma^2/2} \tag{D.43}$$

The value of a firm has the same structure of production in perpetuity plus the option value of adoption,

$$v(z) = ae^{(\varpi-1)z} + \frac{(\varpi-1)a}{\nu}e^{-\nu z} \tag{D.44}$$

Proof. The full equilibrium and specification is in Technical Appendix B. □

To provide an illustrative example, consider an example with no drift or stochastic innovation, and no exogenous exit on the BGP (i.e. $\sigma = \Lambda = \gamma = \delta = 0$). Furthermore, choose $\varpi = 2$ to simplify the algebra. With this, given an α, the growth rate and number of varieties is,

$$g = \frac{r(\theta/\zeta - \alpha)}{(\alpha-1)^2} \tag{D.45}$$

$$N = \frac{(\alpha-1)^2}{r\alpha\zeta(1-2\alpha+\alpha\theta/\zeta)} \tag{D.46}$$

The key result is that growth rates are determined by the ratio θ/ζ. For this reason, a model with an exogenous number of firms, where the cost of adoption is interpreted as being relative to the cost of entry, delivers the same qualitative results as this model.

For a fixed cost ζ, an increasing θ acts as deterrent to entry, raising profits of incumbents and increasing growth rates. On the other hand, if $\delta > 0$, there would be a force acting in the opposite direction as more costly entry takes labor away from marginal production. The elasticity, ϖ, would also effect growth as it changes the relative value of entry.

From (D.41), higher markups lead to changes in the tail parameter of the size distribution, $\hat{\alpha}$, compared to the underlying productivity distribution. Therefore, when comparing growth rates to tail parameters of the firm size distribution in the data, it is important to adjust for the elasticity and markup. Define the markup as $\bar{\varpi} \equiv (\varpi-1)/\varpi > 0$. Given an estimated $\bar{\alpha}$ from the firm size, profits, or revenue empirical distribution, the underlying tail index of the productivity distribution is

$$\alpha = \frac{1-\bar{\varpi}}{\bar{\varpi}}\hat{\alpha}. \tag{D.47}$$

This adjustment might explain some of the differences between the calibrated α in our model and those of the firm size distribution in the data. For example, with 33% markups, an empirically estimated $\hat{\alpha} = 1$ in the size or profits distribution corresponds to an underlying $\alpha = 2$ in the productivity distribution. We will use equation (D.47) to give a rough conversion from the productivity distribution to those of the empirical revenue/size/profit distribution in the rest of the paper.[45]

[45]When comparing to Luttmer (2007) and some other papers using monopolistic competition, keep in mind that the stochastic process in those papers was placed on profits or revenue directly rather than the underlying productivity distribution used here. Therefore, the tail parameters from those papers have something like (D.47) already built in, and require no adjustment.