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

Jess Benhabib  Jesse Perla  Christopher Tonetti
NYU  UBC  Stanford GSB

May 8, 2015
Draft Version: 33

VERY PRELIMINARY, ROUGH, AND INCOMPLETE
[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

The productivity distribution plays a critical role in many studies in international trade (e.g., Eaton and Kortum (2002) and Melitz (2003)), macroeconomics (e.g., Hsieh and Klenow (2009)), industrial organization (e.g., Hopenhayn (1992), Foster, Haltiwanger, and Syverson (2008)), and other areas of economics. In much of this literature, the productivity of firms evolves exogenously according to some shock process, and thus key determinants of this essential object are not studied. There is a theoretical literature that does focus on productivity growth, as pioneered by Romer (1986,
1990), Segerstrom, Anant, and Dinopoulos (1990), Rivera-Batiz and Romer (1991), Grossman and Helpman (1991, 1993), and Aghion and Howitt (1992). The key forces that generate productivity growth in these papers are innovation and imitation. Following Kortum (1997) recent papers such as Perla and Tonetti (2014) and Lucas and Moll (2014) used search theory to develop a new microfoundations for technology diffusion. In these papers all firms are alike except for their initial productivity and have no ex ante comparative advantage in innovation or imitation. However, these papers abstract away from innovation, so growth cannot continue forever if the existing technology distribution has a finite frontier, or even if it has infinite support but thin tails.\(^1\) Thus long-run growth in these models relies on the counter-factual assumption that at all times there are firms producing with arbitrarily large productivities, described by a distribution with infinite support.

In this paper, we build on this new microfoundation of technology diffusion by introducing endogenous innovation to explain economic growth through within-firm productivity improvements. The shape of the distribution of active technologies defines the opportunities for adoption. Innovation and adoption interact to determine the shape of the distribution of productivities, which in turn determines the incentives to adopt and to innovate. One of the aims of this paper is to model the evolution of productivity distributions with a frontier that is finite for all times, \(t < \infty\). For growth to continue forever, the frontier must grow through innovation. A second important aim of this paper is to model the interaction of adoption and innovation decisions. For simplicity, we start by modeling a deterministic innovation process, and continue by introducing exogenous stochastic innovation through geometric Brownian motion or discrete-state Markov chains, and finally model stochastic innovations that are subject to firm choice.

A common interesting feature of this class of models with initial distributions that have fat tails is the existence of a continuum of stationary distributions, i.e., hysteresis. We explore the conditions under which multiple stationary distributions occur in the various models that we consider. In a number of the models considered, the shape of the stationary distribution depends on the initial distribution (e.g., Perla and Tonetti (2014)), the properties of the exogenous shock process, or both (e.g., Luttmer (2007)). However, when there is an innovation decision, the shape of the stationary distribution is endogenous and depends on the parameters of the model and the optimal adoption and innovation choices of agents.

In Appendix D we develop a model of stochastic technology adoption and deterministic innovation. Technology adoption takes the form of draws from the distribution of existing technology in use, while innovation is simply modeled as exogenous multiplicative growth for all agents. We characterize the full dynamic path of the economy starting from arbitrary initial productivity distributions. An important result is that with a finite technology frontier or a thin tailed initial productivity distribution, eventually all adoption stops and long-run growth is entirely driven by the exogenous innovation process. Furthermore, the economy exhibits hysteresis: there exists a continuum of stationary distributions parameterized by the tail index of the initial distribution.

If innovation is stochastic, incentives to adopt are renewed as successful firms pull away and unsuccessful firms fall behind. We find that there are three distinct sources of growth in these classes of models, which this paper will decompose growth into contributions from: (1) firm level research decisions, i.e., “innovation”; (2) incentives for the relatively unproductive to catchup to the aggregate distribution, i.e., “catchup diffusion”; and (3) firms receiving a sequence of bad shocks relative to the growing distribution, i.e., “stochastic diffusion”. We find catch-up diffusion can only occur during transition dynamics, or in the long-run with a thick-tailed productivity distribution. Stochastic diffusion can only occur in models with risky innovation.

The forces of diffusion and innovation may interact when firms endogenously choose both, as innovation investment changes with the internalized option value of future technology diffusion.

---

\(^1\)For a formal demonstration that growth cannot continue forever if the existing technology distribution has infinite support but thin tails, see the section on Thin and Fat Tailed Distributions in Technical Appendix D.4.
opportunities. A novel element of this model is that since firms internalize the value of diffusion, there are interesting trade-offs between innovation and diffusion, which can affect the optimal growth rates. This is in contrast to papers like Luttmer (2007), where diffusion effects incumbents by increasing their fixed costs relative to profits and forcing more exit—effectively, “stochastic diffusion”. Those models provide a different, and more directly Schumpeterian, mechanism.\(^2\)

In Sections 2.1 and 2.2 we describe the basic structure of our model, including the stochastic processes for innovation and technology diffusion, the adoption decisions by firms, and the resulting law of motion for the productivity distribution.

In Section 2 we assume that the firms that are in the innovative state all grow at the same exogenous rate \(\gamma\). In Section 2.3 we study and characterize the stationary distribution of this model when the initial productivity distribution has finite support. If the initial distribution has finite support, it will maintain a finite support at all finite times. However, even though the support remains finite, the normalized stationary distribution will be unbounded. That is, the ratio of the frontier productivity to the mean productivity, \(\bar{z}\), will grow towards infinity as time progresses. We obtain a unique stationary distribution with long-run adoption, and since the frontier remains finite \(\forall \ t < \infty\), the long-run growth rate is equal to the exogenous innovation rate. Looking at data, however, we do not observe a perpetual spreading of the productivity distribution (see, e.g., Luttmer (2010), a significant shortcoming of this specification of the model.

To achieve a stationary distribution that is both finite and bounded, in that the ratio of frontier productivity to the lowest productivity converges to a finite constant \(\bar{z}\), in Section 2.4 we modify the innovation process to allow some agents to leap-frog to the frontier with a positive probability. This is a continuous time analog of the traditional quality-ladder model, in which successful innovators jump to the technology frontier. In this case, we again have a unique stationary distribution, where the long-run growth rate is equal to the exogenous innovation rate, while firms below the frontier continue to adopt. The leapfrogging/quality ladder process creates a locomotive effect where laggards do not perpetually remain behind, so \(\bar{z}\) is finite in the stationary distribution.

For completeness in Section 2.5 we also study the case, without leap-frogging, where the initial productivity distribution has infinite support. Unlike the cases where the initial distribution has finite support and remains finite for all \(t\), there is no requirement that the innovation rate \(\gamma\) equals the growth rate \(g\), as growth can also be driven by adoption from the unbounded tail. Proposition 3 characterizes the stationary equilibria. Unlike the previous cases, now there exists a continuum of stationary equilibria, as in Perla and Tonetti (2014).

Thus far, we have considered innovation processes that were exogenous. In Section 3.1, we generalize the model to allow firms to choose their innovation growth rate at some cost. Firms optimally choose to either invest in adoption or innovation, with the result that innovation growth rates are increasing with a firm’s productivity. Starting with a finite initial distribution, if firms cannot leap-frog to the frontier, the unique stationary distribution is unbounded. In the long run, the growth rate of the economy is equal to the innovation growth rate chosen by the frontier firm.

However, if we also allow some firms to leap-frog to the frontier, there now exists a continuum of stationary distributions that are bounded in relative terms (i.e., the ratio of most to least productive does not diverge). (See Section 3.2.) The growth rate of the economy is endogenous and equal to the innovation growth rate chosen by the firm at the frontier. However, in contrast to all previous cases, even with finite initial distributions, there exists a continuum of stationary distributions. This is because there is an important interaction between the incentives to adopt and to innovate that generates a self-sustaining feedback. We can index the stationary distributions by the relative

\(^2\)A side-effect of introducing stochastic innovation through Geometric Brownian Motion (GBM) is that the support of the productivity distribution becomes infinite instantly. A desirable model property is that at any point in time the technologies in use for production and available for adoption are characterized by a distribution with a finite frontier. To achieve this, in Section 2, we depart from GBM and instead model stochastic innovation as a discrete Markov process. We begin our analysis studying an exogenous uncontrolled innovation process.
frontier, \( \bar{z} \). Even with a finite frontier, a distribution that has more weight in higher productivities produces stronger incentives to adopt. The optimal innovation policy is increasing in productivity, as opposed to the exogenous innovation policy that was flat. This innovation policy generates more mass in the right tail, and thus generates more incentives to adopt, as adopters internalize the option value of innovation.

An interesting feature of the class of models with initial distributions that have fat tails is the existence of a continuum of stationary distributions, i.e., hysteresis. In the model considered in Section 2.5, the shape of the stationary distribution depends on the initial distribution (e.g., Perla and Tonetti (2014)), the properties of the exogenous shock process, or both (e.g., Luttmer (2007)). However, when there is an innovation decision, the shape of the stationary distribution is endogenous and depends on the parameters of the model and the optimal adoption and innovation choices of agents. We summarize our results on hysteresis in Section 3.3.

While the baseline model is written with an exogenous number of firms and linear profits in productivity, see Appendix D.6 in Appendix D.6 for a qualitatively similar version of the model with monopolistic competition, free-entry, all costs denoted in labor, and Geometric Brownian motion for firm dynamics.

### 1.1 Recent Literature


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 result 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. 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.

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

3Mechanically, 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.

4Luttmer (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.

5Lucas and Moll (2014) provides an extension of their baseline model with the addition of exogenous innovators in the distribution in order to discuss finite support.
of the \(\Phi_i(t, Z)\) distribution.\(^6\) The cost of adoption scales with the economy, and for simplicity is proportional to the endogenous scale of the economy, \(M(t)\).\(^8\)

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.\(^9\) Assume that an adopting firm draws a \((i, Z)\) from distributions \(\Phi(t, Z)\) and \(\Phi_h(t, Z)\), both of which will be determined by the equilibrium \(\Phi(t, Z)\) and \(\Phi_h(t, Z)\). Assume that this gives a proper cdf, so \(\Phi(t, 0) = \Phi_h(t, 0) = 0\) and \(\Phi(t, B(t)) + \Phi_h(t, B(t)) = 1\).

**Stochastic Process for Innovation** The jump intensity from low to high is \(\lambda_e > 0\) 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 \(\ell\) 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 \(i = \ell\).\(^10,11\) 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 \(i\). For the two discrete states, the Bellman equations in the continuation region are,\(^12\)

\[
\begin{align*}
rv_{\ell}(t, Z) &= Z + \lambda_e \left( V_{\ell}(t, Z) - V_{\ell}(t, Z) \right) + \frac{\partial V_{\ell}(t, Z)}{\partial t} \\
rv_{h}(t, Z) &= Z + \gamma Z \frac{\partial V_{h}(t, Z)}{\partial Z} + \lambda_h \left( V_{\ell}(t, Z) - V_{h}(t, Z) \right) + \frac{\partial V_{h}(t, Z)}{\partial t}
\end{align*}
\]

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 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 \(\gamma\).

---

\(^6\)The technical appendix is located at [Download Technical Appendix].

\(^7\)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.

\(^8\)In Technical Appendix B, a more elaborate version of this is derived in general equilibrium where \(\zeta\) 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.

\(^9\)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.

\(^10\)In Section 3, the growth rate \(\gamma\) will become a control variable for a firm, with the choice subject to a convex cost.

\(^11\)For notational simplicity, define the differential operator \(\partial\) such that \(\partial_z \equiv \frac{\partial}{\partial z}\) and \(\partial_z z \equiv \frac{\partial^2}{\partial z^2}\). When a function is univariate, derivatives will be denoted as \(r(z) \equiv \frac{dr(z)}{dz}\).

\(^12\)Ordering the states as \(\{t, h\}\), the infinitesimal generator for this continuous time Markov chain is \(Q = \begin{bmatrix} -\lambda_t & \lambda_t \\ \lambda_h & -\lambda_h \end{bmatrix}\), with adjoint operator \(Q^\ast\). The KFE and Bellman equations can be formally derived using these operators and the drift process.
**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 \( \Phi(t, Z) \) 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 \( \ell \) 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,

\[
\begin{align*}
\partial_t \Phi_\ell(t, Z) &= -\lambda_\ell \Phi_\ell(t, Z) + \lambda_h \Phi_h(t, Z) + (S_\ell(t) + S_h(t)) \Phi_\ell(t, Z) - S_\ell(t) \\
\partial_t \Phi_h(t, Z) &= -\gamma Z \partial_Z \Phi_\ell(t, Z) - \lambda_h \Phi_h(t, Z) + \lambda_\ell \Phi_\ell(t, Z) + (S_\ell(t) + S_h(t)) \Phi_h(t, Z) - S_h(t)
\end{align*}
\]

(3)  (4)

Recognizing that the \( \lambda_i \) 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,\(^\text{13}\)

\[
\begin{align*}
S_\ell(t) &\equiv M'(t) \partial_Z \Phi_\ell(t, M(t)) \\
S_h(t) &\equiv \left( M'(t) - \gamma M(t) \right) \partial_Z \Phi_h(t, M(t))
\end{align*}
\]

(5)  (6)  (7)

**Adoption Decision**  Firms choose thresholds, below which they adopt a new technology through the technology diffusion process.\(^\text{14}\) Necessary conditions for the optimal stopping problem include value matching and smooth pasting conditions at the endogenously chosen adoption boundary, \( M(t) \),

\[
\begin{align*}
\bar{V}_\ell(t, M(t)) &= \int_{M(t)}^{B(t)} \bar{V}_\ell(t, Z) d\Phi_\ell(t, Z) + \int_{M(t)}^{B(t)} \bar{V}_h(t, Z) d\Phi_h(t, Z) - \zeta M(t) \\
\partial_Z \bar{V}_\ell(t, M(t)) &= 0, \text{ if } M'(t) > 0 \\
\partial_Z \bar{V}_h(t, M(t)) &= 0, \text{ if } M'(t) - \gamma M(t) > 0
\end{align*}
\]

(8)  (9)  (10)

\(^\text{13}\)This is consistent with the solution to the ODEs in (3) and (4) at \( Z = M(t) \), and is clear in the normalized (16) and (17).

\(^\text{14}\)While the threshold could depend on the type \( i \), see Appendix A.2 for a proof that \( \ell \) and \( h \) agents choose the same threshold, \( M(t) \), if the net value of adoption is independent of the current innovation type.
change of variables, normalized distribution, and normalized value functions as,

\[ z \equiv \log(Z/M(t)) \]  
\[ F_i(t, z) = F_i(t, \log(Z/M(t))) \equiv \Phi_i(t, Z) \]  
\[ v_i(t, z) = v_i(t, \log(Z/M(t))) \equiv \frac{V_i(t, Z)}{M(t)} \]

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 \( F(z) \equiv F_\ell(z) + F_h(z) \).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{normalized_vs_unnormalized_distributions.png}
\caption{Normalized vs. Unnormalized Distributions}
\end{figure}

With the above normalizations, the value function, productivity distribution, and growth rates can be stationary and independent of time.\(^{15}\) 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

\(^{15}\)An important example is when \( \Phi(t, Z) \) is Pareto with minimum of support \( M(t) \) and tail parameter \( \alpha \):

\[ \Phi(t, Z) = 1 - \left( \frac{M(t)}{Z} \right)^\alpha, \text{ for } M(t) \leq Z \]  
\[ F(t, z) = 1 - e^{-\alpha z}, \text{ for } 0 \leq z < \infty. \]  

This is the cdf of an exponential distribution, with parameter \( \alpha > 1 \). From a change of variables, if \( X \sim \text{Exp}(\alpha) \), then \( e^X \sim \text{Pareto}(1, \alpha) \). Hence, \( \alpha \) is the tail index of the unnormalized Pareto distribution for \( Z \).
the distribution,
\[ 0 = gF'_\ell(z) - \lambda_\ell F_\ell(z) + \lambda_h F_h(z) + (S_\ell + S_h)\hat{F}_\ell(z) - S_\ell \] (16)
\[ 0 = (g - \gamma)F'_h(z) - \lambda_h F_h(z) + \lambda_\ell F_\ell(z) + (S_\ell + S_h)\hat{F}_h(z) - S_h \] (17)
\[ 0 = F_\ell(0) = F_h(0) \] (18)
\[ 1 = F_\ell(\bar{z}) + F_h(\bar{z}) \] (19)
\[ S_\ell = gF'_\ell(0) \] (20)
\[ S_h = (g - \gamma) F'_h(0) \] (21)

The summary of necessary conditions for the firm’s problem are
\[ (r - g)v_\ell(z) = e^z - g v'_\ell(z) + \lambda_\ell (v_h(z) - v_\ell(z)) \] (22)
\[ (r - g)v_h(z) = e^z - (g - \gamma)v'_h(z) + \lambda_h (v_\ell(z) - v_h(z)) \] (23)
\[ v_\ell(0) = v_h(0) = \int_0^\bar{z} v_\ell(z)d\hat{F}_\ell(z) + \int_0^\bar{z} v_h(z)d\hat{F}_h(z) - \zeta \] (24)
\[ v'_\ell(0) = 0, \text{ if } g > 0 \] (25)
\[ v'_h(0) = 0, \text{ if } g > \gamma \] (26)

To interpret: (16) to (19) are the stationary KFE with initial conditions and boundary values. $S_\ell$ in (20) is the flow of $\ell$ agents moving backwards at a relative speed of $g$ across the barrier, while $S_h$ in (21) 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.

(22) and (23) are the Bellman Equations in the continuum region, where (24) is the value matching condition between the continuation and technology adoption regions. The smooth pasting conditions in (25) and (26) 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.

![Figure 2: Normalized, Stationary Value Functions](image)

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.
Adoption compresses

Stochastic innovation spreads

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., \( \text{sup support } \{ F(0, \cdot) \} = \infty \)). 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 is 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 \( \ell \) type regardless of its former type, and the \( Z \) is drawn from a distortion of the unconditional distribution.\(^\text{16}\) 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 \( \kappa \), 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 \( \kappa \) is exogenous here, in Section 3.2, we solve a version of the model with directed technology diffusion where \( \kappa \) is endogenous.

\[ \hat{\Phi}_{\ell}(t, Z) \equiv (\Phi_{\ell}(t, Z) + \Phi_{h}(t, Z))^\kappa \]  \hspace{1cm} \text{(27)}

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

\[ \hat{F}_{\ell}(z) = (F_{\ell}(z) + F_{h}(z))^\kappa = F(z)^\kappa \]  \hspace{1cm} \text{(28)}

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

\(^\text{16}\)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.
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 $\gamma$ 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 $v_i(z)$, $v_h(z)$, $F_i(z)$, $F_h(z)$, $S$, such that,

\[(r - g) v_i(z) = e^z - g v_i'(z) + \lambda_i (v_h(z) - v_i(z)) \quad (29)\]
\[(r - g) v_h(z) = e^z + \lambda_h (v_i(z) - v_h(z)) \quad (30)\]

\[v_i(0) = \int_0^z v_i(z) dF(z)^\kappa - \zeta \quad (31)\]
\[v_i'(0) = 0 \quad (32)\]
\[0 = gF_i'(z) + SF(z)^\kappa + \lambda_i F_h(z) - \lambda_i F_i(z) - S \quad (33)\]
\[0 = \lambda_i F_i(z) - \lambda_h F_h(z) \quad (34)\]
\[0 = F_i(0) = F_h(0) \quad (35)\]
\[1 = F_h(z) + F_i(z) \quad (36)\]
\[S = gF_i'(0) \quad (37)\]

Define the constants, $\hat{\lambda} \equiv \frac{\lambda_i}{\lambda_h}, \tilde{\lambda} \equiv \frac{\lambda_i}{r - \gamma + \lambda_h} + 1$. The following characterizes the equilibrium,

**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 $\tilde{z} \to \infty$.

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

\[F_i(z) = \frac{1}{1 + \lambda} e^{-\alpha z} \quad (38)\]
\[F_h(z) = \tilde{\lambda} F_i(z), \quad (39)\]

where $\alpha$ is the tail index of the power law distribution:

\[\alpha \equiv (1 + \hat{\lambda}) F_i'(0), \quad (40)\]

and $F_i'(0)$ is determined by model parameters:

\[F_i'(0) = \frac{\lambda_i (cr + \lambda_h + \lambda_i) - \sqrt{z(4\gamma + r^2 \zeta(z - \gamma + r + \lambda_h)^2 + 2(-2\gamma + (\gamma - r)\zeta)(\gamma - r - \lambda_h)\lambda_i + (\gamma - r)^2 \zeta \lambda_i^2)} + \sqrt{\gamma} 2 + \zeta (\gamma - r - \lambda_h))}{\gamma 2 \zeta (\lambda_i + \lambda_h + \lambda_i)/r} (41)\]

The firm value functions are,

\[v_i(z) = \frac{\tilde{\lambda}}{\gamma + (r - \gamma)\lambda} e^z + \frac{1}{(r - \gamma)(\nu + 1)} e^{-\nu z} \quad (43)\]
\[v_h(z) = \frac{e^z + \lambda h v_i(z)}{r - \gamma + \lambda h}, \quad (44)\]

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

\[\nu \equiv \frac{(r - \gamma)\tilde{\lambda}}{\gamma}. \quad (45)\]
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 $\ell$ 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 $\ell$ 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 $\gamma = .02, r = .06, \lambda_\ell = .01, \lambda_h = .03, \text{and } \zeta = 25$. 

![Figure 4: Normalized, Stationary Value Functions and PDFs for the Unbounded Cases](image-url)
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 $t \to \infty$. This is because a diminishing, yet strictly positive number of firms keep getting lucky and grow at $\gamma$ 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. The jumps can occur either for firms adopting through diffusion, or for innovating firms that successfully 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 distribution. 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 accommodate 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. To consider the case where it is adopters–rather than just innovating firms–who jump to the frontier, see Section 3.2 where an endogenous jump probability is chosen by adopting firms.

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:

$$\Delta_{\ell} = \lim_{\epsilon \to 0} (F_{\ell}(\bar{z}) - F_{\ell}(\bar{z} - \epsilon))$$  \hspace{1cm} (46)

$$\Delta_{h} = \lim_{\epsilon \to 0} (F_{h}(\bar{z}) - F_{h}(\bar{z} - \epsilon))$$  \hspace{1cm} (47)

Set $\kappa = 1$ for simplicity, and define $\mathbb{H}(z)$ as the Heaviside operator. The following characterizes

17This 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.

18In a model of leapfrogging arrivals and a multiplicative step above the frontier, in continuous time the frontier would become infinite immediately. Alternatively, it could be recast as a step-by-step innovation model in the spirit of Aghion, Acquisti, and Howitt (2013) with the same qualitative results.

19Rather than being the autarkic process improvement of the $\gamma$ growth, this is leapfrogging 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.

20The assumption of a jump to the $\ell$ 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 $\epsilon$, 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 $\theta$ probability in Section 3.2, who jump, this doesn’t occur.
the necessary conditions for a stationary equilibrium,

\[ (r - g)v_i(z) = e^z - gv'_i(z) + \lambda_e(v_h(z) - v_e(z)) + \eta(v_e(\bar{z}) - v_e(z)) \]  

\[ (r - g)v_h(z) = e^z + \lambda_h(v_e(z) - v_h(z)) + \eta(v_e(\bar{z}) - v_h(z)) \]  

\[ v_e(0) = \int_0^\bar{z} v_e(z)(F'_e(z) + F'_h(z))dz - \zeta \]  

\[ v'_e(0) = 0 \]  

\[ 0 = gF'_e(z) + \lambda_hF_h(z) - \lambda_eF_e(z) - \eta F_e(z) + \eta \bar{z}(z - \bar{z}) + SF(z) - S \]  

\[ 0 = \lambda_eF_e(z) - \lambda_hF_h(z) - \eta F_h(z) \]  

\[ 0 = F_e(0) = F_h(0) \]  

\[ 1 = F_e(\bar{z}) + F_h(\bar{z}) \]  

\[ S = gF'_e(0) \]  

Let \( r > \gamma \) and define the constants, \( \hat{\lambda} \equiv \frac{\lambda_e}{\eta + \lambda_h} \), \( \bar{\lambda} \equiv \frac{r - \gamma + \lambda_e + \lambda_h}{r - \gamma + \lambda_h} \), and \( \nu = \frac{r - \gamma + \eta}{r} \bar{\lambda} \). Furthermore, assume that the value of \( F'_e(0) \) that solves (61) is larger than \( \eta/\gamma \).

**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,

\[ F_e(z) = \frac{F'_e(0)}{(F'_e(0) - \eta/\gamma)(1 + \hat{\lambda})} (1 - e^{-\alpha z}) \]  

\[ F_h(z) = \hat{\lambda}F_e(z) \]  

where

\[ \alpha \equiv (1 + \hat{\lambda})(F'_e(0) - \eta/\gamma) \]  

\[ \bar{z} = \log\left(\frac{\eta F'_e(0)}{\alpha}\right) \]  

The equilibrium \( F'_e(0) \) solves the following implicit equation substituting for \( \alpha \) and \( \bar{z} \),

\[ \zeta + \frac{1}{r - \gamma} = \frac{\gamma F'_e(0)\alpha \bar{\lambda} \left( -\frac{e^{-\nu \bar{z}}(1 + e^{-\bar{z}})}{(-\gamma + r)\alpha \nu} + \frac{e^\nu e^{-\alpha \bar{z}} - 1}{\alpha (r - \gamma)} + \frac{e^{-\nu(\alpha + \nu)}e^{\bar{z}} + 1}{\nu(\alpha + \nu)} + \frac{e^{\bar{z} - \alpha \bar{z} + 1}}{\alpha - 1} \right)}{\gamma (F'_e(0) - \eta)(\nu + 1)} \]  

Figure 5: Normalized, Stationary Value Functions with Bounded Support
The value functions for the firm are,

\[ \begin{align*}
\nu_\ell(z) &= \frac{\lambda}{\gamma(1+\nu)} \left( e^z + \frac{1}{\nu} e^{-\nu z} + \frac{\eta}{r-\gamma} \left( e^z + \frac{1}{\nu} e^{-\nu z} \right) \right) \\
\nu_h(z) &= \frac{e^z + (\lambda_h - \eta)\nu_\ell(z) + \eta\nu_\ell(\bar{z})}{r-\gamma + \lambda_h}.
\end{align*} \] (62)

(63)

Proof. See Appendix B.2.

In the above, note that \( F_\ell(z) \) are continuous, and \( \Delta_\ell = \Delta_h = 0 \). This is because leapfrogging firms become type \( \ell \), 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 (62), now contain the value of production in perpetuity and the option value for both the current \( z \) and the frontier \( z \).

The above \( \alpha \) 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. 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 adoption threshold. The \( \alpha \) computed from this example is 1.2, close to the empirical Zipf’s law.

![Figure 6: Exogenous \( v_\ell(z) \), and \( F'_\ell(z) \) with a Bounded Frontier](image)

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 \( \ell \) 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 $\alpha$ are shown in Figure 7 for changes in $\eta, \gamma, \zeta,$ and $\lambda_h$. 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 (60), a relationship can be found that determines the range of the productivity distribution for any particular $F'_\ell(0)$:

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

(64)

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).\textsuperscript{21}

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 $\ell$ state.\textsuperscript{22} The normalized adoption distributions are then, $\tilde{F}_\ell(z) \equiv F_\ell(z)$ and $\tilde{F}_h(z) \equiv F_h(z)$, which can be verified to yield a proper CDF: $\tilde{F}_\ell(\infty) + \tilde{F}_h(\infty) = 1$.

\textsuperscript{21}Uniqueness of related models with Geometric Brownian Motion is discussed in Luttmer (2012).

\textsuperscript{22}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.
By construction, and the parameter restrictions given in Proposition 3 require $g$ to ensure a stationary non-degenerate productivity distribution. However, $g \geq \gamma$ is necessary to ensure that $S_h \geq 0$. Summarizing the stationary equations,

\[ \begin{align*}
(r - g)v'_l(z) &= e^z - gv'_l(z) + \lambda_l (v_h(z) - v_l(z)) \\
(r - g)v'_h(z) &= e^z - (g - \gamma)v'_h(z) + \lambda_h (v_l(z) - v_h(z)) \\
v_l(0) &= v_h(0) = \frac{1}{r - g} = \int_0^\infty v_l(z) dF_l(z) + \int_0^\infty v_h(z) dF_h(z) - \zeta
\end{align*} \]  

(65) (66) (67)

Adopt both $i$ and $Z$ of draw $A$ defines

\[ F_l(z) = \begin{bmatrix} F_l(z) \\ F_h(z) \end{bmatrix}, \quad v(z) = \begin{bmatrix} v_l(z) \\ v_h(z) \end{bmatrix}, \quad A = \begin{bmatrix} \frac{1}{r - g} \\ g - \gamma \end{bmatrix}, \quad B = \begin{bmatrix} \frac{r + \lambda_l - g}{g - \gamma} \\ -\lambda_l \end{bmatrix}, \quad \varphi = \sqrt{(\lambda_h - \alpha \gamma)^2 + 2\lambda_l (\alpha \gamma + \lambda_h) + \lambda_l^2} \]  

(70) (71) (72) (73) (74) (75)

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

\[ \begin{align*}
\bar{F}(z) &= \begin{bmatrix} F_l(z) \\ F_h(z) \end{bmatrix} \\
\varphi &= \sqrt{(\lambda_h - \alpha \gamma)^2 + 2\lambda_l (\alpha \gamma + \lambda_h) + \lambda_l^2} \\
C &= \begin{bmatrix} -\alpha \gamma + 2ag + \lambda_h - \lambda_i + \varphi \\ 2g \\ \frac{\lambda_h}{g - \gamma} \\ -\alpha \gamma + 2ag + \lambda_h + \lambda_l + \varphi \end{bmatrix} \\
D &= \begin{bmatrix} \lambda_h g (\alpha \gamma + \lambda_h - \lambda_i) + 2g \lambda_l \\ \lambda_h g (\alpha \gamma + \lambda_h - \lambda_i + 2g \lambda_l) \\ g (\alpha \gamma - \lambda_h + \varphi - \lambda_l) + 2g \lambda_l \\ 2g (g - \gamma) \end{bmatrix}
\end{align*} \]  

(76) (77) (78) (79) (80)

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

\[ \frac{1}{r - g} + \zeta = \int_0^\infty \left[ (I + B)^{-1} (e^{Iz} + e^{-Bz} B^{-1}) A \right]^T e^{-Cz} D \, dz \]  

(81)

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

\[ \begin{align*}
\bar{F}(z) &= (I - e^{-Cz}) C^{-1} D \\
\bar{F}'(z) &= e^{-Cz} D \\
v(z) &= (I + B)^{-1} (e^{Iz} + e^{-Bz} B^{-1}) A
\end{align*} \]  

(82) (83) (84)

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

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 \( r > g \) 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 asymptotically goes to 0 for large \( z \).

See Figures 8 and 9 for an example of infinite support with \( \gamma = .01, r = .05, \lambda_{t} = .0004, \lambda_{h} = .03, \) and \( \zeta = 6.0 \). In this equilibrium, \( g = .029, \alpha = 2.5 \) and \( F_{h}(\infty) = .988 \). The relationship between \( g \) and \( \alpha \) 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.

![Figure 8: Exogenous \( v_{i}(z) \), and \( F'_{i}(z) \) with a Infinite Frontier](image)

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.
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). At first we assume that the arrival rate of jumps to the frontier, $\eta$, is zero in order to analyze the unbounded case with $\bar{z} \to \infty$. Then in Section 3.2 we move to the model with endogenously chosen innovation rates and jump probabilities and show that the frontier $\bar{z}$ is finite.

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.\(^{23}\)

3.1 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 $\gamma$. Adapting the equations in Section 2.3 after normalizing the

\(^{23}\)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.
innovation cost, and using the result that $\tilde{z} \to \infty$ in the absence of jumps to the frontier,

$$(r - g)v_\ell(z) = e^z - gv_\ell'(z) + \lambda_\ell(v_h(z) - v_\ell(z))$$

(85)

$$(r - g)v_h(z) = \max_{\gamma \geq 0} \left\{ e^z - \frac{1}{\chi} e^{-z} \gamma^2 - (g - \gamma) v_h'(z) + \lambda_h(v_\ell(z) - v_h(z)) \right\}$$

(86)

$$v_\ell(0) = v_h(0) = \int_0^\infty v_\ell(z)dF(z)^\kappa - \zeta$$

(87)

$$v_\ell'(0) = v_h'(0) = 0$$

(88)

$$0 = gF'_\ell(z) + (S_\ell + S_h)F(z)^\kappa + \lambda_hF_h(z) - \lambda_\ell F_\ell(z) - S_\ell$$

(89)

$$0 = (g - \gamma(z))F_h'(z) + \lambda_\ell F_\ell(z) - \lambda_h F_h(z) - S_h$$

(90)

$$0 = F_\ell(0) = F_h(0)$$

(91)

$$1 = F_h(\infty) + F_\ell(\infty)$$

(92)

$$S_\ell = gF'_\ell(0)$$

(93)

$$S_h = gF'_h(0)$$

(94)

Define the constant,

$$\bar{\lambda} = \frac{r + \lambda_\ell + \lambda_h}{r + \lambda_\ell}$$

(95)

With this setup, instead of all firms growing at rate $\gamma$ exogenously, $h$-type firms are choosing a growth rate $\gamma$ that is a function of their current productivity level, $z$. As the choice of $\gamma$ 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). If $r > \sqrt{\frac{\lambda}{\chi}}$, then a unique equilibrium exists with a growth rate of,

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

(96)

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

$$(r - g)v_\ell(z) = e^z - gv_\ell'(z) + \frac{1}{\chi} e^{-z}v_\ell'(z)^2 + \lambda_h(v_\ell(z) - v_h(z))$$

(97)

$$(r - g)v_h(z) = e^z - gv_h'(z) + \lambda_\ell(v_h(z) - v_\ell(z))$$

(98)

$$v_\ell(0) = v_h(0) = \frac{1}{r - g}.$$ 

(99)

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

$$\gamma(z) = \frac{\bar{\lambda}}{2}e^{-z}v_h'(z).$$

(100)

With this $\gamma(z)$, $F_\ell(z)$ solves the KFEs in (89) to (92).

**Proof.** See Appendix C.1. A numerical method to compute the equilibrium is described in Technical Appendix E.3
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%.

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

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

(101)

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.

**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 (2011), consider that 1) technology adoption is a key component growth of small firms but is not measured by $\gamma$,
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 (2011) 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 (2011), as discussed).

3.2 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 3.1, 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 $\gamma$, 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 $\theta$ is $\frac{1}{\zeta} \theta^2$. 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 $\frac{1}{\eta} \kappa^2$. We can adapt Section 2.4 to include the same endogenous intensive innovation choice of $\gamma$, a new
The directed technology diffusion solves the implicit equation
\begin{equation}
(r - g)v_{\ell}(z) = e^z - gu_{\ell}(z) + \lambda_{\ell}(v_h(z) - v_{\ell}(z)) + \eta(v_{\ell}(\bar{z}) - v_{\ell}(z)) \tag{102}
\end{equation}
\begin{equation}
(r - g)v_h(z) = \max_{\gamma \geq 0} \left\{ e^z - \frac{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{103}
\end{equation}
\begin{equation}
v_s \equiv v_s(0) = v_h(0) = \max_{\theta \geq 0, \kappa > 0} \left\{ (1 - \theta) \int_0^\bar{z} v_{\ell}(z) dF(z)^\kappa + \theta v_{\ell}(\bar{z}) - \zeta - \frac{1}{\chi} \gamma^2 - \frac{1}{\nu} \kappa^2 \right\} \tag{104}
\end{equation}
\begin{equation}
v_{\ell}'(0) = v_h'(0) = 0 \tag{105}
\end{equation}
\begin{equation}
0 = gF_{\ell}'(z) + \lambda_hF_h(z) - \lambda_{\ell}F_{\ell}(z) - \eta F_{\ell}(z) \tag{106}
\end{equation}
\begin{equation}
0 = (g - \gamma(z))F_h'(z) + \lambda_{\ell}F_{\ell}(z) - \lambda_hF_h(z) - \eta F_h(z) - S_{\ell} \tag{107}
\end{equation}
\begin{equation}
0 = F_{\ell}(0) = F_h(0) \tag{108}
\end{equation}
\begin{equation}
1 = F_h(\bar{z}) + F_{\ell}(\bar{z}) \tag{109}
\end{equation}
\begin{equation}
S_{\ell} = gF_{\ell}'(0) \tag{110}
\end{equation}
\begin{equation}
S_h = gF_h'(0) \tag{111}
\end{equation}

Note that when a particular firm is choosing \( \theta \) or \( \kappa \), it does not influence the \( \theta \) or \( \kappa \) 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 \( \theta \), which will induce a \( F_{i}(z) \) that is consistent with firm beliefs about \( F_{i}(z) \) in equilibrium. Define \( \lambda \) as
\begin{equation}
\bar{\lambda} = \frac{r + \eta + \lambda_{\ell} + \lambda_h}{r + \eta + \lambda_{\ell}}. \tag{112}
\end{equation}

**Proposition 5** (Stationary Equilibrium with Continuous Endogenous Innovation and Bounded Support). A continuum of equilibria exist, parameterized by a \( \bar{z} \). To ensure \( g < r \) there is a maximum feasible \( g \)—which is decreasing in \( \eta \)—given by
\begin{equation}
g_{\text{max}} = \bar{\lambda}(r + \eta) \left[ 1 - \sqrt{1 - \frac{\chi}{\bar{\lambda}(r + \eta)^2}} \right]. \tag{113}
\end{equation}

The accompanying \( \gamma(z), v_{\ell}(z) \), and \( F_{i}(z) \) solve the system of non-linear ODEs,
\begin{equation}
(r - g)v_{\ell}(z) = e^z - gu_{\ell}(z) + \lambda_{\ell}(v_h(z) - v_{\ell}(z)) + \eta(u_{\ell}(\bar{z}) - v_{\ell}(z)) \tag{114}
\end{equation}
\begin{equation}
(r - g)v_h(z) = e^z - g\bar{v}_{\ell}'(z) + \frac{\lambda_{\ell}}{4} e^{-z} v_{\ell}'(z)^2 + \lambda_h(v_{\ell}(z) - v_h(z)) + \eta(v_{\ell}(\bar{z}) - v_h(z)) \tag{115}
\end{equation}
\begin{equation}
v_{\ell}'(0) = v_h'(0) = 0 \tag{116}
\end{equation}
\begin{equation}
v_{\ell}(0) = v_h(0) = \frac{1 + \eta v_{\ell}(\bar{z})}{r - g + \eta} \tag{117}
\end{equation}

The endogenous innovation choice is such that \( \gamma(0) = 0, \gamma(\bar{z}) \equiv g \) and
\begin{equation}
\gamma(z) = \frac{\lambda_{\ell}}{4} e^{-z} v_{\ell}'(z) \tag{118}
\end{equation}

Given the innovation intensity, the distribution solves (104), (106) and (107) where the chosen intensity of jumps to the frontier for adopting agents is,
\begin{equation}
\theta = 1 - \sqrt{1 - \zeta(v_{\ell}(\bar{z}) - v_{\ell}(0) - \zeta)}. \tag{119}
\end{equation}

The directed technology diffusion solves the implicit equation
\begin{equation}
\kappa = -\frac{\theta - 1}{2} \int_0^{\bar{z}} v_{\ell}(z) \log(F(z)) F(z)^\kappa dz \tag{120}
\end{equation}
Proof. See Appendix C.2. A numerical method to solve for the continuum of equilibria is described in Technical Appendix E.2.

In the limit, as $\eta \to 0$, the upper bound on $g$ in (113) becomes the limiting case in Proposition 4. Comparing to Stokey (2014), here the endogenous choice of $\gamma$ is complicated by the option value. In the case of Proposition 4, the asymptotic $\gamma$ 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 11: Endogenous $\gamma(z), v_l(z)$, and $F_i(z)$ with an Bounded Frontier](image1)

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

Figure 13 plots the maximum growth rate of the set of admissible $g$ as a function of $\eta$ using (113) 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 3.2 asymptotically becomes that in Section 3.1. 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.

### 3.3 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 (113)), 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
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

tail of the initial distribution, the richer will be the opportunities to adopt superior technologies, 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.\(^{24}\) 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-frogs, 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 \(\bar{\tau}\), the ratio of frontier productivity to the mean of the distribution (Proposition 5).\(^{25}\) In this case the support and the position of the stationary distribution, as well as the growth rate, are parametrically determined \(\bar{\tau}\).

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.\(^{26}\) 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.

\(^{24}\)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.

\(^{25}\)Note that unlike crefprop:stationary-equilibrium-exogenous-finite-jumps \(\bar{\tau}\) can be jointly solved with \(F_\ell(0)\), using (60) and (61). \(\bar{\tau}\) Proposition 5 \(\bar{\tau}\) can be chosen parametrically to define the stationary distribution.

\(^{26}\)See for example Propositions 1 to 3, 6 and 7 where the tail index is denoted as \(\alpha\).
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 3.3). 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.
References


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 $M(t)$:

$$
\Phi_i(t, Z) \equiv F_i(t, \log(Z/M(t)))
$$

(A.1)

Differentiating to obtain the pdf yields

$$
\partial_Z \Phi_i(t, Z) = \frac{1}{Z} \frac{\partial F_i(t, \log(Z/M(t)))}{\partial z} = \frac{1}{Z} \partial_Z F_i(t, z)
$$

(A.2)

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

$$
\partial_t \Phi_i(t, Z) = \frac{\partial F_i(t, \log(Z/M(t)))}{\partial t} - \frac{M'(t)}{M(t)} \frac{\partial F_i(t, \log(Z/M(t)))}{\partial z}
$$

(A.3)

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

$$
\partial_t \Phi_i(t, Z) = \partial_t F_i(t, z) - g(t) \partial_z F_i(t, z)
$$

(A.4)

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

$$
\frac{\partial F_i(t, \log(Z/M(t)))}{\partial t} - g(t) \frac{\partial F_i(t, \log(Z/M(t)))}{\partial z} = -\lambda_t F_i(t, \log(Z/M(t))) + \lambda_h F_h(t, \log(Z/M(t))) + (S_i(t) + S_h(t)) F_i(t, \log(Z/M(t))) - S_i(t)
$$

(A.5)

$$
\frac{\partial F_h(t, \log(Z/M(t)))}{\partial t} - g(t) \frac{\partial F_h(t, \log(Z/M(t)))}{\partial z} = -\lambda_h F_h(t, \log(Z/M(t))) + \lambda_t F_i(t, \log(Z/M(t))) - \gamma Z \frac{\partial F_h(t, \log(Z/M(t)))}{\partial z} + (S_i(t) + S_h(t)) F_i(t, \log(Z/M(t))) - S_h(t)
$$

(A.6)

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

$$
\partial_t F_i(t, z) = -\lambda_t F_i(t, z) + \lambda_h F_h(t, z) + g(t) \partial_z F_i(t, z) + S_i(t) F_i(t, z) - S_i(t)
$$

(A.7)

$$
\partial_t F_h(t, z) = \lambda_t F_i(t, z) - \lambda_h F_h(t, z) + (g(t) - \gamma) \partial_z F_h(t, z) + S_i(t) F_h(t, z) - S_h(t)
$$

(A.8)

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

$$
S_i(t) = g(t) \partial_z F_i(t, 0)
$$

(A.9)

$$
S_h(t) = (g(t) - \gamma) \partial_z F_h(t, 0)
$$

(A.10)
Normalizing the Value Function  
Define the normalized value of the firm as,
\[ v_i(t, \log(Z/M(t))) = \frac{V_i(t, Z)}{M(t)} \]  
(A.11)

Rearrange and differentiating (A.11) with respect to \( t \)
\[ \partial_t V_i(t, Z) = M'(t)v_i(t, \log(Z/M(t))) - M'(t)\frac{\partial v_i(t, \log(Z/M(t)))}{\partial z} + M(t)\frac{\partial v_i(t, \log(Z/M(t)))}{\partial t} \]  
(A.12)

Divide by \( M(t) \) and use the definition of \( g(t) \)
\[ \frac{1}{M(t)} \partial_t V_i(t, Z) = g(t)v_i(t, z) - g(t)\partial_z v_i(t, z) + \partial_t v_i(t, z) \]  
(A.13)

Differentiating (A.11) with respect to \( Z \) and rearranging
\[ \frac{1}{M(t)} \partial_Z V_i(t, Z) = \frac{1}{Z} \partial_z v_i(t, z) \]  
(A.14)

Divide (2) by \( M(t) \) and then substitute from (A.13) and (A.14),
\[ \frac{r}{M(t)} V_h(t, Z) = \frac{Z}{M(t)} + \frac{M(t)}{M(t)} \frac{Z}{Z} \partial_z v_h(t, z) + g(t)v_h(t, z) - g(t)\partial_z v_h(t, z) + \lambda_h(v(t, z) - v_h(t, z)) + \partial_t v_h(t, z) \]  
(A.15)

Use (A.11) and the definition of \( z \) and rearrange,
\[ (r - g(t))v_h(t, z) = e^z + (\gamma - g(t))\partial_z v_h(t, z) + \partial_t v_h(t, z) \]  
(A.16)

Similarly, for (1)
\[ (r - g(t))v(t, z) = e^z - g(t)\partial_z v(t, z) + \lambda_e(v(t, z) - v(t, z)) + \partial_t v(t, z) \]  
(A.17)

Optimal Stopping Conditions  
Divide the value matching condition in (8) by \( M(t) \),
\[ \frac{V_i(t, M(t))}{M(t)} = \int_{M(t)}^{B(t)} \frac{V_i(t, Z)}{M(t)} \partial_Z \hat{\Phi}_e(t, Z) dZ + \int_{M(t)}^{B(t)} \frac{V_h(t, Z)}{M(t)} \partial_Z \hat{\Phi}_h(t, Z) dZ - \frac{M(t)}{M(t)} \zeta \]  
(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 \( dz = \frac{1}{Z} dZ \). Note that the bounds of integration change to \([\log(M(t)/M(t)), \log(B(t)/M(t))] = [0, \bar{z}(t)]\)
\[ v_i(t, 0) = \int_0^{\bar{z}(t)} v(t, z) d\hat{F}_e(t, z) + \int_0^{\bar{z}(t)} v_h(t, z) d\hat{F}_h(t, z) - \zeta \]  
(A.19)

Evaluate (A.14) at \( Z = M(t) \), and substitute this into (10) to give the smooth pasting condition,
\[ \partial_z v_i(t, 0) = 0 \]  
(A.20)

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 \( \partial_z v(t, 0) = -\zeta \). 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 $\gamma$ (which could be a vector or a scalar). Assume that there is some control $u$ which controls the infinitesimal generator $Q_u$ of the Markov process on type $i$ and potentially $\gamma$. Also assume that the agent can control the growth rate $\dot{\gamma}$ at some cost. The feasibility set of the controls is $(u, \dot{\gamma}) \in U(t, i, z, \gamma)$ 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 $\gamma$, (b) The minimum of the cost function is 0 and in the interior of the feasibility set: $\min_{(\dot{\gamma}, u) \in U(t, i, z, \gamma)} c(t, z, \dot{\gamma}, i, \gamma, u) = 0$, for all $t, \gamma, i$ and (c) the value of a jump to the frontier, $\bar{v}(t)$, is identical for all agent states (e.g. $\bar{v}(t) = v(t, \bar{z}(t)) = v_h(t, \bar{z}(t))$).\(^{27}\)

Then the normalization of the firm’s problem gives the following set of necessary conditions,

\[
(r - g(t))v_i(t, z, \gamma) = \max_{(\dot{\gamma}, u) \in U(t, i, z, \gamma)} \left\{ e^z - c(t, z, \dot{\gamma}, i, \gamma, u) + (\dot{\gamma} - g) \frac{\partial v_i(t, z, \gamma)}{\partial z} + \frac{\partial v_i(t, z, \gamma)}{\partial t} + e_1 \cdot Q_u \cdot v(t, z, \gamma) + \eta(\bar{v}(t) - v_i(t, z, \gamma)) \right\} \tag{A.21}
\]

\[
v_i(t, \bar{z}(t, i, \gamma), \gamma) = v_s(t) \tag{A.22}
\]

\[
\frac{\partial v_i(t, \bar{z}(t, i, \gamma), \gamma)}{\partial z} = 0 \tag{A.23}
\]

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

\[
v_i(t, 0, \gamma) = v_s(t) \tag{A.24}
\]

\[
\frac{\partial v_i(t, 0, \gamma)}{\partial z} = 0 \tag{A.25}
\]

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

\[
(r - g(t))v_s(t) = \max_{(\dot{\gamma}, u) \in U(t, i, z, \gamma)} \left\{ 1 - c(t, z, \dot{\gamma}, i, \gamma, u) + e_1 \cdot Q_u \cdot v_s(t) + \eta(\bar{v}(t) - v_s(t)) + v'_s(t) \right\} \tag{A.26}
\]

Since in order to be a valid intensity matrix, all rows in $Q_u$ add to 0 for any $u$, the last term is 0 for any $i$ or control $u$,

\[
(r - g(t))v_s(t) = \max_{(\dot{\gamma}, u) \in U(t, i, z, \gamma)} \left\{ 1 - c(t, z, \dot{\gamma}, i, \gamma, u) + v'_s(t) + \eta(\bar{v}(t) - v_s(t)) \right\} \tag{A.27}
\]

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

\[
(r - g(t))v_s(t) = 1 + v'_s(t) + \eta(\bar{v}(t) - v_s(t)) \text{ for all } i \tag{A.28}
\]

Therefore, the necessary conditions for optimal stopping are identical for all $i, \gamma, z$, confirming our guess. Furthermore, (A.28) provides an ODE for $v_s(t)$ based on aggregate $g(t)$ and $\bar{v}(t)$ changes. Solving this in a stationary environment gives an expression for $v_s$ in terms of equilibrium $g$ and $\bar{v}$,

\[
v_s = \frac{1 + \eta \bar{v}}{r - g + \eta} \tag{A.29}
\]

\(^{27}\)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.
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,

\[ \alpha \equiv (1 + \lambda) \frac{S}{g} \]  \hspace{1cm} (B.1)

\[ \dot{\lambda} \equiv \frac{\lambda_f}{\lambda_h} \]  \hspace{1cm} (B.2)

\[ \bar{\lambda} \equiv \frac{\lambda_f}{r - g + \lambda_h} + 1 \]  \hspace{1cm} (B.3)

\[ \nu = \frac{(r - g) \dot{\lambda}}{g} \]  \hspace{1cm} (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 (34) and solve for \( F_h(z) \)

\[ F_h(z) = \dot{\lambda} F_f(z) \]  \hspace{1cm} (B.5)

Substitute into (33)

\[ S = g F_f(z) + \left( \dot{\lambda} + 1 \right) S F_f(z) \]  \hspace{1cm} (B.6)

Solve this as an ODE in \( F_f(z) \), subject to the \( F_f(0) = 0 \) boundary condition in (35)

\[ F_f(z) = \frac{1}{1 + \lambda} e^{-\alpha z} \]  \hspace{1cm} (B.7)

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

\[ \lim_{z \to \infty} (F_f(z) + F_h(z)) = 1 \]  \hspace{1cm} (B.8)

Differentiating (B.7),

\[ F_f'(z) = \frac{\alpha}{1 + \lambda} e^{-\alpha z} \]  \hspace{1cm} (B.9)

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

\[ F'(z) = \alpha e^{-\alpha z} \]  \hspace{1cm} (B.10)

Solve (30) for \( v_h(z) \)

\[ v_h(z) = \frac{e^z + \lambda_h v_f(z)}{r - g + \lambda_h} \]  \hspace{1cm} (B.11)

Substituting into (29) gives the following ODE in \( v_f(z) \)

\[ (r - g)v_f(z) = e^z + \lambda_h \dot{\lambda} \left( -v_f(z) + \frac{e^z + \lambda_h v_f(z)}{r - g + \lambda_h} \right) - g v_f'(z) \]  \hspace{1cm} (B.12)
Using the constant definitions and simplifying

\[(r - g)v'_\ell(z) = e^z - \frac{gv'_\ell(z)}{\lambda}\]  \hspace{1cm} (B.13)

Solve this ODE subject to the smooth pasting condition in (32) and simplify,

\[v'_\ell(z) = \frac{\tilde{\lambda}}{g + (r - g)\lambda}e^z + \frac{1}{(r - g)(\nu + 1)}e^{-z'\nu}\]  \hspace{1cm} (B.14)

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

\[v'_\ell(0) = \frac{1}{r - g}\]  \hspace{1cm} (B.15)

Substitute (B.10), (B.14) and (B.15) into the value matching condition in (31) and simplify

\[
\frac{1}{r - g} = \int_0^\infty \left[ \frac{e^z\left(\lambda - \frac{\alpha - \frac{\lambda}{r}}{g}\right)\alpha g}{(g - r)(-r\lambda + g(\lambda - 1))} + \frac{e^{z-2\alpha}\alpha \tilde{\lambda}}{g + r\lambda - g\lambda} \right] dz - \zeta \hspace{1cm} (B.16)
\]

Evaluating the integral,

\[
\zeta = \frac{\alpha \left(-r\lambda + g\left(\lambda - \alpha + 1\right)\right)}{(g - r)(r\lambda + g(\alpha - \lambda))\left(\alpha - 1\right)} - \frac{1}{(r - g)(\nu + 1)} - \frac{\tilde{\lambda}}{g + r\lambda - g\lambda} \hspace{1cm} (B.17)
\]

Substitute for \(\alpha\) gives an implicit equation in \(S\)

\[0 = \zeta + \frac{g}{(r - g)\left(\frac{1}{r - g} + \frac{\lambda}{s - g + s\lambda} - \frac{\lambda}{s - g\lambda + r\lambda + s\lambda}\right)} - \frac{1}{(r - g)(\nu + 1)} \hspace{1cm} (B.18)\]

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,

\[
S = \lambda_h \left(\frac{r(r + \lambda_h + \lambda_e) - \sqrt{\zeta((4g + r^2\zeta)(-g + r + \lambda_h)^2 + 2(-2g + (r - g)r\zeta)(g - r - \lambda_h)\lambda_e + (g - r)^2\zeta\lambda_e^2 + \zeta g^2 + \zeta(-g)(3r + 2\lambda_h + \lambda_e))}}{2\zeta(\lambda_h + \lambda_e)(g - r - \lambda_h)}\right) \hspace{1cm} (B.19)
\]

From this \(S\), \(\alpha\) can be calculated through (B.1) and the the rest of the equilibrium follows. \(\Box\)

### B.2 Bounded Support

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

\[\alpha \equiv \frac{(1 + \tilde{\lambda})S - \eta}{g}\]  \hspace{1cm} (B.20)

\[\dot{\lambda} \equiv \frac{\lambda_e}{\eta + \lambda_h}\]  \hspace{1cm} (B.21)

\[\tilde{\lambda} \equiv \frac{r - g + \lambda_e + \lambda_h}{r - g + \lambda_h}\]  \hspace{1cm} (B.22)

\[\nu \equiv \frac{\dot{\lambda}}{g}\]  \hspace{1cm} (B.23)
Solve for $F_h(z)$ in (53),

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

Substitute this back into (52) to get an ODE in $F_\ell$

$$0 = g F'_\ell(z) + (S - \eta)(1 + \hat{\lambda}) F_\ell(z) + \eta H (z - \bar{z}) - S \quad \text{(B.25)}$$

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

$$F_\ell(z) = \begin{cases} 
\frac{S}{(S - \eta)(1 + \lambda)} \left(1 - e^{-\alpha z}\right) & 0 \leq z < \bar{z} \\
\frac{S}{(S - \eta)(1 + \lambda)} \left(1 - e^{-\alpha \bar{z}}\right) & z = \bar{z}
\end{cases} \quad \text{(B.26)}$$

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

$$F(z) = (1 + \hat{\lambda}) F_\ell(\bar{z}) = \frac{S}{S - \eta} \left(1 - e^{-\alpha \bar{z}}\right) \quad \text{(B.27)}$$

$$F'(z) = \frac{\alpha S}{S - \eta} e^{-\alpha z} \quad \text{(B.28)}$$

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

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

The pdf of the unconditional distribution is,

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

To solve for the value solve (49) for $v_h(z)$,

$$v_h(z) = e^z + (\lambda_h - \eta) v_\ell(z) + \eta \nu_\ell(\bar{z}) \quad \text{(B.32)}$$

Substitute into (48) and simplify

$$(r - g + \eta) v_\ell(z) = e^z + \eta \nu_\ell(\bar{z}) - \frac{g}{\alpha} v'_\ell(z) \quad \text{(B.33)}$$

Solving (51) and (B.33) and simplifying,

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

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

$$v_\ell(\bar{z}) = \left(-\frac{\eta}{g - r} + 1\right) \left(\frac{e^{\bar{z} \bar{\lambda}}}{g + (\eta + r - g) \lambda} + \frac{e^{-\nu \bar{z}}}{(\eta + r - g)(\nu + 1)}\right) \quad \text{(B.35)}$$
Substitute (B.35) into (B.34) to find an expression for \(v_\ell(z)\)

\[
v_\ell(z) = \frac{\bar{\lambda}}{g(1 + \nu)} \left( e^{\bar{z}} + \frac{1}{\nu} e^{-\nu \bar{z}} + \frac{\eta}{r - g} \left( e^{\bar{z}} + \frac{1}{\nu} e^{-\nu \bar{z}} \right) \right)
\]

Substitute (B.30) and (B.36) into the value matching condition in (50) and evaluate the integral,

\[
\zeta + \frac{1}{r - g} = \frac{S\alpha\bar{\lambda} \left( -e^{-\nu \bar{z}}(-1+e^{-\alpha \bar{z}})\eta + e^{\bar{z}} \eta(e^{-\alpha \bar{z}} - 1) - e^{-\alpha z} + 1 - e^{\bar{z} - \alpha z} + 1 - e^{\bar{z} - \alpha z} + 1 \right)}{g(S - \eta)(\nu + 1)}
\]

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

\[
\zeta + \frac{1}{r - g} = \frac{S\lambda(\lambda + 1) \left( -\left(\frac{S}{\eta}\right)^{-1} - \frac{S}{\eta(1+\lambda)} + 1 \right) + g \left( \frac{1}{\nu + (S - \eta)(\lambda + 1)} \frac{\eta}{g(S - \eta)(\lambda + 1)} - \frac{g}{g + (S - \eta)(\lambda + 1)} \right)}{g^2(\nu + 1)}
\]

### B.3 Stationary Stochastic Innovation Equilibrium with Infinite Support

**Proof of Proposition 3.** Define \(0, 1, I\) as a vector of 0, 1, and the identity matrix and the following:

\[
A = \begin{bmatrix} \frac{1}{\nu + (g - \gamma)} \\ \frac{g}{\nu + (g - \gamma)} \end{bmatrix}, \quad B = \begin{bmatrix} r + \lambda_c - g & -\frac{\lambda_c}{g - \gamma} \\ \frac{\lambda_c}{g - \gamma} & -\frac{\lambda_c}{g - \gamma} \end{bmatrix}, \quad C = \begin{bmatrix} gF_\ell'(0) + (g - \gamma)F_h'(0) - \lambda_c \\ \frac{\lambda_c}{g - \gamma} \end{bmatrix}, \quad D = \begin{bmatrix} F_\ell'(0) \\ F_h'(0) \end{bmatrix}
\]

\[
\bar{F}(z) = \begin{bmatrix} F_\ell(z) \\ F_h(z) \end{bmatrix}, \quad v(z) = \begin{bmatrix} v_\ell(z) \\ v_h(z) \end{bmatrix}
\]

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

\[
v'(z) = Ae^z - Bv(z)
\]

\[
v'(0) = 0
\]

\[
\bar{F}'(z) = -C\bar{F}(z) + D
\]

\[
\bar{F}'(0) = 0
\]

\[
\bar{F}(\infty) \cdot 1 = 1
\]

\[
v_\ell(0) = v_h(0) = \int_{0}^{\infty} (v(z)^T \cdot \bar{F}'(z)) \, dz - \zeta
\]

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

\[
\bar{F}(z) = \left( e^{Az} - 1 \right) A^{-1} b
\]

\[\text{The equation } \bar{F}(z) = AF[z] + b \text{ subject to } \bar{F}(0) = 0 \text{ has the solution,}
\]

\[
\bar{F}(z) = \left( e^{Az} - 1 \right) A^{-1} b
\]
\[ v(z) = (I + B)^{-1} \left( e^{Iz} + e^{-Bz} B^{-1} \right) A \]  \hfill (B.50)

Evaluate at \( z = 0 \),

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

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

\[ \bar{F}(z) = (I - e^{-Cz}) C^{-1} D \]  \hfill (B.52)

Take the derivative,

\[ \bar{F}'(z) = e^{-Cz} D \]  \hfill (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 (72) and (73) 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

\[ S_l + S_h > \frac{(g - \gamma) \lambda_l + g \lambda_h}{(g - \gamma) + g} \]  \hfill (B.54)

\[ S_h + S_l > \lambda_h + \lambda_l \]  \hfill (B.55)

and for \( B \) the corresponding conditions are

\[ r > g > \gamma \]  \hfill (B.56)

\[ r - g + \lambda_h + \lambda_l > 0 \]  \hfill (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 \( z = 0 \),

\[ \bar{F}'(0) = D \]  \hfill (B.58)

Take the limit of (B.52)

\[ \bar{F}(\infty) = C^{-1} D \]  \hfill (B.59)

The derivation of these results uses that \( \int_0^T e^{tA} dt = A^{-1} (e^{TA} - I) \). With appropriate conditions on eigenvalues, this implies that \( \int_0^\infty e^{tA} dt = -A^{-1} \).

Equations of the form \( v'(z) = Ae^z - B \cdot v(z) \) with the initial condition \( v'(0) = 0 \) have the solution,

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

This derivation exploits commutativity, as both \( e^{Bz} \) and \( (I + B)^{-1} \) can be expanded as power series of \( B \).
And (B.46) becomes

$$1 = C^{-1}D \cdot 1$$  \hspace{1cm} (B.60)

We can check that, by construction, with the $C$ and $D$ defined by (B.40), (B.60) is fulfilled for any $F'_h(0), F'_l(0), \lambda_l$, and $\lambda_h$.

For $\tilde{F}'(z)$ to define a valid pdf it is necessary for $\tilde{F}'(z) > 0$ 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 $C$. The unique constant of proportionality is determined by (B.60). The two eigenvectors of $C$ fulfilling this proportionality are,

$$\nu_i \equiv \begin{bmatrix} -F'_h(0)(\gamma(g-\gamma)F'_l(0)+2(\gamma-\gamma)\lambda_l(\gamma(g\gamma F'_h(0)+F'_l(0)\gamma F'_h(0))-\gamma F'_h(0))-\gamma \lambda_h)\gamma F'_h(0)+\gamma F'_l(0) - \gamma \lambda_h)^2 + (\gamma - \gamma)^2 \lambda_l^2 + \gamma \lambda F'_h(0) + \gamma \lambda_l + \gamma \lambda_l^2 \\ F'_h(0) \end{bmatrix}^{2 \lambda_l}$$  \hspace{1cm} (B.61)

Denote $\nu$ 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,

$$D = \nu$$  \hspace{1cm} (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,

$$F'_l(0) = \frac{F'_h(0)\lambda_h}{\gamma F'_h(0) + \lambda_l}$$  \hspace{1cm} (B.63)

Solving this equation for $F'_l(0)$ and choosing the positive root

$$F'_l(0) = \frac{F'_h(0)\lambda_h}{\gamma F'_h(0) + \lambda_l}$$  \hspace{1cm} (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,

$$\frac{1}{r-g} + \zeta = \int_0^\infty \left[ [(I+B)\zeta][e^{Iz} + e^{-Bz}B^{-1}] \right] A^T e^{-CzD} dz$$  \hspace{1cm} (B.65)

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

The tail index of the unconditional distribution, $F(z) \equiv F_l(z) + F_h(z)$ 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 $F'_l(0)$ from above, the smallest eigenvalue of $C$ is

$$\alpha \equiv \frac{(g-\gamma)F'_h(0) - \lambda_l}{g(\gamma - \gamma)F'_h(0) + \lambda_l}$$  \hspace{1cm} (B.66)

\footnote{Since $C > 0$ and irreducible (in this case off diagonals not zero), then by Perron-Frobenius it has a simple dominant real root $\alpha$ and an associated eigenvector $\nu > 0$. Hence, as $\tilde{F}(0) = 0, F_l(\infty) + F_h(\infty) = 1$, and $\tilde{F}'(z) > 0$, we have a valid pdf. This uniqueness of the $\nu$ solution only holds if the other eigenvector of $C$ has a positive and negative coordinate, which always holds in our model.}
Solving (B.66) for $F'_h(0)$ as a function of $\alpha$,

$$F'_h(0) = \frac{g \left( \alpha \gamma - \lambda_h + \sqrt{(\lambda_h - \alpha \gamma)^2 + 2\lambda_l (\alpha \gamma + \lambda_h) + \lambda_l^2 - \lambda_l} \right) + 2\gamma \lambda_l}{2\gamma (g - \gamma)}$$

(B.67)

Substituting for $F'_h(0)$ into $C$ and $D$ gives a function in terms of $g$ and $\alpha$,

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

(B.68)

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

(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.

\[\square\]

Comparing (D.31) with (82) shows that the positivity of the tail index $\alpha$ is now equivalent to $C$ having positive eigenvalues. For the decomposition of the option value, comparing (D.32) with (84) 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 (70) and (71), a system of linear ODEs. If we define the unconditional distribution $F(z) \equiv F_i(z) + F_h(z)$, and if both $F_i(z)$ 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, $\alpha$, is defined as that of the unconditional distribution, $F(z)$. The ODE solution for the vector $\vec{F}(z)$ given in Proposition 3 by (82) 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 $\alpha$ by the construction of (79).

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 $F'_i(0) > 0$ 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 $F_i(z)$ and $F_h(z)$ remain positive and satisfy (16) and (17) 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 $\alpha$, to solve for $F'_h(0)$, which then determines $F'_i(0)$ from the eigenvector restriction. This then allows us to obtain the expressions (79) and (80) in terms of parameters, $\alpha$ and $g$. Then value matching, (B.47) and (B.65), gives us expression (81) to define $g$ in terms $\alpha$, so we end up with a continuum of stationary equilibria parametrized by $\alpha$. 

40
Appendix C  Endogenous Markov Innovation

C.1 Stationary BGP with a Endogenous, Continuous Choice, and a Finite, Unbounded Technology Frontier

Proof of Proposition 4. Assuming an interior solution, taking the first order necessary condition of the Hamilton-Jacobi-Bellman equation in (86), and reorganizing

\[ \gamma(z) = \frac{1}{2} e^{-z} v'_h(z) \] (C.1)

Substituting this back into (86) gives a non-linear ODE,

\[ (r - g)v_h(z) = e^z - gv'_h(z) + \frac{1}{4} e^{-z} v'_h(z)^2 + \lambda_h (v_{\ell}(z) - v_h(z)) \] (C.2)

As the equilibrium has finite support, the growth rate of the economy will be the growth rate of the frontier,

\[ g \equiv \lim_{z \to \infty} \gamma(z) \] (C.3)

For large \( z \), the option value of diffusion approaches 0, and the choice of \( \gamma \) can be done assuming only a transversality condition, rather than the initial value problem with optimal stopping. Guess asymptotic solutions of the following form,\(^{30}\)

\[ v_{\ell}^{\infty}(z) = c_{\ell} e^{z} \] (C.4)
\[ v_{h}^{\infty}(z) = c_{h} e^{z} \] (C.5)

Substituting this guess into (85) and (C.2) and simplifying leads to the following system of quadratic equations after using undetermined coefficients,

\[ (r + \lambda_{\ell}) c_{\ell} = 1 + \lambda_{\ell} c_{h} \] (C.6)
\[ (r + \lambda_{h}) c_{h} = 1 + \lambda_{h} c_{\ell} + \frac{1}{4} c_{h}^2 \] (C.7)

Solving this system for \( c_{\ell} \) and \( c_{h} \) gives a pair of roots,

\[ c_{\ell} = \frac{\lambda_{\ell} \left( \frac{1}{2} \pm \sqrt{\lambda_{h} + \lambda_{\ell} + r} \left( r^2 (\lambda_{h} + \lambda_{\ell} + r) - \chi (\lambda_{\ell} + r) \right) \right) + r \left( \frac{1}{2} (\lambda_{h} + \lambda_{\ell} + r) + \frac{1}{2} \right) + r^2 \lambda_{\ell}}{\lambda_{\ell}} \] (C.8)

\[ c_{h} = \frac{r (\lambda_{h} + \lambda_{\ell} + r) \pm \sqrt{\lambda_{h} + \lambda_{\ell} + r} \left( r^2 (\lambda_{h} + \lambda_{\ell} + r) - \chi (\lambda_{\ell} + r) \right)}}{\frac{1}{2} (\lambda_{\ell} + r)} \] (C.9)

The two roots to the system have matching plus (or minus) signs for these constants. The verification step shows that this is a particular solution of this system. Use of the transversality conditions would eliminate terms for the general solution. Using the solution to this system of equations and (C.3) and (C.5),

\[ g = \frac{1}{2} e^{-z} (c_{h} e^{z}) \] (C.10)

\(^{30}\)More rigorously, do a change of variables \( w_{i}(z) \equiv v'_{i}(z) e^{-z} \). This is a stationary ODE subject to the new initial condition \( w_{i}(z) = 0 \), where the solutions finds the stationary value is \( w_{i}(\infty) = c_{i} \) and \( g \equiv \gamma(\infty) = \frac{1}{2} c_{h} \). To convert back to \( v_{i}(z) \), use (87).
Substituting from (C.9) and simplifying
\[ g = \frac{r (\lambda_h + \lambda_\ell + r) \pm \sqrt{(\lambda_h + \lambda_\ell + r) (r^2 (\lambda_h + \lambda_\ell + r) - \chi (\lambda_\ell + r))}}{\lambda_\ell + r} \]  

(C.11)

Define
\[ \bar{\lambda} = \frac{r + \lambda_\ell + \lambda_h}{r + \lambda_\ell} \]  

(C.12)

Then, choose the negative root (to ensure \( g < r \)) and simplify,
\[ g = \bar{\lambda} r \left[ 1 - \sqrt{1 - \frac{\chi}{\lambda r^2}} \right] \]  

(C.13)

As \( \gamma(0) = 0 \), Evaluating (89) and (90) at \( z = 0 \) gives,
\[ S_{\ell} = gF_{\ell}'(0) \]  

(C.14)
\[ S_{h} = gF_{h}'(0) \]  

(C.15)

For the initial values, note that if \( v'_{\ell}(0) = v'_{h}(0) = 0 \), then to fulfill the ODE, the initial value must be,
\[ v_{\ell}(0) = v_{h}(0) = \frac{1}{r-g} \]  

(C.16)

**Parameter Requirements**  From (C.13), \( g > 0 \) always holds, and a necessary condition for \( r > g \) is,
\[ r > \sqrt{\frac{\lambda}{\bar{\lambda}}} \]  

(C.17)

\[ \square \]

C.2 Stationary BGP with a Endogenous, Continuous Choice, and a Finite, Bounded Technology Frontier

**Proof of Proposition 5.** Note that Section 3.2 nests Section 3.1 when \( \eta = 0 \)

**Nested Derivation of Stationary HBJE**  Assuming an interior solution, take the first order necessary condition of the Hamilton-Jacobi-Bellman equation in (103), and reorganize
\[ \gamma(z) = \frac{\lambda}{2} e^{-z} v'_{h}(z) \]  

(C.18)

Substitute this back into (103) to get a non-linear ODE,
\[ (r - g)v_{h}(z) = e^{-z} - g v'_{h}(z) + \frac{\lambda}{2} e^{-z} v'_{h}(z)^2 + \lambda_h (v_{\ell}(z) - v_{h}(z)) + \eta (v_{\ell}(z) - v_{h}(z)) \]  

(C.19)

To create a stationary solution for the value function define a change of variables,
\[ w_{\ell}(z) \equiv e^{-z} v'_{\ell}(z) \]  

(C.20)

From (105),
\[ w_{\ell}(0) = w_{h}(0) = 0 \]  

(C.21)
Differentiate and reorganize (C.20),

\[ e^{-z}v''_i(z) = w'_i(z) + w_i(z) \] (C.22)

Differentiate (102),

\[ (r - g)v'_i(z) = e^{-z} - gv''_i(z) + \lambda_e(v'_h(z) - v'_i(z)) - \eta v'_i(z) \] (C.23)

Multiply by \( e^{-z} \) and use (C.20) and (C.22)

\[ (r + \lambda_e + \eta)w(z) = 1 - gw'_i(z) + \lambda ew_h(z) \] (C.24)

Note that,

\[ e^{-z} \partial_z (e^{-z}v_h(z)) = 2e^{-z}v''_h(z)e^{-z}v'_h(z) - (e^{-z}v'_h(z))^2 \] (C.25)

\[ = 2w_h(z)w'_h(z) + w_h(z)^2 \] (C.26)

Differentiate (C.19), multiply by \( e^{-z} \), and use (C.20), (C.22) and (C.26)

\[ (r + \lambda_h + \eta)w_h(z) = 1 - (g - \frac{1}{2}w_h(z))w'_h(z) + \lambda hw_e(z) + \frac{1}{2}w_h(z)^2 \] (C.27)

From (C.18),

\[ \gamma(z) = \frac{2}{r}w_h(z) \] (C.28)

\[ g = \frac{2}{r}w_h(z) \] (C.29)

Integrate (C.20) with the initial value from (104) to get,

\[ v_i(z) = v_s + \int_0^ze^{\tilde{z}}w_i(\tilde{z})d\tilde{z} \] (C.30)

Substitute (C.30) into (A.29) and rearrange to get an expression for \( v_s \) in terms of \( w_e \) and intrinsics,

\[ v_s = \frac{1 + \eta w_e(z)}{r - g + \eta} = \frac{1 + \eta \int_0^ze^{\tilde{z}}w_e(\tilde{z})d\tilde{z}}{r - g} \] (C.31)

**Endogenous choice of \( \theta \) and \( \kappa \):** Take the value matching condition for the choice of the idiosyncratic \( \hat{\theta} \) and \( \hat{\kappa} \) given equilibrium \( \theta \) and \( \kappa \) choices of the other firms.

\[ v_s \equiv v_s(0) = v_h(0) = \max_{\theta \geq 0, \kappa > 0} \left\{ (1 - \hat{\theta}) \int_0^z v_i(z; \theta)dF(z; \theta)^\kappa + \hat{\kappa} v_i(z; \theta) - \zeta - \frac{1}{\xi} \hat{\theta}^2 - \frac{1}{3} \kappa^2 \right\} \] (C.32)

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 and then letting \( \hat{\theta} = \theta \) in equilibrium gives,

\[ \theta = \zeta \left( v_i(\tilde{z}; \theta) - \int_0^\tilde{z} v_i(z; \theta)dF(z; \theta)^\kappa \right) \] (C.33)

Reorganize,

\[ \int_0^\tilde{z} v_i(z)dF(z)^\kappa = v_i(\tilde{z}) - \frac{\theta}{\zeta} \] (C.34)
Substitute into (C.32) at the optimal $\theta$

$$v_\ell(0) = (1 - \theta) \left( v_\ell(\bar{z}) - \frac{\theta}{\zeta} \right) + \theta v_\ell(\bar{z}) - \zeta - \frac{1}{\zeta} \theta^2$$

(C.35)

Solve the quadratic for $\theta$ and choose the interior (i.e., negative) root,

$$\theta = 1 - \sqrt{1 - \zeta (v_\ell(\bar{z}) - v_\ell(0) - \zeta)}$$

(C.36)

Substitute from (C.30),

$$\theta = 1 - \sqrt{1 - \zeta \left( \int_0^\bar{z} e^z w_\ell(\bar{z}) d\bar{z} - \zeta \right)}$$

(C.37)

This equation provides an equilibrium expression for the optimal choice of $\theta$. Similarly for $\kappa$, first write down (C.32) with the conversion to $w_i(z)$ space

$$\zeta = \max_{\theta \geq 0, \kappa > 0} \left\{ \hat{w}_\ell(\bar{z}) - (1 - \theta) \int_0^\bar{z} e^z w_\ell(z) F(z)^\kappa dz - \frac{1}{\zeta} \hat{\theta}^2 - \frac{1}{\zeta} \hat{\kappa}^2 \right\}$$

(C.38)

Take the first order condition and equation $\hat{\kappa} = \kappa$ in the economy. Note that $F(z)^\kappa = \exp(\kappa \log F(z))$, and assume conditions to differentiate under the integral

$$\kappa = -\frac{\vartheta (1 - \theta)}{2} \int_0^\bar{z} e^z w_\ell(z) \log(F(z)) F(z)^\kappa dz$$

(C.39)

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

**KFE and Value Matching** From (106), for $z < \bar{z}$ the KFE is,

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

(C.40)

Note that since $\gamma(0) = 0$, $S_\ell = g F_\ell'(0)$ and $S_h = g F_h'(0)$. For the value matching, first use the generic result from (D.80) that,

$$\mathbb{E} [v_\ell(z)] = \int_0^\bar{z} v_\ell(z) dF(z) dz = \int_0^\bar{z} v_\ell(z)(1 - F(z)^\kappa) dz + v_\ell(0)$$

(C.41)

Substitute (C.20) and (C.41) into (104) at the optimal $\theta$

$$v_s = (1 - \theta) \left( \int_0^\bar{z} e^z w_\ell(z)(1 - F(z)^\kappa) dz + v_s \right) + \theta \left( v_s + \int_0^\bar{z} e^z w_\ell(z) dz \right) - \zeta - \frac{1}{\zeta} \theta^2 - \frac{1}{\zeta} \kappa^2$$

(C.42)

Simplifying gives an expression in terms of $F(z)$ and $w(z)$,

$$\zeta = \begin{cases} \int_0^\bar{z} e^z w_\ell(z)(1 - F(z)^\kappa) dz - \frac{1}{\zeta} \kappa^2 & \text{if } \theta = 0 \\ \int_0^\bar{z} e^z w_\ell(z) dz - (1 - \theta) \int_0^\bar{z} e^z w_\ell(z) F(z)^\kappa dz - \frac{1}{\zeta} \theta^2 - \frac{1}{\zeta} \kappa^2 & \text{if } \theta > 0 \end{cases}$$

(C.43)

For the unbounded case where $\eta = \theta = 0$, and $\bar{z} \to \infty$, we can check the asymptotic value comes from (C.5), (C.8), (C.9) and (C.20)

$$\lim_{z \to \infty} w_i(z) = c_i$$

(C.44)
Upper bound on \( g \): 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, \( \lim_{\bar{z} \to \infty} w_i'(z) = 0 \) as \( w_i(z) \) have been constructed to be stationary. Therefore, looking at the asymptotic limit of (C.24) and (C.27),

\[
(r + \lambda_\ell + \eta)w_\ell(\infty) = 1 + \lambda_\ell w_h(\infty) \tag{C.45}
\]

\[
(r + \lambda_h + \eta)w_h(\infty) = 1 + \lambda_h w_\ell(\infty) + \frac{\lambda}{4} w_h(\infty)^2 \tag{C.46}
\]

Comparing this system of 2 equations in \( w_\ell(\infty) \) and \( w_h(\infty) \) to (C.6) and (C.7) shows the equations to be of the identical form, except \( r \) has been augmented by \( \eta \). Using the solution procedure in that section, define an equivalent to (C.12),

\[
\bar{\lambda} \equiv \frac{r + \eta + \lambda_\ell + \lambda_h}{r + \eta + \lambda_\ell} \tag{C.47}
\]

Where an equivalent to (C.13) gives an upper bound on the growth rate

\[
g < \bar{\lambda}(r + \eta) \left[ 1 - \sqrt{1 - \frac{X}{\bar{\lambda}(r + \eta)^2}} \right] \tag{C.48}
\]

The full set of equations to solve for \( F_i(z) \) and \( w_i(z) \) is, Summarizing the full set of equations to solve for \( F_i(z) \) and \( w_i(z) \) from (107) to (109), (C.14), (C.15), (C.21), (C.24), (C.27) to (C.31), (C.37), (C.39), (C.40), (C.43) and (C.47).

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).\(^{31,32}\)

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 \( \kappa \), 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 \( \gamma \) and variance \( \sigma \),


\(^{32}\)For a version of the model using monopolistic competition and associated equilibrium conditions, see Appendix D.6 and Technical Appendix B.
the Kolmogorov Forward Equation (KFE) below (in cdfs) is\(^3\)

\[
\partial_t \Phi(t, Z) = -\left(\gamma - \sigma^2/2\right)Z \frac{\partial}{\partial Z} \Phi(t, Z) + \frac{\sigma^2}{2} Z^2 \frac{\partial^2}{\partial Z^2} \Phi(t, Z) + \frac{S(t)}{Z} \Phi(t, Z) - S(t),
\]

for \(M(t) \leq Z \leq B(t)\) \hspace{1cm} (D.1)

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 \(Z\).\(^4\)

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

\[
rV(t, Z) = Z + (\gamma + \sigma^2/2)Z \frac{\partial}{\partial Z} V(t, Z) + \frac{\sigma^2}{2} Z^2 \frac{\partial^2}{\partial Z^2} V(t, Z) + \partial_t V(t, Z)
\]

\[
V(t, M(t)) = \int_{M(t)}^{B(t)} V(t, Z) d\Phi(t, Z) - \zeta M(t)
\]

\[
\frac{\partial}{\partial t} \Phi(t, Z) = -\left(\gamma - \sigma^2/2\right)Z \frac{\partial}{\partial Z} \Phi(t, Z) + \frac{S(t)}{Z} \Phi(t, Z) - S(t)
\]

\[
\Phi(t, M(t)) = 0
\]

\[
\Phi(t, B(t)) = 1
\]

\[
B'(t)/B(t) = \gamma, \text{ if } \sigma > 0
\]

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.

\(^3\)To derive from the more common KFE written in pdfs, use the adjoint of the infinitesimal generator of GBM,

\[
\frac{\partial \phi(t, Z)}{\partial t} = -\frac{\partial}{\partial Z} \left((\mu - \sigma^2/2)Z \phi(t, Z)\right) + \frac{\partial}{\partial Z} \left(\frac{\sigma^2}{2} Z^2 \phi(t, Z)\right)
\]

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,

\[
\frac{\partial \Phi(t, Z)}{\partial t} = (\nu^2 - \sigma^2)Z \frac{\partial \Phi(t, Z)}{\partial Z} + \frac{\nu^2}{2} \frac{\partial^2 \Phi(t, Z)}{\partial Z^2} + \ldots
\]

\(^4\)The sequential formulation and connection to a recursive optimal stopping of a deterministic process is given on page 110-112 of Stokey (2009).
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

\[
(r - g(t))v(t, z) = e^z + (\gamma - g(t))\frac{\partial}{\partial z}v(t, z) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial z^2}v(t, z) + \partial_\nu v(t, z)
\]

\[
v(t, 0) = \int_0^\infty v(t, z) dF(t, z)^\kappa - \zeta
\]

\[
\frac{\partial}{\partial z}v(t, 0) = 0
\]

\[
0 = (g(t) - \gamma)\frac{\partial}{\partial z}F(t, z) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial z^2}F(t, z) + S(t)F(t, z)^\kappa - S(t)
\]

\[
F(t, 0) = 0
\]

\[
F(t, B(t)) = 1
\]

\[
S(t) = (g(t) - \gamma)\frac{\partial}{\partial z}F(t, 0) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial z^2}F(t, 0)
\]

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

\[
(r - g)v(z) = e^z + (\gamma - g)v'(z) + \frac{\sigma^2}{2} v''(z)
\]

\[
v(0) = \int_0^\infty v(z) dF(z)^\kappa - \zeta
\]

\[
v'(0) = 0
\]

\[
0 = (g - \gamma)F'(z) + \frac{\sigma^2}{2} F''(z) + SF(z) - S
\]

\[
F(0) = 0
\]

\[
F(\infty) = 1
\]

\[
S = (g - \gamma)F'(0) + \frac{\sigma^2}{2} F''(0)
\]

The value matching condition in (D.17) can also be written, using (C.41), as

\[
\zeta = \int_0^\infty v'(z) (1 - F(z)^\kappa) \, dz
\]

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 \( \sigma \) 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 \( r > \gamma + (\zeta(\alpha - 1))^{-1} > 0 \), then*

\[
g = \frac{1 - \zeta(\alpha - 1)(r - \alpha\gamma)}{\zeta(\alpha - 1)^2},
\]

\[
v(z) = \frac{1}{r - \gamma} e^z + \frac{1}{\nu(r - \gamma)} e^{-\nu z},
\]

47
where,
\[
\nu \equiv \frac{r - g}{g - \gamma} > 0, \quad (D.26)
\]
and the stationary distribution in logs is
\[
F(z) = 1 - e^{-\alpha z}. \quad (D.27)
\]

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, \(\nu\) 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 \(\alpha\). 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.\(^{35}\) 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.

\(^{35}\)Without endogenous adoption there is no “absorbing” or “reflecting” barrier and geometric random shocks lead to a diverging variance in the KFE.
**Proposition 7** (Equilibrium with Geometric Brownian Motion Innovations). A continuum of equilibria parameterized by $\alpha$ exist satisfying

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

and

\[
0 > (\alpha - 1)^2 \alpha \zeta \sigma^4 - 2(\alpha - 1)\zeta \sigma^2 ((\alpha - 3)\alpha + (\alpha^2 - 1) \zeta(r - \gamma)) + 4((\alpha - 1)\zeta(r - \gamma) - 1)^2.
\]  
(D.29)

For a given $\alpha$, the growth rate is

\[
g = \gamma + \frac{1 - (\alpha - 1)\zeta(r - \gamma)}{(\alpha - 1)^2\zeta} + \frac{\sigma^2 \alpha \left( (\alpha - 1) \left( r - \gamma - \frac{\sigma^2}{2} \right) \zeta - 2 \right) + 1}{2 \left( (\alpha - 1) \left( r - \gamma - \frac{\sigma^2}{2} \right) \zeta - 1 \right)}
\]  
(D.30)

Furthermore the stationary distribution and value function are

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

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

where,

\[
\nu = \frac{\gamma - g}{\sigma^2} + \sqrt{\left( \frac{g - \gamma}{\sigma^2} \right)^2 + \frac{r - g}{\sigma^2/2}}.
\]  
(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 $\alpha$, the flow of adopters, $S$, satisfies

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

that is,

\[
S = \frac{\sigma^2}{2} \left( \alpha - \frac{g - \gamma}{\sigma^2} \right)^2 + \left( \frac{g - \gamma}{\sigma^2} \right)^2.
\]  
(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 $\nu$ 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 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 $\alpha$ 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 $\alpha \approx 1.62$ 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.$^{36}$

The minimum growth rate of around 1% occurs at the maximum $\alpha$, 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

$^{36}$In determining whether these $\alpha$ are empirically plausible, consider the crude adjustment to tail indices based on for markups discussed in (D.47).
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.37

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.

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, $\zeta > 0$. Future utility is discounted at a rate $\hat{r} > 0$. 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 perfectly diversified portfolio of firms.

A competitive final goods sector produces a good with elasticity of substitution $\beta$ 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, $\rho/(\rho - 1)$, with period profits $\Pi(t, Z) \propto Z^{\rho - 1}$.

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 $r = \delta + \hat{r} + \zeta g$.

Productivity of operating firms evolves according to the exogenous GBM in Appendix D. Adopting a new technology requires hiring $\zeta$ 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 $\delta N$ 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.38

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, $\theta W(t) = E[V(t, Z)]$, 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 $(\theta - \zeta) W(t) = V(t, M(t))$.

37See Perla, Tonetti, and Waugh (2015) for a related derivation of the deterministic equilibrium with richer cost functions and international trade.

38While 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).
The labor market clearing condition distributes the inelastic supply of labor between variable production, adoption, and entry. Define the following,

\[ \tilde{\pi} \equiv \frac{1 + \alpha - \rho}{\alpha(\rho - 1)} \left( \frac{1}{N} - \zeta \alpha \left( g - \gamma - \alpha \frac{\sigma^2}{2} \right) - \delta \theta \right) \]  

\[ \nu \equiv \frac{4\pi \alpha - 2(\alpha - 1)\rho (\tilde{\pi} + \alpha \zeta (g - \gamma)) - 2\tilde{\pi} + (\alpha - 1)\alpha \zeta (\rho - 1)^2 \sigma^2 + 2(\alpha - 1)\alpha \zeta (-\gamma + 2g - r)}{(\alpha - 1)\zeta (2(-\gamma \rho + \gamma + g(\rho - 2) + r) - (\rho - 1)^2 \sigma^2) - 2\tilde{\pi}} \]  

\[ a = \frac{\tilde{\pi}}{r - g - (\rho - 1)(\gamma - g + (\rho - 1)\sigma^2/2)} \]

Proposition 8 (Monopolistic Competition with Free Entry on a BGP). There exist a continuum of equilibria parameterized by \( \alpha \) where,

\[ F(z) = 1 - e^{-\alpha z}. \]  

The tail parameter of the underlying productivity distribution is \( \alpha \), while the tail parameter of the profit and firm size distributions, is given by

\[ \hat{\alpha} \equiv (\rho - 1)\alpha. \]  

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

\[ 0 = -g + \frac{2\tilde{\pi}(\alpha - 1)(\rho - 1)\sigma^2 \nu}{(\alpha - 1)\zeta (2(-\gamma \rho + \gamma + g(\rho - 2) + r) - (\rho - 1)^2 \sigma^2) - 2\tilde{\pi}} + \alpha \sigma^2 + \gamma + \sqrt{(g - \gamma)^2 + 2\sigma^2(r - g)} \]

\[ \theta - \zeta = \frac{\tilde{\pi}(\nu + \rho - 1)}{\nu(-\gamma \rho + \gamma + g(\rho - 2) + r) - \nu(\rho - 1)^2 \sigma^2/2} \]  

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

\[ v(z) = ae^{(\rho - 1)z} + \frac{(\rho - 1)a}{\nu}e^{-\nu z} \]  

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. \( \sigma = \zeta = \gamma = \delta = 0 \)). Furthermore, choose \( \rho = 2 \) to simplify the algebra. With this, given an \( \alpha \), the growth rate and number of varieties is,

\[ g = \frac{r(\theta / \zeta - \alpha)}{(\alpha - 1)^2} \]  

\[ N = \frac{(\alpha - 1)^2}{r\alpha \zeta (1 - 2\alpha + \alpha \theta / \zeta)} \]

The key result is that growth rates are determined by the ratio \( \theta / \zeta \). 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 \( \zeta \), an increasing \( \theta \) 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, $\rho$, 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 $\tilde{\rho} \equiv (\rho - 1)/\rho > 0$. Given an estimated $\hat{\alpha}$ from the firm size, profits, or revenue empirical distribution, the underlying tail index of the productivity distribution is

$$\alpha = \frac{1 - \tilde{\rho}}{\tilde{\rho}} \hat{\alpha}. \tag{D.47}$$

This adjustment might explain some of the differences between the calibrated $\alpha$ 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.\footnote{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.}