NBER WORKING PAPER SERIES

CATCH-UP AND FALL-BACK THROUGH INNOVATION AND IMITATION

Jess Benhabib
Jesse Perla
Christopher Tonetti

Working Paper 18091
http://www.nber.org/papers/w18091

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2012

We would like to thank Boyan Jovanovic and Tom Sargent for useful feedback. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Catch-up and Fall-back through Innovation and Imitation
Jess Benhabib, Jesse Perla, and Christopher Tonetti
NBER Working Paper No. 18091
May 2012
JEL No. O14,O30,O31,O33,O40

ABSTRACT

Will fast growing emerging economies sustain rapid growth rates until they "catch-up" to the technology frontier? Are there incentives for some developed countries to free-ride off of innovators and optimally "fallback" relative to the frontier? This paper models agents growing as a result of investments in innovation and imitation. Imitation facilitates technology diffusion, with the productivity of imitation modeled by a catch-up function that increases with distance to the frontier. The resulting equilibrium is an endogenous segmentation between innovators and imitators, where imitating agents optimally choose to "catch-up" or "fall-back" to a productivity ratio below the frontier.

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

Jesse Perla
New York University
19 W 4th St, 6th Floor
NY, NY 10012
jesse.perla@nyu.edu

Christopher Tonetti
Department of Economics
New York University
19 W. 4th Street 6 Floor
New York, NY 10012-1119
chris.tonetti@nyu.edu

1 Introduction

During the last decade, a number of newly industrialized countries, most prominently Russia, India, and China, have exhibited rapid sustained growth, often at rates of 5 to 10 percent annually. While part of this growth is attributable to factor accumulation, innovation, and rising revenues from natural resources, the unusually fast pace is also due to the diffusion of technology from more advanced countries that has been termed "catch-up." Will these countries continue to grow quickly until they catch-up to the technology frontier and then grow through innovation, or will growth slow down while these countries are significantly behind the frontier where they can better exploit externalities from technology diffusion?

Emerging economies are not the only countries that have experienced catch-up. Many European countries grew faster than the United States during the 1970s and 1980s. However, early in the 1990s their GDP per capita relative to the U.S. began to stabilize at a value significantly below unity, perhaps providing evidence for the future path of these rapidly growing low income countries.

Additionally, over the previous two decades a number of upper-middle income developed countries like Portugal and Italy have been growing at rates slower than innovating countries at the technology frontier like the United States. Are they falling behind despite their best efforts to keep up with the leader, or are there free-rider incentives to let others innovate and push out the frontier that might make falling behind optimal? If the intensity of technology diffusion is endogenous, might some middle-income countries choose to "fall-back" in their relative technology level in order to consume more and grow through the diffusion of new innovations from the frontier?[1]

Empirical explorations of the early Nelson and Phelps (1966) model based on the theory of technological catch-up (Gerschenkron (1962), Lucas (2009)) suggest that the rate of productivity growth depends on the distance to the technology frontier, modulated by factors like human capital (Benhabib and Spiegel (1994, 2005)) or political constraints and inertia (Parente and Prescott (1994)). Despite presenting compelling empirical findings on the role of technology diffusion in explaining growth, these early papers do not explicitly model economic variables that affect the rate of diffusion as subject to economic choice. Levels of human capital or the political institutions, even if measured, are taken as exogenous. Some of the later literature, for example Barro and Sala-i-Martin (1997), endogenize the investment expenditures in "imitation," but typically differentiate between leaders, with exogenously specified low innovation costs, and followers, with high innovation but low imitation costs. A related approach is to endogenize the decision to undertake research expenditures that increase the probability of adopting the leader's technology. For example, in Aghion, Howitt, and Mayer-Foulkes (2005) countries differ in their profitability as well as their research costs: some find it profitable to engage in investments to optimize their probability of technology adoption, while others with high costs or low profitability may choose zero investments in adoption and keep falling behind.[2,3]

The main contribution of our paper is a stylized model in which agents optimally choose the amount to invest in improving growth through innovation as well as through technology adoption. We examine how the efficiency of technology diffusion determines whether agents will ultimately catch-up to the frontier, catch-

[1]Chapter 5 of the World Bank Golden Growth publication (Gill and Raiser (2012)) focuses on "Europe's innovation deficit," discussing the American "innovation machine" and the European "convergence machine." Some of the main questions in this publication are "whether Europe has fundamental flaws in its economic environment that make its innovation deficit a fact of life" and "what should European governments do to increase innovation?" From the perspective of this theory, imitation and fall-back may be an optimal outcome. This force is discussed in the context of the "benefits to backwardness" in Klenow and Rodriguez-Clare (2005).

[2]For an exposition see Aghion and Howitt (2009), pp. 152–158.

[3]Acemoglu, Aghion, and Zilibotti (2006) introduce a model where high skill entrepreneurs can enhance productivity growth through innovation but not through technology adoption. When the distance to the frontier is large, firms may retain a low-skill entrepreneur but replace the entrepreneur at some cost when the distance closes and innovation becomes important. This introduces productivity growth that depends on the distance to the frontier. In our model, we allow the agents to optimally choose any level of investment in innovation as well as in technology adoption, rather than having firms make discrete decisions on the retention of low skill entrepreneurs. In both models, productivity growth and optimal actions depend on the distance to the frontier.

up part of the way to the frontier and then slow down, fall-back and continue to grow through technology diffusion, or conduct autarkic innovation with no contribution from technology diffusion.

Agents in our model are identical except for their initial productivity levels: they do not differ in their costs of investment, production or adoption technologies, or preferences. All agents can affect the rate at which their productivity deterministically grows with expenditures that facilitate technology adoption or innovation at some consumption cost, and they do so optimally. The optimal choice for an agent is a portfolio of imitation and innovation investments, both of which affect the rate at which their productivity grows depending on the relative return of each investment. Innovation is a simple technology with constant marginal productivity that does not benefit from externalities. As imitation facilitates technology diffusion from an agent at the frontier, the productivity of imitation will depend on characteristics of both the agent itself and the agent it is imitating.[4] This relationship is modeled with a catch-up function that determines the amount of technology diffusion that occurs for a marginal investment in imitation.

This catch-up function for technology diffusion nests some the standard diffusion models, including the logistic, the Gompertz, and the Nelson-Phelps (confined exponential). In this framework technology adoption is gradual. The *rate of growth* of productivity due to imitation depends on investments that facilitate technology diffusion as well as on a measure of the ease of adopting existing superior technologies—the distance to the technology frontier.[5,6]

Engaging in imitation has a trade-off: the agent gains the advantage of higher productivity as they exploit the diffusion externality, but if they move closer to the frontier, diffusion becomes less efficient and more investment is necessary to maintain their productivity relative to the frontier. These competing forces can eventually balance, where an agent chooses an equilibrium productivity that is permanently below the technology frontier. Moreover, if agents have high relative productivity but the returns to innovation are insufficient, they may choose to endogenously fall-back to a lower relative productivity level to gain the advantage of more efficient technology diffusion. Very low productivity agents, on the other hand, may invest to increase their productivity until the trade-offs are in balance.

We find that there is a threshold such that all agents with a ratio of their productivity to the technology frontier above this threshold choose to invest only in innovation. All agents whose ratio of productivity to the frontier is below the threshold entirely focus investment expenditures on facilitating technology diffusion. The threshold will depend on the return of innovation relative to imitation for the agents, where the catch-up function determines the rate of technology flow for a unit investment in imitation.

We define a balanced growth path (BGP) as a path where the productivities of all agents grow at the same rate. Starting from an arbitrary initial distribution, we study the long run distribution of productivities on the BGP, its dependence on fundamental cost and efficiency parameters, and the uniqueness and stability of this equilibrium. We explore how the productivity distribution becomes segmented between those who pursue imitation and those who specialize in innovation. All innovators grow at the same rate, and therefore the ratio

[4] The literature on technology and knowledge spillovers also explores a similar trade-off between an agent's own investment in R&D and exploiting externalities by investing in imitating the innovations of more productive firms. In Eeckhout and Jovanovic (2002), the total factor productivity of an agent depends on both its own capital and that of all better firms in the economy. Kortum (1997) instead has agents investing in a stochastic research technology, where the efficiency of new technologies depends on a spillover from the entire stock of existing research.

[5] Technologies to produce goods could consist of various processes, not all of which are instantly adopted. For example, production may start with an assembly plant and the manufacturing of parts. Associated processes may be gradually adopted as know-how develops and accumulates. In 1977, on visiting a Samsung research lab in Korea, Ira Magaziner wrote: "[It] reminded me of a dilapidated high-school science room. . .They'd gathered color televisions from every major company in the world—RCA, GE, Hitachi—and were using them to design a model of their own." See Magaziner and Patinkin (1990), pp. 24.

[6] Jones (2005) explores idea based models of growth and the role of non-rivalrous inputs in generating endogenous growth. In our model, innovating agents have non-rivalrous productivity but are given no legal method to profit from imitation and face no negative effects from being copied. Moreover, technology diffusion is slow, copying modest improvements of the frontier sequentially as described in footnote 5. This is in contrast to models of directed imitation at lower levels of aggregation, where agents make a rapid jump to the frontier by copying the most productive individual firms. In general, patent systems typically require active monitoring and expensive litigation that are only optimal for cutting-edge products with high rents. Last, our model exhibits no scale effects due to the structure of the catch-up function.

of their productivity to the technology frontier remains unchanged. In the limit, the ratio of every imitator's productivity to the technology frontier converges to a common constant that lies below the threshold. All imitators grow at the same rate as the frontier technology with growth only driven by technology diffusion. Imitators do not necessarily disappear in the limit, nor is there universal convergence to the frontier. However, if investment in productivity growth is significantly more efficient through innovation than through adoption, the threshold ratio above which agents only invest in innovation may be zero. Similarly, if the technology for diffusion is sufficiently inefficient, agents who optimally specialize their expenditures on technology diffusion are nevertheless left progressively behind, so that their productivity ratio to the technology frontier converges towards zero. Last, we find that a uniform increase in the productivity of both innovation and imitation will not distort the proportion of agents conducting innovation, but will change the equilibrium productivity of imitating agents.

An alternative approach to technology diffusion is to consider how agents might adopt the technology of agents spread throughout the productivity distribution, on average adopting technologies much closer to the mean productivity than the frontier. The recent works of Lucas and Moll (2012) and Perla and Tonetti (2011) endogenize technology adoption as an optimal search decision made by firms taking draws from the distribution of existing technologies. They then study the endogenous evolution of the productivity distribution that results from this search process, without giving agents access to an innovation technology. Both Lucas and Moll (2012) and Perla and Tonetti (2011) work within a stochastic search framework where searching agents copy the technology of a random agent in the economy and immediately make a discrete jump to the new productivity level. Imitation here is more in the spirit of Benhabib and Spiegel (1994, 2005). Another approach, taken by Luttmer (2011), studies how imitation by new entrants can drive aggregate growth, with the productivity of incumbents evolving according to an exogenous stochastic process as opposed to a controlled innovation decision. In our framework we are able to focus on the portfolio choice of allocating expenditures between imitation and innovation and on the endogenous evolution of the frontier, allowing the key economic incentives for catch-up and fall-back through innovation and imitation to be clearly illuminated.

2 The Model

Agents are heterogeneous over their productivity level, z, and choose to invest in improving productivity through imitation and innovation. The cost of generating productivity growth through imitation (or technology adoption) is sz and the cost of generating productivity growth through innovation is γz. The parameters $\sigma > 0,\ c > 0$ measure the efficiency of innovation and technology adoption expenditures, r is the discount rate, and utility is logarithmic. Output is given by $Bz,\ B > 0$. The initial distribution of productivities is $\tilde{\Phi}(0, z)$, with min support > 0 and max support $\equiv F(0) > 0$. The distribution of productivities at time t is $\tilde{\Phi}(t, z)$ on $(0, F(t)]$, where $F(t)$ is the frontier. It is often convenient to work in productivities as ratios compared to the frontier: $x(t) \equiv \frac{z(t)}{F(t)}$. In that case, the distribution of x is $\Phi(t, x)$ defined on $(0, 1]$.[7] This distribution fully characterizes the aggregate state of the economy.

Catch-up Technology The efficiency of expenditure on imitation depends on the distance of the imitating agent to the technology frontier. Our catch-up function (or technology adoption function) parametrically specifies how technology adoption expenditures affect productivity growth as a function of distance to the technology frontier. Our specification nests the logistic and Nelson-Phelps technology diffusion processes and is flexible enough to allow for the flow of technology diffusion to decrease, for a fixed investment, as the distance to the frontier increases. Since the catch-up function only depends on distance to the frontier and

[7] A $\tilde{\cdot}$, such as $\tilde{\Phi}(t, z)$, in our notation indicates objects in (t, z) space, while its absence indicates the transformation to (t, x) space.

no other statistics of the productivity distribution, the evolution of each initial productivity relative to the frontier can be studied independently.[8]

All agents choose their expenditures in innovation and in technology adoption optimally, both of which promote the rate of productivity growth. Spending γz on innovation produces growth of z at rate $\gamma\sigma$. For an investment of sz in imitation, the agent produces growth of z at a rate $s\tilde{D}(t,z)$. Given parameters controlling the rate of technology diffusion, c and m, the catch-up function for technology growth at time t is given by

$$\tilde{D}\left(t,z;m\right)=\tfrac{c}{m}\left(1-\left(\tfrac{z(t)}{F(t)}\right)^{m}\right) \tag{1}$$

where expenditures $s\geq 0$ and $m\in(-\infty,\infty)$. Since the catch-up function specifies the marginal return to an investment in imitation it critically affects agents dynamic behavior. Setting $m=1$ gives the familiar logistic diffusion, while $m=-1$ gives the Nelson-Phelps (confined exponential) diffusion.[9] Letting $m\to 0$ yields, after using L'Hopital's rule, the Gompertz diffusion model:

$$\tilde{D}(t,z;0)=-c\ln\left(\tfrac{z(t)}{F(t)}\right)$$

Note that $\tilde{D}\left(t,F(t);m\right)=0$ for all t,m. Hence, the returns to imitation for the frontier agent are always 0.

The General Optimization Problem Our model allows for the endogenous growth of heterogeneous agents: productivity growth depends linearly on optimally chosen expenditures on innovation and adoption, γ and s. The benefits of innovation and diffusion, $\gamma\sigma$ and $s\tilde{D}\left(t,z;m\right)$, and their respective costs, $\gamma(t)z(t)$ and $s(t)z(t)$, are flow variables in continuous time. Consumption is $Bz(t)-s(t)z(t)-\gamma(t)z(t)$. Given an initial idiosyncratic productivity level, $z(0)$, and an initial productivity distribution across agents, $\tilde{\Phi}(0,z)$, all agents solve the same problem:

$$\max_{s(t),\gamma(t),z(t)}\int_{0}^{\infty}\left(\ln\left(Bz-sz-\gamma z\right)\right)e^{-rt}dt \tag{2}$$

$$s.t.\quad \frac{\dot{z}}{z}=\sigma\gamma+\tilde{D}(t,z;m)s \tag{3.a}$$

$$s\geq 0,\;\gamma\geq 0 \tag{3.b}$$

Calculating the current value Hamiltonian for this system with $\tilde{\lambda}$ as the Lagrange multiplier on the law of motion, assuming $m\neq 0$ and plugging in our $\tilde{D}(\cdot)$ function yields

$$\tilde{H}=\ln\left(\left(B-s-\gamma\right)z\right)+\tilde{\lambda}\sigma\gamma z+\tilde{\lambda}zs\frac{c}{m}\left(1-\left(\frac{z}{F(t)}\right)^{m}\right) \tag{4}$$

The first order necessary conditions of this Hamiltonian are augmented with complementarity conditions to ensure $s\geq 0$ and $\gamma\geq 0$. H is concave in s,γ, but the full second order conditions are addressed in Section

[8] Using the frontier (the maximum statistic) to summarize the distribution provides a great deal of analytical tractability. Perla and Tonetti (2011) and Lucas and Moll (2012) use random search models to focus on the feedback relationship between the entire endogenously determined productivity distribution and optimal technology adoption decisions. Alvarez, Buera, and Lucas (2011) combine a similar random search structure with a model of trade to explicitly micro-found technology diffusion across countries. These papers do not feature an innovation technology or an endogenously evolving technology frontier. Dependence only on the frontier can be thought of as directed search with no differential search costs. The effect of directed investment in imitation, with imitation costs increasing as the distance to the imitated grows, is left for future research.

[9] Note that in the logistic case ($m=1$) the increase in productivity $\dot{z}\left(t\right)$ from diffusion, $z(t)\tilde{D}\left(t,z;1\right)=cz\left(t\right)\left(1-\left(\tfrac{z(t)}{F(t)}\right)\right)$, is zero when $z\left(t\right)=0$ or when $z\left(t\right)=F\left(t\right)$, and it is maximized at $z\left(t\right)=F\left(t\right)/2$. By contrast in the Nelson-Phelps case ($m=-1$), $z(t)\tilde{D}\left(t,z;-1\right)=cz\left(t\right)\left(\left(\tfrac{F(t)}{z(t)}\right)-1\right)=c\left(F(t)-z\left(t\right)\right)$, the larger the distance to the frontier, the larger is the diffusion flow.

4. Note that given log utility, it will not be optimal to set $(B - s - \gamma) = 0$.

3 Frontier

Because the imitation policies of all agents not at the frontier depend on the behavior of the frontier, the frontier agent's problem must be solved first. Since improving productivity through imitation is not possible at the frontier, the leader sets $s = 0$. The problem for the leader then simplifies to

$$\tilde{H} = \ln\left((B - \gamma) z\right) + \tilde{\lambda}\sigma\gamma z \tag{5}$$

with

$$\dot{z} = \sigma\gamma z \tag{6}$$

and

$$\dot{\tilde{\lambda}} = -\frac{\partial \tilde{H}}{\partial z} + r\tilde{\lambda} = -z^{-1} + \tilde{\lambda}\left(r - \sigma\gamma\right) \tag{7}$$

Defining $\tilde{\mu} = \tilde{\lambda} z$, we have

$$\dot{\tilde{\mu}} = \dot{\tilde{\lambda}} z + \dot{z}\tilde{\lambda} = -1 + \tilde{\mu}\left(r - \sigma\gamma\right) + \sigma\gamma\tilde{\mu} = -1 + r\tilde{\mu} \tag{8}$$

The solution is

$$\tilde{\mu}(t) = e^{rt}\left(\tilde{\mu}(0) - \frac{1}{r}\right) + \frac{1}{r} \tag{9}$$

and the transversality condition

$$\lim_{t\to\infty} e^{-rt}\tilde{\mu}(t) = 0 \tag{10}$$

immediately requires

$$\tilde{\mu}(t) = \frac{1}{r} \tag{11}$$

Maximizing the Hamiltonian with respect to γ yields

$$\frac{-1}{B - \gamma} + \tilde{\lambda} z\sigma \leq 0 \tag{12}$$

Using equation 11 we get

$$\gamma \geq B - \frac{r}{\sigma} \tag{13}$$

Assumption 1: $\sigma B - r > 0$

From equation 13, under Assumption 1 it is clear that the optimal investment in innovation at the frontier is strictly positive.[10] Assumption 1 ensures that for the frontier agent the discount rate is small enough that the benefits of increasing future productivity through innovation are larger than the benefits of immediately

[10]It is straightforward to show that the maximized Hamiltonian, $\max_z \tilde{H}\left(z, \tilde{\lambda}\right)$ is concave in z, so that this is indeed an optimal solution. See Appendix B for details.

consuming more by not innovating. With Assumption 1 the growth rate at the frontier is a strictly positive constant: $g \equiv \frac{\dot{F}(t)}{F(t)} = \sigma\gamma = \sigma B - r > 0$. Section 4.2 shows that agents initially behind the frontier will chose optimal growth rates such that they remain behind the leader at all times.

4 Followers

For $z(t) < F(t)$, the returns to imitation are strictly positive. Since the productivity of imitation depends on the distance to the frontier, followers take into account that the frontier grows at a constant rate $g = B\sigma - r$. The following results are derived for the $m \neq 0$ case. As is discussed in footnote 14, the solution is continuous for the Gompertz case at $m = 0$.

Before solving the follower's problem, it is useful to transform the problem with a change of variables.

Optimization Problem in x A major advantage of this form of catch-up function is that since it only depends on the ratio of an agent's productivity to the frontier, the optimization problem becomes time invariant given a simple change of variables. Using the ratio of productivity to the frontier

$$x(t) \equiv \frac{z(t)}{F(t)} \tag{14}$$

examination of equation 1 shows that there is no dependence on time in the catch-up function except through x

$$\tilde{D}\left(t, x(t)F(t); m\right) = \tfrac{c}{m}\left(1 - \left(\tfrac{x(t)F(t)}{F(t)}\right)^m\right) = \tfrac{c}{m}\left(1 - x^m\right) \tag{15}$$

$$\equiv D(x; m) \tag{16}$$

The growth of x is

$$\frac{\dot{x}(t)}{x(t)} = \sigma\gamma(t) + D(x(t); m)s(t) - \tfrac{\dot{F}(t)}{F(t)} \tag{17}$$

Since the frontier grows at a constant rate, $\frac{\dot{F}(t)}{F(t)} = g$ as proved in Section 3, the law of motion for x is time homogeneous.

$$\frac{\dot{x}(t)}{x(t)} = \sigma\gamma(t) + D(x(t); m)s(t) - g \tag{18}$$

Transforming the objective function in equation 2 using the change of variables yields

$$\int_0^\infty \left(\ln(x) + \ln(B - s - \gamma)\right) e^{-rt}\mathrm{d}t + \int_0^\infty \ln(F(t))e^{-rt}\mathrm{d}t$$

Given an $F(t)$, which this agent does not control, the last term is a constant. Since this constant is finite, $\dot{F}(t)/F(t) = g = \sigma B - r$, the term does not affect optimal controls. Thus the optimization problem in x space is

$$\max_{s(t),\gamma(t),x(t)} \int_0^\infty (\ln x + \ln(B - s - \gamma))\, e^{-rt}\mathrm{d}t + \frac{g + r\ln(F(0))}{r^2} \tag{19}$$

$$\frac{\dot{x}}{x} = \sigma\gamma + D(x;m)s - g \tag{20.a}$$

$$s \geq 0,\ \gamma \geq 0 \tag{20.b}$$

The current value Hamiltonian ignoring the constant in the objective function and the complementarity conditions is given by

$$H = \ln(x) + \ln(B - s - \gamma) + \lambda x \left(\sigma\gamma + \tfrac{c}{m}(1 - x^m)s - g\right) \tag{21}$$

The Lagrange multiplier on the law of motion in x is λ. As previously mentioned, this problem is now autonomous when compared to equation 4.

Marginal Productivity of Innovation and Imitation Since equation 20.a is linear in both s and γ, the marginal productivity for an investment in either s or γ is constant. That is, an investment in γ has a marginal productivity of innovation σ. Given a fixed value of x, an investment in s has a marginal productivity of imitation $D(x)$. Since $\frac{\partial D(x;m)}{\partial x} < 0$, the marginal productivity of imitation decreases with x. Because the efficiency of technology adoption depends on the distance to the frontier, forward looking agents will consider the negative effect on the marginal productivity of imitation from increasing x.

4.1 First Order Necessary Conditions

The first order necessary conditions can be derived using the Hamiltonian in x space (equation 21).

$$H = \ln(x) + \ln(B - s - \gamma) + \lambda x \left(\sigma\gamma + \tfrac{c}{m}(1 - x^m)s - g\right) \tag{22}$$

The Euler equation is

$$\begin{aligned}
\dot{\lambda} &= -\frac{\partial H}{\partial x} + r\lambda \\
&= -\frac{1}{x} - \lambda\left(\sigma\gamma + sD(x) + sxD'(x) - g\right) + \lambda r && (23)\\
&= -\frac{1}{x} - \lambda\left(\sigma\gamma + \tfrac{c}{m}(1 - x^m)s - g\right) + \lambda c s x^m + \lambda r && (24)\\
&= -\frac{1}{x} - \lambda\frac{\dot{x}}{x} + \lambda c s x^m + \lambda r && (25)
\end{aligned}$$

Note that $D'(x) = -cx^{m-1}$ is embedded in equation 23, so the agent considers the effect of varying x on future catch-up costs.

The first-order conditions for s, including the complementarity condition, are

$$0 = s\frac{\partial H}{\partial s}, \ \frac{\partial H}{\partial s} \leq 0$$

$$\frac{1}{B-s-\gamma} \geq \lambda x D(x) \tag{26}$$

$$\frac{1}{B-s-\gamma} \geq \lambda x \frac{c}{m}\left(1-x^m\right) \tag{27}$$

The first-order conditions for γ, including the complementarity condition, are

$$0 = \gamma\frac{\partial H}{\partial \gamma}, \ \frac{\partial H}{\partial \gamma} \leq 0$$

$$\frac{1}{B-s-\gamma} \geq \lambda x \sigma \tag{28}$$

The transversality condition is

$$\lim_{t\to\infty} e^{-rt}\lambda(t)x(t) = 0 \tag{29}$$

It can be seen from the the first-order conditions in equations 26 and 28 that an agent considers the marginal productivity of innovation or imitation when optimizing its investment portfolio, where λ is the shadow price of x.

To solve for the equilibrium, it is useful to define a new variable: $\mu(t) \equiv \lambda(t)x(t) \implies \dot{\mu}(t) = \dot{\lambda}x + \lambda\dot{x}$. $\mu(t)$ is the agent's valuation of its state $x(t)$ at shadow price $\lambda(t)$. Let $\bar{\cdot}$ denote steady state values (e.g. $\bar{s}$ and $\bar{x}$). The Euler equation (equation 25) can be rewritten as

$$\dot{\lambda}x = -1 - \lambda\dot{x} + \lambda x c s x^m + \lambda x r$$

$$\dot{\mu} = -1 + \mu(r + csx^m) \tag{30}$$

At steady state $\dot{\mu} = 0$, so

$$\bar{\mu} = \frac{1}{r + c\bar{s}\bar{x}^m} \tag{31}$$

4.2 Solution to Followers Problem

The followers problem can be solved by analyzing all cases of the complementarity conditions for s and γ.

4.2.1 Simultaneous Innovation and Imitation

For the case where agents do both innovation and imitation, $s > 0$ and $\gamma > 0$. This can only happen at a unique x^*, which separates regions of innovation and imitation. For both s and γ to be interior, the first order conditions, equations 27 and 28, must be equal.

$$\lambda x^* \sigma = \lambda x^* \frac{c}{m}\left(1 - x^{*m}\right)$$

$$\sigma = D(x^*) \tag{32}$$

$$\sigma = \frac{c}{m}\left(1 - x^{*m}\right)$$

$$x^* \equiv (1 - \tfrac{\sigma m}{c})^{1/m} \tag{33}$$

Hence, an agent is indifferent between innovating and imitating at a time-invariant knife edge productivity ratio, x^*, which is strictly between zero and one for all $m \neq 0$ given the following assumption.

Assumption 2: $\frac{\sigma m}{c} < 1$

$x^* \in (0,1)$ means that there are both followers that prefer to innovate and followers that prefer to imitate. Assumption 2 guarantees this by ensuring that the productivity of innovation is not too large compared to the productivity of imitation and that the productivity of imitation is not too large compared to the productivity of innovation, since these forces are modulated by σ, m, and c. At any x other than x^* both equation 27 and equation 28 cannot hold with equality simultaneously, and therefore both s and γ cannot be simultaneously strictly positive except at x^*.[11]

When making the portfolio choice on the extensive margin, equation 32 shows that the agent only needs to compare the instantaneous marginal productivity of innovation to the marginal productivity of imitation. In particular, the change in the marginal product of imitation as x changes, $D'(x)$, does not affect the agent's decision.

Lemma 1. *Given Assumptions 1 and 2, there exists an interior time-invariant x^* such that:*

$$x^* = \left(1 - \frac{\sigma m}{c}\right)^{1/m} \tag{34}$$

$$x < x^* \implies s(x) \geq 0,\ \gamma(x) = 0 \tag{35}$$

$$x > x^* \implies s(x) = 0,\ \gamma(x) > 0 \tag{36}$$

Proof. Section 4.2.1 has shown that both $s > 0$ and $\gamma > 0$ can only occur at a single x. From the FOCs, for innovation to be preferred over imitation, it must be that $\sigma > \frac{c}{m}(1 - x^m)$, which is satisfied by all $x > x^*$ for both positive and negative m. The imitation region follows directly. The fact that γ is strictly positive in the innovation region is proved in Section 4.2.2.[12] □

When the marginal productivity of innovation is high enough to violate Assumption 2, it is possible for x^* to be $= 0$, and no imitation will take place, as can be seen from equation 34. As long as Assumption 2 is satisfied, no matter how high the marginal productivity of imitation, there is always a region where innovation is optimal.

4.2.2 Innovation Only: $x > x^*$

In the region $x > x^*$ only innovation will occur, as stated in Lemma 1.

[11] If $\frac{m\sigma}{c} > 1$ then $x^* = 0$ and there is no imitation region. In this case, all agents choose to innovate at the same rate, as is shown in Section 4.2.2.

[12] Since x and λ are continuous on an optimal path, at the threshold $x^* = \left(1 - \frac{\sigma m}{c}\right)^{1/m}$, from the left,

$$\frac{\dot{x}}{x} = s\frac{c}{m}(1 - x^m) - g = s\frac{c}{m}\left(\frac{m\sigma}{c}\right) - g = s\sigma - g$$

and from the FOC, coming from the left,

$$\begin{aligned} B - s &= \left(\lambda x \frac{c}{m}(1 - x^m)\right)^{-1} = \left(\frac{1}{r}\frac{c}{m}\left(\frac{m\sigma}{c}\right)\right)^{-1} \\ s &= B - \frac{r}{\sigma} > 0 \\ \frac{\dot{x}}{x} &= s\sigma - g = B\sigma - r - g = 0 \end{aligned}$$

So the growth rate is continuous and equal to the leader's growth rate from the left as $x \to (1 - \frac{m\sigma}{c})^{1/m}$. Note though that s will not be continuous at the threshold.

Using equation 28 and $s = 0$ yields:

$$\frac{1}{B-\gamma} = \lambda x \sigma \tag{37}$$

Assuming that an agent with $s = 0$ will never enter the region $x < x^*$ where $s \geq 0$, as is later verified, equation 11 can be substituted into equation 37 to find

$$\gamma = \frac{B\sigma - r}{\sigma} \tag{38}$$

This is the same level of investment as that of the frontier agent, given in equation 13. Plugging this into the LOM for x

$$\frac{\dot{x}}{x} = \sigma \frac{B\sigma - r}{\sigma} - g = 0 \tag{39}$$

Given the solution for the frontier, $g = B\sigma - r$, this proves that $\frac{\dot{x}}{x} = 0$. All agents in this region grow at the same rate as the frontier by choosing the same level of investment in innovation. Since $\frac{\dot{x}}{x} = 0$ and x^* is time-invariant, it is verified that an agent in $x > x^*$ chooses $\gamma > 0$ and will never enter the $x < x^*$ region.

To verify transversality, note that since an agent will never change regions, $s = 0$ for all t for $x > x^*$. Thus, using $\bar{\mu}$ from equation 31 with $s = 0$ yields

$$\lim_{t\to\infty} e^{-rt}\lambda(t)x(t) = \frac{1}{r}\lim_{t\to\infty} e^{-rt} = 0 \tag{40}$$

4.2.3 Imitation Only: $x < x^*$

In the region $x < x^*$, only imitation will occur, as stated in Lemma 1.

To solve for the policy functions of agents in this region it is useful to perform another change of variables, where the relative distance of the agent to the frontier is distorted by the diffusion parameter m: $q(t) \equiv x(t)^m$. Using the chain rule:

$$\dot{q} = mq\frac{\dot{x}}{x}$$

The law of motion for q is obtained using this change of variables together with the law of motion for x (equation 20.a) when $\gamma = 0$:

$$\frac{\dot{q}}{q} = m\left(\tfrac{c}{m}(1-q)s - g\right) \tag{41}$$

The first order necessary condition in equation 27, with $\gamma = 0$, gives

$$\mu = \frac{m}{c(B-s)\,(1-q)} \tag{42}$$

The law of motion for μ in terms of q is obtained from equation 30

$$\dot{\mu} = -1 + \mu(r + csq) \tag{43}$$

Equations 41, 42, and 43 form a system in the variables μ, q, s. Eliminating μ via substitution will create a 2x2 system. First, differentiating equation 42

$$\dot{\mu} = \frac{m\left((B-s)\dot{q} + (1-q)\dot{s}\right)}{c(1-q)^2(B-s)^2} \tag{44}$$

Substituting equations 42 and 44 into 43, and rearranging for $\dot{s}$

$$\dot{s} = \frac{(B-s)\left((-1+q)(-Bc+mr+cs+cq(B+(-1+m)s)) + m\dot{q}\right)}{m(-1+q)}$$

Substituting for $\dot{q}$ from the law of motion in equation 41, generates a first order ODE in s and q

$$\dot{s} = \frac{(B-s)\left(Bc - mr + q(-2Bc + m(-gm+r) + Bcq) - c(-1+q)^2 s\right)}{m(-1+q)} \tag{45}$$

Equations 41 and 45 are a system of 2 nonlinear first-order ODEs in q and s. Stacking the equations and defining the vector valued function Ψ yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} s(t) \\ q(t) \end{pmatrix} = \Psi(s(t), q(t)) \equiv \begin{pmatrix} \frac{(B-s(t))\left(Bc-mr+q(t)(-2Bc+m(-gm+r)+Bcq(t))-c(-1+q(t))^2 s(t)\right)}{m(-1+q(t))} \\ mq(t)\left(\frac{c}{m}(1-q(t))s(t) - g\right) \end{pmatrix} \tag{46}$$

Stationary solution for $s \geq 0, \gamma = 0$ Define the stationary points of this system as $\bar{s}$ and $\bar{q}$ (or equivalently $\bar{x}$). Evaluating the first order condition in equation 42 at the stationary value $\bar{s}$ yields

$$\bar{s} = B - \frac{1}{c/m\bar{\mu}(1-\bar{q})} \tag{47}$$

Substituting for $\bar{\mu}$ from equation 31 and solving for $\bar{s}$ gives

$$\bar{s} = \frac{B\frac{1}{m}(1-\bar{q}) - \frac{r}{c}}{\frac{1}{m}(1-\bar{q}) + \bar{q}} \tag{48}$$

From Assumption 1 and Lemma 1 it follows that $\bar{s} > 0$. Thus agents choose strictly positive expenditure on imitation at the stationary point in equilibrium.[13]

To determine the stationary solution for $q(t)$, $\bar{q}$, substitute this expression into the stationary law of motion in equation 41

$$\frac{\dot{q}}{q} = 0 = m\left(\frac{c}{m}(1-\bar{q})\frac{B\frac{1}{m}(1-\bar{q}) - \frac{r}{c}}{\frac{1}{m}(1-\bar{q}) + \bar{q}} - g\right) \tag{49}$$

Using the known value of g, this becomes a quadratic equation in $\bar{q}$

$$\frac{\sigma B - r}{B} = \left(\frac{\frac{c}{m}(1-\bar{q}) - \frac{r}{B}}{1 + \bar{q}(m-1)}\right)(1-\bar{q}) \tag{50}$$

This can be further simplified (see Appendix E)

$$0 = \frac{c}{m}(1-\bar{q})^2 - \frac{r}{B}(1-\bar{q}) - \frac{(\sigma B - r)}{B}(1+\bar{q}(m-1)) \tag{51}$$

$$0 = \bar{q}^2 - \left(2 - \left(\frac{m\sigma}{c}\right) + \frac{m^2}{c}\left(\sigma - \frac{r}{B}\right)\right)\bar{q} + \left(1 - \frac{m\sigma}{c}\right) \tag{52}$$

The roots of this quadratic are

$$\bar{q}_{1,2} = \tfrac{1}{2}\left(\left(2 - \tfrac{m\sigma}{c} + \tfrac{m^2}{c}\left(\sigma - \tfrac{r}{B}\right)\right) \pm \left(\left(2 - \tfrac{m\sigma}{c} + \tfrac{m^2}{c}\left(\sigma - \tfrac{r}{B}\right)\right)^2 - 4\left(1 - \tfrac{m\sigma}{c}\right)\right)^{1/2}\right) \tag{53}$$

As is proven in Appendix A.1, the unique interior stationary solution, $\bar{x} \in (0, x^*)$, is

$$\bar{x} = \left[\tfrac{1}{2}\left(\left(2 - \tfrac{m\sigma}{c} + \tfrac{m^2}{c}\left(\sigma - \tfrac{r}{B}\right)\right) - \operatorname{sign}(m)\left(\left(2 - \tfrac{m\sigma}{c} + \tfrac{m^2}{c}\left(\sigma - \tfrac{r}{B}\right)\right)^2 - 4\left(1 - \tfrac{m\sigma}{c}\right)\right)^{1/2}\right)\right]^{1/m} \tag{54}$$

In Proposition 1 and Proposition 2 we show that the unique stationary point $(\bar{x}, \bar{s})$ is stable. Therefore paths that converge to $(\bar{x}, \bar{s})$ automatically satisfy the transversality condition. Substituting these values into equation 31 yields

$$\lim_{t\to\infty} e^{-rt}\lambda(t)x(t) = \lim_{t\to\infty} e^{-rt}\bar{\lambda}\bar{x} = \frac{1}{r + c\bar{s}\bar{x}^m}\lim_{t\to\infty} e^{-rt} = 0 \tag{55}$$

It remains to show that the agent's optimization problem satisfies the second order sufficiency conditions. It is sufficient to prove the concavity of the maximized Hamiltonian with respect to x (see Seierstad and Sydsaeter (1977) Theorems 3 and 10). The following parameter restrictions are sufficient to satisfy the second order conditions. See Appendix B for details.[14]

Assumption 3: $m \geq -1$ and $\frac{c}{\sigma} < m + 2$

Note that both conditions in Assumption 3 restrict the efficiency of technology diffusion. Concavity may be lost as technology diffusion becomes too easy, with high c or low m.

Summary Lemma 1 summarizes the threshold productivity ratio, x^*, that separates agents into innovators and imitators. As proven in Section 4.2.2, for $x > x^*$ all agents choose to innovate at the same rate and grow at the same rate. As proven in Section 4.2.3, for $x < x^*$ all agents exclusively imitate, and stability of the system $\Psi(s(t), q(t))$ implies that there exists a unique steady state productivity ratio, $\bar{x}$, to which all imitating agents converge. Agents below $\bar{x}$ invest in imitation to take advantage of the high efficiency of technology diffusion and catch up to $\bar{x}$. Agents below x^* but above $\bar{x}$ endogenously choose to fall back to $\bar{x}$ to reach the optimal trade-off between consumption and investment in imitation, dictated by the strength

[13] By Assumption 1, $\sigma B > r > 0$, so it is sufficient to show $B\frac{1}{m}(1-\bar{q}) - \frac{\sigma B}{c} > 0$. By Lemma 1, in this region $\frac{1}{m}(1-\bar{q}) > \frac{\sigma}{c}$.

[14] The solution for the $m = 0$ case needs to be solved with the Hamiltonian directly in terms of $D(x;0) = -c\ln(x)$ as discussed in Section 2. Using similar methods to those in Section 4.2, it can be shown that the stable solution is:

$$x^* = \exp\left(-\tfrac{\sigma}{c}\right), \; \bar{x} = \exp\left(-\tfrac{\sigma}{2c} + \tfrac{\sqrt{-4cr+4Bc\sigma+B\sigma^2}}{2c\sqrt{B}}\right), \; \bar{s} = \tfrac{-B\sigma+\sqrt{B}\sqrt{-4cr+4Bc\sigma+B\sigma^2}}{2c}$$

Finally, it can be shown that this solution is the same as the limit of the $m \neq 0$ case, ensuring that there is no discontinuity at $m = 0$.

of the incentive to adopt technology that is governed by the catch-up function. Imitators do not disappear in the limit, nor is there universal convergence to the frontier.[15]

5 The Dynamic Solution

Proposition 1 states the main result of the paper: the existence, uniqueness, and stability of the stationary solution.

Proposition 1. *Let Assumptions 1, 2, and 3 hold, then for arbitrary $\Phi(0,x)$*
***A**: There exists a threshold ratio $x^* = \left(1 - \frac{m\sigma}{c}\right)^{1/m}$ such that all agents with $x > x^*$, including the leader, do only research. They set $s = 0$, $\gamma = B - \frac{r}{\sigma} > 0$, and grow at the rate $\sigma B - r$. For $x > x^*$, the distribution $\Phi(t,x) = \Phi(0,x)$ for all t.*

***B**: All agents with initial conditions $x < x^*$ invest only in technology adoption with $s(t) \geq 0$ and $\gamma(t) = 0$. There exists a strictly positive BGP productivity ratio $0 < \bar{x} < x^*$ such that all initial productivity ratios $x < x^*$ converge to $\bar{x}$.*

Proof. See the cases in Section 4 and the conditions in Appendix A. □

If we denote the cases where $m > 0$ as logistic diffusions and the cases where $m < 0$ as Nelson-Phelps diffusions, Figure 1 illustrates system dynamics and Proposition 1 for several values of σ and $D(x;m)$. The stylized arrows show how convergence might look towards steady state values for x. As can be seen in the lower part of the graph, when the marginal productivity of innovation is greater than the marginal productivity of imitation, the agent chooses to invest entirely in γ.[16]

Proposition 2 characterizes the asymptotic dynamics of the distribution of productivities.

Proposition 2. *A balanced growth path equilibrium is a growth rate $g = \sigma B - r$, a steady state imitator productivity ratio $\bar{x}$, an innovator threshold x^*, and an asymptotic distribution $\Phi(\infty, x)$ such that*

1. *The growth rate of all agents is $\frac{\dot{z}}{z} = g = \sigma B - r$, so $\frac{\dot{x}}{x} = 0$ for all x.*

2. *For arbitrary $\Phi(0,x)$,*

$$\Phi(\infty, x) = \begin{cases} 0 & \textit{for} \quad 0 \leq x < \bar{x} \\ \Phi(0, x^*) & \textit{for} \quad \bar{x} \leq x < x^* \\ \Phi(0,x) & \textit{for} \quad x^* \leq x \leq 1 \end{cases} \tag{56}$$

Proof. See Proposition 1. □

Any initial distribution $\Phi(0,x)$ will converge to this unique balanced growth path. In particular, all initial distributions that agree over support $x > x^*$ will converge to the same BGP, since all agents initially below x^* converge to $\bar{x}$.

This proposition is illustrated in Figure 2 for an arbitrary initial distribution $\Phi(0,x)$ and the unique BGP distribution $\Phi(\infty,x)$. The dynamics and asymptotic distribution are independent of the initial frontier

[15] There is no mixing in the evolution technology, in that the relative productivity rankings of agents is constant for all time. Additionally, innovators never become imitators and imitators never become innovators. This is because the model is deterministic and investments gradually improve growth rates as opposed to facilitating discrete advances in productivity levels. Shocks to productivity would introduce leapfrogging and switches between innovation and imitation.

[16] For some numerical examples that fulfill Assumptions 1, 2, and 3:

Nelson-Phelps ($m = -1$)**:** $x^* = 0.5263$ and $\bar{x} = 0.3148$, when $\sigma = 1; c = .9; B = 2; r = 0.05$

Gompertz ($m = 0$)**:** $x^* = 0.4066$ and $\bar{x} = 0.2258$, when $\sigma = .9; c = 1; B = 2; r = 0.05$

Logistic ($m = 1$)**:** $x^* = 0.1$ and $\bar{x} = 0.0520$, when $\sigma = .9; c = 1; B = 2; r = 0.05$

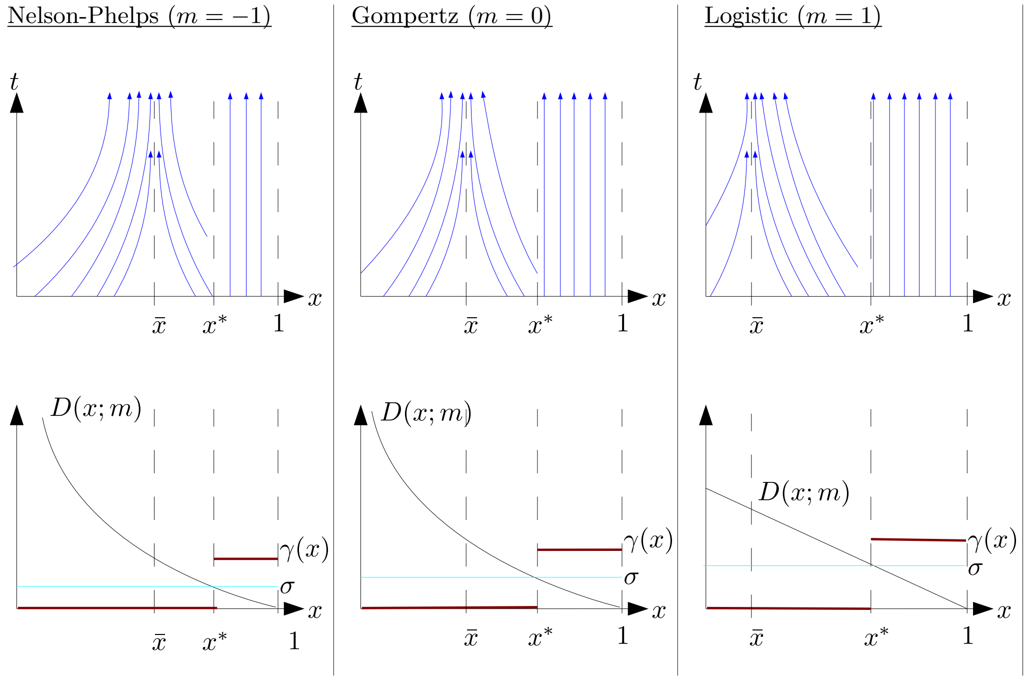

Figure 1: Equilibrium for $D(x;m)$ and σ

$F(0)$. For any initial distribution, the CDF of the upper tail does not change, while a mass of $\Phi(0, x^*)$ will accumulate at the equilibrium point $\bar{x}$.

One of the main findings of Benhabib and Spiegel (2005) was that in the logistic case, if their distance to the frontier is too large, "the follower will not be able to keep up, growth rates will diverge, and the income ratio of the follower to the leader will go to zero." This is because there exists a range of productivities such that the flow of technology diffusion for a given investment in imitation is decreasing in the distance to the frontier. Once the imitation expenditure is endogenized this finding is reversed, as agents invest in technology diffusion such that they converge to a positive BGP productivity ratio, $\bar{x}$, and grow at the same rate as the leader.

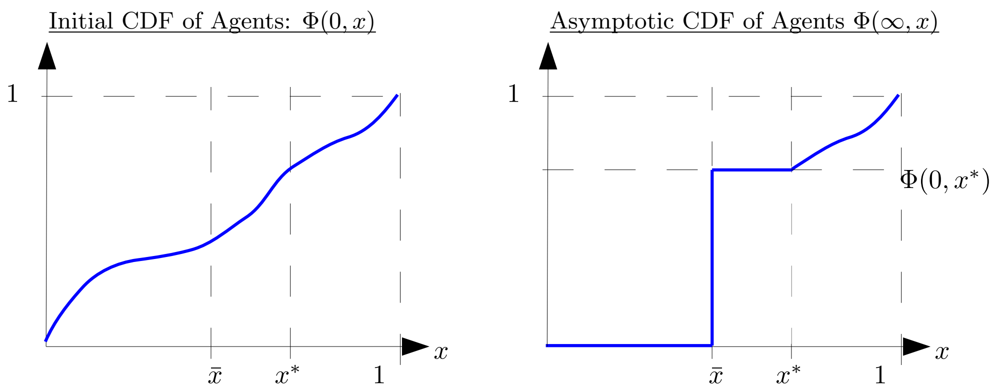

Figure 2: Balanced Growth Path Distribution

5.1 Catch-up, Fall-back, and the Shadow Value of $x(t)$

The Euler equation characterizes the dynamic trade-offs an agent faces. Joint analysis of the Euler equation and the shadow price and shadow value of x provides intuition for the determinants of equilibrium and the forces that lead to the BGP.

From the Euler equation in 23,

$$r = \frac{\dot{\lambda}}{\lambda} + \frac{1}{x} + (\sigma\gamma + sD(x) - g) + sxD'(x) \tag{57}$$

As usual, the Euler equation states that at each point in time the total benefit of a marginal unit of x, represented by sum of the terms on the right, must equal the discount rate. The term $\frac{\dot{\lambda}}{\lambda}$ is the appreciation in the shadow price of x while $\frac{1}{x}$ is marginal utility of x. The terms $(\sigma\gamma + sD(x) - g)$ represent the marginal product of x, both through innovation and through technology diffusion. The final term, $sxD'(x)$, tempers the marginal product of x because a higher x reduces the distance to the frontier and therefore the productivity of future imitation. To see this, note that in our model $D'(x) < 0$. The agents optimally choose s to determine the rate at which they catch-up, or alternatively fall-back. Hence the strength of catch-up or fall-back depends on the magnitude of $sxD'(x)$ with $D'(x)$ exerting a negative force on catch-up or, equivalently, a positive force for fall-back. If however the agent optimally sets $s = 0$, then $D'(x)$ has no direct effect on the change.

Recall $\mu(t)$ can be interpreted as the agent's valuation of its state $x(t)$ at shadow prices $\lambda(t)$. The Euler equation at steady state can be arranged as

$$\bar{\mu} = \frac{1}{r - \bar{s}\bar{x}D'(\bar{x})} = \frac{1}{r + c\bar{s}\bar{x}^m} \leq \frac{1}{r} \tag{58}$$

$$\bar{\lambda} = \frac{1}{\bar{x}}\left(\frac{1}{r - \bar{s}\bar{x}D'(\bar{x})}\right) \tag{59}$$

Equations 58 and 59 show the role of $\bar{s}\bar{x}D'(\bar{x})$ on imitation and innovation in equilibrium. In a region of innovation where $\bar{s} = 0$ the discounted value of the utility flow for a marginal unit of $\bar{x}$ per period is the marginal utility $\frac{1}{\bar{x}}$ discounted by r. In fact, in the innovation region where $s = 0$, logarithmic utility implies that everywhere along the optimal path the marginal utility of an additional unit of x is $\frac{1}{x}$ while its marginal product is constant. This yields the value of the stock $x(t)$ evaluated at the shadow price $\lambda(t)$ as $\mu(t) = \lambda(t)x(t) = \frac{1}{r}$, and $\lambda(t) = \frac{1}{r}\frac{1}{x(t)}$. In the imitation region where $s > 0$ the marginal product of x is no longer constant because, as discussed earlier, $D(x)$ is not constant and $D'(x) < 0$. An increase in the BGP $\bar{x}$ brings the agent closer to the frontier and diminishes the marginal productivity of imitation. This is reflected by the term $\bar{s}\bar{x}D'(\bar{x}) < 0$ which augments the effective discount rate that must be applied to the flow of utility. At the BGP productivity ratio in the imitation region the effective discount rate is $r - \bar{s}\bar{x}D'(\bar{x}) > r$ (see equation 58). This captures the negative impact of higher productivity on the endogenously chosen flow of diffusion. This is the active force in the region of imitation below the innovator threshold x^* that causes agents close to the threshold to choose to fall-back to $\bar{x}$. These agents let their relative productivity slip in order to benefit from a higher flow of technology diffusion.

6 Comparative Dynamics

Propositions 1 and 2 document that the unique BGP can be characterized by g, x^*, and $\bar{x}$, but how do these equilibrium values depend on fundamental parameters? This section studies how the balanced growth path equilibrium changes as parameters modulating the productivity of investment—σ, m, and c—are varied.

6.1 Comparative Dynamics for Hicks-Neutral Technical Change

The ratio $\theta \equiv \frac{\sigma}{c}$ is a measure of the relative productivity of innovation to imitation and is an important determinant of the equilibrium. As can be seen by rewriting equation 34,

$$x^* = (1 - m\theta)^{1/m} \tag{60}$$

the threshold for imitation versus innovation depends on the ratio θ rather than on the levels of σ and c independently. This is because x^* is determined by the ratio of instantaneous marginal productivities that are linear in σ and c, as captured in equation 32. This is analogous to the optimal capital to labor ratio in a neoclassical growth model, where a Hicks-neutral technology shock does not alter the optimal relative expenditures. Here, changes in σ and c that keep θ constant may act like a Hicks-neutral technical change, since they keep this measure of relative productivity constant.

However, equation 54 shows the lack of neutrality on $\bar{x}$.

$$\bar{x} = \left[\frac{1}{2}\left(2 + (m^2 - m)\theta - m^2\tfrac{r}{cB} - \operatorname{sign}(m)\left(\left(2 + (m^2 - m)\theta - m^2\tfrac{r}{cB}\right)^2 - 4\left(1 - m\theta\right)\right)^{1/2}\right)\right]^{1/m} \tag{61}$$

The direct dependence of $\bar{x}$ on c comes from the fall-back incentive. Since $\frac{\partial^2 D(x;c)}{\partial x \partial c} < 0$, as c increases, the strength of the fall-back incentive, $D'(x;c)$, increases. Hence, a seemingly Hicks-neutral increase in the productivity of growth technologies can change the equilibrium outcome. To see this, fix the ratio $\bar{\theta} = \frac{\bar{\sigma}}{\bar{c}}$ and multiply these by a total factor productivity term, A: $\sigma = A\bar{\sigma}$ and $c = A\bar{c}$.

Proposition 3. *An increase in the TFP of growth technologies*

1. *Does not alter the equilibrium ratio of innovators:* $\frac{\mathrm{d}x^*}{\mathrm{d}A} = 0$
2. *Decreases the equilibrium BGP productivity ratio of imitators:* $\frac{\mathrm{d}\bar{x}}{\mathrm{d}A} < 0$
3. *Increases the equilibrium expenditure on imitation:* $\frac{\mathrm{d}\bar{s}}{\mathrm{d}A} > 0$

Proof. $\frac{\mathrm{d}x^*}{\mathrm{d}A} = 0$ follows directly from the definition of x^*. For $\frac{\mathrm{d}\bar{x}}{\mathrm{d}A} < 0$ and $\frac{\mathrm{d}\bar{s}}{\mathrm{d}A} > 0$ see Appendix C. ☐

A seemingly Hicks-neutral change in the technology level of the economy has no effect on x^*, but by increasing the strength of the fall-back incentive the change is not neutral in its effect on $\bar{x}$. Even though the growth technology has improved, the growth rate $B\bar{\sigma}A - r$ increases with A, so imitators have to increase investment in imitation, $\bar{s}$, to keep up with the higher growth rate. The optimal response is to increase investment some to keep up, but, in response to a better imitation technology, to fall back a bit and benefit from the more efficient flow.

To see how the economy responds to ever increasing TFP in growth technologies take the limit of $A \to \infty$. Clearly x^* is unchanged. For $\bar{x}$, the terms with a c in equation 61 go to zero. Hence,

$$\lim_{A\to\infty} \bar{x}(A) = \left[\frac{1}{2}\left(2 + (m^2 - m)\bar{\theta} - \operatorname{sign}(m)\left(\left(2 + (m^2 - m)\bar{\theta}\right)^2 - 4\left(1 - m\bar{\theta}\right)\right)^{1/2}\right)\right]^{1/m} > 0$$

Thus $\bar{x}$ decreases as TFP grows, but stays strictly above 0. This result is in contrast to the comparative dynamics of taking c to ∞, where both x^* and $\bar{x}$ converge to 1, as discussed in Section 6.4.

Figure 3 provides a numerical simulation of these results for both $m = .5$ and $m = -.5$, demonstrating $\bar{x}(A)$ is decreasing in A. In order to better compare the two m values, the figure is normalized by $A = 1$.[17]

[17] The parameters used in this simulation are $\bar{\sigma} = 1.25, \bar{c} = 1, B = 2$ and $r = 0.05$. Using these parameters, Figure 3 plots $\bar{x}(A)$ for $m = -.5$ and $m = .5$. When $m = 0.5$, $\frac{\bar{x}(\infty)}{\bar{x}(1)} = 0.989$. When $m = -.5$, $\frac{\bar{x}(\infty)}{\bar{x}(1)} = 0.991$. For visual clarity, the plotted axes do not intersect at the origin.

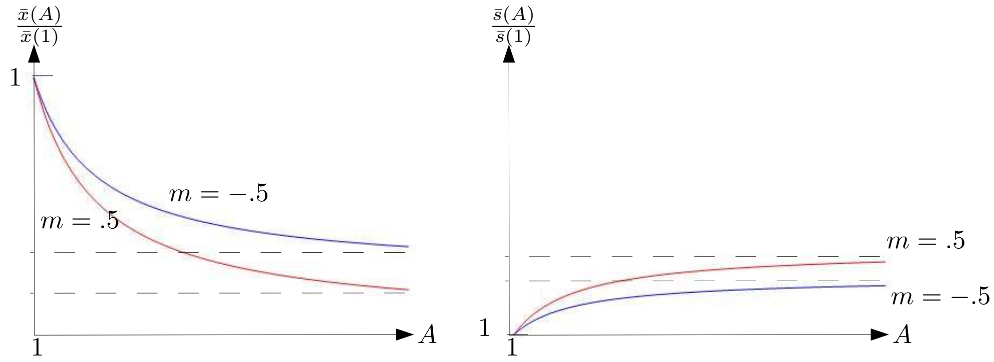

Figure 3: Equilibrium with Hicks-Neutral Technical Change

6.2 Comparative Dynamics of m

m is a key parameter that affects the efficiency of technology diffusion and distorts imitation incentives depending on the distance to the frontier.

As discussed in Section 4.2.1, the decision of agents to imitate or innovate, summarized by x^*, depends only on the instantaneous marginal productivity of innovation and imitation (i.e., $D(x;m)$ vs. σ). A change in m alters the incentives for agents to choose to imitate, as an increase in m decreases the efficiency of technology diffusion. Since $\frac{\partial D(x;m)}{\partial m} < 0$, as m increases the marginal productivity of imitation decreases.[18] As imitation becomes less productive, or equivalently as technology diffusion becomes less efficient, more agents will choose to be innovators and there will be less imitators. This is captured in Proposition 4.

Proposition 4. *As curvature in the catch-up function changes to make imitation less efficient, the innovator threshold decreases. That is* $\frac{\mathrm{d}x^*}{\mathrm{d}m} < 0$.

Proof. Taking the derivative of x^* from equation 34 with respect to m

$$\frac{\mathrm{d}x^*}{\mathrm{d}m} = \frac{\left(1 - \frac{m\sigma}{c}\right)^{\frac{1}{m}} \left(m\sigma + (c - m\sigma)\log\left(1 - \frac{m\sigma}{c}\right)\right)}{m^2(-c + m\sigma)}$$

First, we analyze the case of $m \neq 0$. Assumption 2 ensures that $1 - \frac{\sigma m}{c} > 0$, making the denominator negative and the first term of the numerator positive. It remains to show that the second term in the numerator is positive. That is, we need to show

$$m\sigma + (c - m\sigma)\ln\left(1 - \frac{m\sigma}{c}\right) = c\left[\frac{m\sigma}{c} + (1 - \frac{m\sigma}{c})\ln\left(1 - \frac{m\sigma}{c}\right)\right] > 0$$

This term as a function of $\frac{m\sigma}{c}$ achieves a unique global minimum of 0 at $\frac{\sigma m}{c} = 0$.[19] Since Assumption 2 restricts $\frac{\sigma m}{c} > 0$, this term is strictly positive, and thus $\frac{\mathrm{d}x^*}{\mathrm{d}m} < 0$.[20] $\square$

Contrary to x^*, $\bar{x}$ captures the optimal intensity of investment in imitation. An increase in m decreases the instantaneous marginal productivity of imitation, $D(x;m)$, as well as the fall-back incentive, $D'(x;m)$, which will cause agents to adjust their optimal $\bar{x}$ and $\bar{s}$ levels.

[18] To see this, first note, using a first order Taylor expansion of the convex function x^m around $m = 0$, that $1 - x^m > -mx^m \ln x$. Then using this inequality, $\frac{d\left(\frac{c}{m}(1-x^m)\right)}{dm} = \frac{c}{m^2}\left(x^m - 1 - mx^m \ln x\right) < \frac{c}{m^2}\left(x^m - 1 + (1 - x^m)\right) = 0$.

[19] The first derivative of the term is $\ln(1 - \frac{\sigma m}{c})$ and the second derivative is positive: $\frac{1}{1 - \frac{\sigma m}{c}} > 0$ for $1 - \frac{\sigma m}{c} > 0$.

[20] Note: $\lim_{m \to 0^+} \frac{\mathrm{d}x^*}{\mathrm{d}m} = \lim_{m \to 0^-} \frac{\mathrm{d}x^*}{\mathrm{d}m} = \frac{-\sigma^2}{2c^2} e^{\frac{-\sigma}{c}} < 0$.

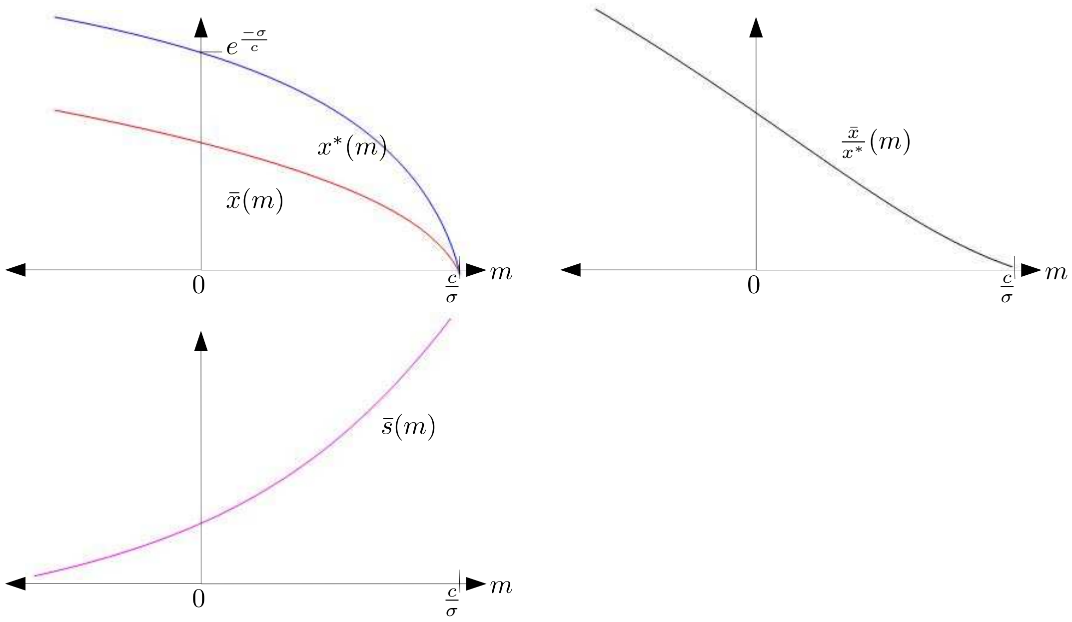

Figure 4: Equilibrium indexed by m

Intuitively, decreasing the efficiency of technology diffusion by increasing m also decreases the incentive to fall-back. Formally, $\frac{\mathrm{d}}{\mathrm{d}m}(D'(x;m)) = \frac{\partial^2 D(x;m)}{\partial x \partial m} = -cx^{-1+m}\ln(x) > 0$. Since $D'(x;m) < 0$, increasing m lowers the absolute value of $D'(x;m)$ which lowers the fall-back incentive and decreases $\bar{x}$.

Proposition 5. *As curvature in the catch-up function changes to make imitation less efficient, the BGP productivity ratio of imitators decreases if* $\frac{m}{\bar{q}}\frac{d\bar{q}}{dm} < \ln\bar{q}$*. That is* $\frac{\mathrm{d}\bar{x}}{\mathrm{d}m} < 0$.

Proof. See Appendix D.2. In Appendix D.1, $\frac{\mathrm{d}\bar{q}}{\mathrm{d}m} < 0$ is shown using just Assumptions 1, 2, and 3. ☐

Proposition 6. *For the parameter restrictions in Appendix D.3, as curvature in the catch-up function changes to make imitation less efficient, the BGP level of expenditure on imitation increases. That is* $\frac{\mathrm{d}\bar{s}}{\mathrm{d}m} > 0$.

Proof. See Appendix D.3. ☐

The parameter restrictions in Propositions 5 and 6 are in addition to those in Assumptions 1, 2, and 3. However, in all numerical exercises we have examined, Assumptions 1, 2, and 3 alone have ensured $\frac{\mathrm{d}\bar{x}}{\mathrm{d}m} < 0$ and $\frac{\mathrm{d}\bar{s}}{\mathrm{d}m} > 0$.

Figure 4 illustrates the comparative dynamics of the economy with respect to changes in the efficiency of technology diffusion, m. Proposition 4 states that decreasing the efficiency of diffusion causes x^* to fall. That is, as technology diffusion becomes less efficient, a larger mass of agents become innovators and fewer choose to be imitators. As Proposition 5 shows, not only are there fewer imitators as m increases, but imitators choose a BGP productivity ratio, $\bar{x}$, that decreases with m. With a lower efficiency diffusion process, imitating agents optimally respond by letting the ratio of their productivity to the technology frontier slip lower and they benefit from the higher diffusion rate that comes from falling behind the frontier. Since $\bar{s}$ is increasing in m, along a new BGP with less efficient diffusion, imitators increase expenditures that aid technology diffusion, even as they fall further behind the technology frontier.

If we think of the recent trend of globalization and improvements in information technology as a decrease in m, this theory suggests that the new BGP would feature fewer innovators and more imitators, with the imitators spending less to increase diffusion and having a relatively higher ratio of productivity to the technology frontier.

6.3 Does Less Efficient Diffusion Technology Always Lower $\bar{x}$?

Proposition 5, shows that as m increases and imitation gets less efficient, the BGP productivity ratio of imitators always decreases. Imitators choose $\bar{x}$ optimally to balance the incentives for catch-up and fall-back. Since m is not the only parameter controlling the efficiency of technology diffusion, do imitating agents always choose to fall-back as imitation becomes less efficient? Equivalently, does there exist a sequence of $D(x)$ functions where imitation becomes strictly less efficient and agents choose to increase $\bar{x}$?

To show that such a non-monotonicity can occur, Figure 5 conducts a variation of the experiment described in Section 6.2. To construct the sequence of $D(x)$, this new experiment will vary m while changing c to keep $x^*(m)$ constant. Hence, instead of the constant $\hat{c}$ with the catch-up function $D(x; m, \hat{c})$, the multiplicative term on the catch-up function will be adjusted to $D(x; m, c(m))$. The $c(m)$ function is implicitly defined by $\left(1 - \frac{m\sigma}{c(m)}\right)^m = x^*(\hat{c})$. For all agents in the imitation region, $\frac{\partial D(x;m,c(m))}{\partial m} < 0$ for all $x < x^*$, so an increase in m makes imitation less productive. This particular way of changing c and m also has the advantage of keeping x^* constant, which isolates the change in economic incentives on the intensive margin.

Figure 5 compares the results of changing both c and m to those of Section 6.2 where only m changes.[21] The left column shows results for an experiment where c is held constant at $\hat{c}$ and only m varies—similar to Figure 4. The top row shows $D(x; m, \hat{c})$ for several values of m and the bottom row shows $\bar{x}(m, \hat{c})$ and $x^*(m, \hat{c})$ as a function of m. The right column presents the new results, where the top row shows $D(x; m, c(m))$ and the bottom row shows $\bar{x}(m, c(m))$ and $x^*(m, c(m))$.

The results show the non-monotonicity of $\bar{x}(m, c(m))$ even while $x^*(m, c(m))$ is constant. For $m < 0$, the fall-back incentive dominates and agents choose to decrease $\bar{x}$. For $m > 0$, the imitation technology becomes so inefficient that the catch-up incentive dominates and agents choose to increase $\bar{x}$.

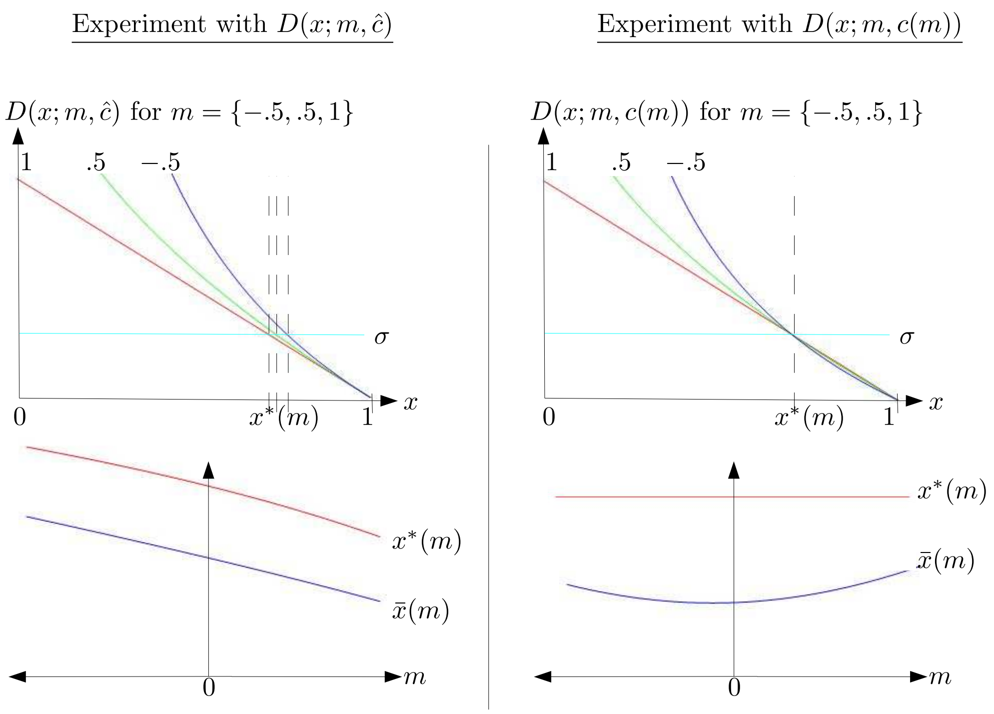

Figure 5: Non-monotonicity of $\bar{x}$

[21] The parameters used in this simulation are $\sigma = 0.9, \hat{c} = 2, B = 2$ and $r = 0.05$. The normalization parameter $\hat{c}$ is chosen such that x^* and $\bar{x}$ are equal across the two experiments at $m = 1$. Figures are not drawn to scale.

6.4 Comparative Dynamics of c

While both c and m determine the marginal productivity of imitation, m has a distorting effect depending on the distance to the frontier, while c changes the marginal productivity of innovation linearly and uniformly for all x.

From equation 52 and Proposition 1, it follows that as c grows large, both $\bar{x}$ and x^* go to 1 and the region where imitators fall behind will disappear. For agents below $\bar{x}$, growth rates increase and approach infinity as they approach the frontier. So for large c, we expect only the frontier to do research and all other agents to be infinitesimally close to the frontier doing imitation.[22]

A critical feature of our catch-up specification is that technology diffusion is not instantaneous; the growth of productivity at each instant depends on some measure of the distance to the frontier. The productivities of firms or countries grow incrementally through diffusion from existing superior technologies at a rate that depends on their investment in diffusion. This formulation differs from instantaneous one-shot adoption which allows the implementation of a new technology drawn from the distribution of existing superior technologies, as in Perla and Tonetti (2011) or Lucas and Moll (2012). The limits for large c give us a way to see how a nearly instantaneous jump occurs, albeit here agents jump to the unique frontier rather than jumping to the productivity of another agent in the economy, as in Perla and Tonetti (2011) or Lucas and Moll (2012).

7 Conclusion

All agents are identical except for their initial levels of productivity and they choose the intensity of their investments to promote growth through imitation and innovation optimally. Innovation is a simple technology with a constant marginal productivity that does not benefit from externalities. Imitation facilitates technology diffusion from the frontier, with the marginal productivity of imitation modeled by a catch-up function that increases with distance to the technology frontier. In equilibrium, agents optimally segment into innovators and imitators and converge to a unique and stable balanced growth path. All innovators grow at the same rate, and therefore the ratio of their productivity to the technology frontier remains unchanged. In the limit, the ratio of every imitator's productivity to the technology frontier converges to a value that lies below the threshold, where the incentives for catch-up and fall-back are balanced. When agents find their productivities far below the technology frontier, they choose investments to facilitate the diffusion of technology and catch-up to the balanced growth path by growing quickly. If agents are close to or at the frontier, they focus their investments on promoting technological innovation. Agents with relatively high productivity, but not high enough to innovate, optimally fall-back to the balanced growth path of imitators to take advantage of more efficient technology diffusion. By focusing on the portfolio choice of allocating expenditures between imitation and innovation we describe the key forces that provide incentives for imitators to fall-back and converge to a balanced growth path that lies below the technology frontier.

[22] These arguments are loose in that the sufficient conditions for concavity do not necessarily hold for large c. Also, note that the flow costs $s(t)z(t)$ of an instantaneous jump to the frontier are negligible in our setup as $c \to \infty$. If there is a discrete as opposed to flow consumption cost proportional to $z(t)$ for enabling technology diffusion, whether a jump occurs or not will depend on the discrete benefit of jumping to the frontier technology relative to the discrete cost.

References

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

Aghion, P., and P. Howitt (2009): *The Economics of Growth*. MIT Press.

Aghion, P., P. Howitt, and D. Mayer-Foulkes (2005): "The Effect of Financial Development on Convergence: Theory and Evidence," *The Quarterly Journal of Economics*, 120, 173–222.

Alvarez, F., F. Buera, and R. E. Lucas (2011): "Idea Flows, Economic Growth, and Trade," *mimeo.*

Barro, R. J., and X. Sala-i-Martin (1997): "Technological Diffusion, Convergence, and Growth," *Journal of Economic Growth*, 2(1), 1–26.

Benhabib, J., and M. M. Spiegel (1994): "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data," *Journal of Monetary Economics*, 34(2), 143–173.

——— (2005): "Human Capital and Technology Diffusion," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf, vol. 1 of *Handbook of Economic Growth*, chap. 13, pp. 935–966. Elsevier.

Eeckhout, J., and B. Jovanovic (2002): "Knowledge Spillovers and Inequality," *American Economic Review*, 92(5), 1290–1307.

Gerschenkron, A. (1962): *Economic Backwardness in Historical Perspective*. Belknap Press of Harvard University Press.

Gill, I. S., and M. Raiser (2012): *Golden Growth: Restoring the Lustre of the European Economic Model*. The World Bank.

Jones, C. I. (2005): "Growth and Ideas," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf, vol. 1 of *Handbook of Economic Growth*, chap. 16, pp. 1063–1111. Elsevier.

Klenow, P. J., and A. Rodriguez-Clare (2005): "Externalities and Growth," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf, vol. 1 of *Handbook of Economic Growth*, chap. 11, pp. 817–861. Elsevier.

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

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

Lucas, R. E., and B. Moll (2012): "Knowledge Growth and the Allocation of Time," *mimeo.*

Luttmer, E. G. J. (2011): "Technology Diffusion and Growth," *Journal of Economic Theory*, 147(2), 602–622.

Magaziner, I., and M. Patinkin (1990): *The Silent War*. Vintage Books: New York.

Milnor, J. W. (1965): *Topology from the Differentiable Viewpoint*. The University Press of Virginia.

Nelson, R. R., and E. S. Phelps (1966): "Investment in Humans, Technological Diffusion, and Economic Growth," *American Economic Review*, 56(1/2), 69–75.

Parente, S. L., and E. C. Prescott (1994): "Barriers to Technology Adoption and Development," *The Journal of Political Economy*, 102(2), 298–321.

Perla, J., and C. Tonetti (2011): "Endogenous Risk and Growth," *NYU mimeo*.

Seierstad, A., and K. Sydsaeter (1977): "Sufficient Conditions in Optimal Control Theory," *International Economic Review*, 18(2), 367–391.

Appendix A Stability and Uniqueness

A.1 Uniqueness of Interior Stationary $\bar{x}$

Lemma 2. *For any quadratic of form $y^2 - ay + b$ with roots y_1 and y_2, if $a, b > 0$ and Discriminant > 0, then both roots are real and positive. Furthermore, $y_1, y_2 \notin [\min\{b, 1\}, \max\{b, 1\}]$ if $a > 1 + b$.*

Proof. The result that roots are real and positive is a standard property of quadratic equations, as b is the product of the roots and a is the sum of the roots. Assume that there exists a root $\hat{y} \in [\min\{b, 1\}, \max\{b, 1\}]$. Note that the product of the roots is given by $P(\hat{y}) = \hat{y}(a - \hat{y})$. Since $P(\hat{y})$ is concave, the argmin of P is either b or 1. If we show that $P(b) > b$ and $P(1) > b$, then we have a contradiction. $P(1) = a - 1$ and $P(b) = ba - b^2$, so $P(1) > b$ if $a > 1 + b$ and $P(b) > b$ if $a > 1 + b$. Thus if $a > 1 + b$, $\hat{y}$ can not be in $[\min\{b, 1\}, \max\{b, 1\}]$. ☐

To show that a unique root $\bar{x} \in (0, x^*)$ exists, we take equation 52 in the form of $\bar{q}^2 - a\bar{q} + b$. For this function,

$$
\begin{aligned}
a &= 2 - \left(\frac{m\sigma}{c}\right) + \frac{m^2}{c}\left(\sigma - \frac{r}{B}\right) \\
b &= 1 - \frac{m\sigma}{c}
\end{aligned}
$$

$b > 0$ from Assumption 2. $a > 0$ using Assumptions 1 and 2. Additionally, the discriminant is positive:

$$
D = \left(\left(2 - \left(\frac{m\sigma}{c}\right) + \frac{m^2}{c}\left(\sigma - \frac{r}{B}\right)\right)\right)^2 - 4\left(1 - \frac{m\sigma}{c}\right) > \left(2 - \left(\frac{m\sigma}{c}\right)\right)^2 - 4\left(1 - \frac{m\sigma}{c}\right) \tag{A.1}
$$

$$
= 4 - 4\left(\frac{m\sigma}{c}\right) + \left(\frac{m\sigma}{c}\right)^2 - 4 + 4\left(\frac{m\sigma}{c}\right) = \left(\frac{m\sigma}{c}\right)^2 \geq 0 \tag{A.2}
$$

Thus the $\bar{q}$ roots are real and positive, and hence $\bar{x}$ roots are real and positive. It remains to show that there is a unique $\bar{x} \in (0, x^*)$. That is, there exists a unique $\bar{q}$ s.t. $\bar{q}^{\frac{1}{m}} < b^{1/m}$. Using Assumption 1, the conditions in Lemma 2 are satisfied, i.e., $a > 1 + b$

$$
\begin{aligned}
a - (1 + b) &= 2 - \frac{m\sigma}{c} + \frac{m^2}{c}\left(\sigma - \frac{r}{B}\right) - (1 + 1 - \frac{m\sigma}{c}) \\
&= \frac{m^2}{c}\left(\sigma - \frac{r}{B}\right) > 0
\end{aligned}
$$

Thus, by Lemma 2, $\bar{q}_1, \bar{q}_2 \notin [\min\{1 - \frac{m\sigma}{c}, 1\}, \max\{1 - \frac{m\sigma}{c}, 1\}]$. Note for $m > 0$, if both roots were less than $1 - \frac{m\sigma}{c} < 1$ or both roots were greater than one, the product could not be $1 - \frac{m\sigma}{c}$. Note for $m < 0$, if both roots were greater than $1 - \frac{m\sigma}{c} > 1$ or both roots were less than one, the product could not be $1 - \frac{m\sigma}{c}$. Let convention be that $\bar{q}_1 < \bar{q}_2$. Thus $\bar{q}_1 < 1 - \frac{m\sigma}{c}$ and $\bar{q}_2 > 1 - \frac{m\sigma}{c}$.

Note if $m > 0$, then $\bar{x} < (1 - \frac{m\sigma}{c})^{1/m}$ iff $\bar{q} < (1 - \frac{m\sigma}{c})$ and if $m < 0$, $\bar{x} < (1 - \frac{m\sigma}{c})^{1/m}$ iff $\bar{q} > (1 - \frac{m\sigma}{c})$. Therefore, the smaller root, $\bar{q}_1$, is the unique stationary solution if $m > 0$ and the larger root, $\bar{q}_2$, is the unique stationary solution if $m < 0$. Thus there exists a unique root $\bar{x} \in (0, x^*)$.

A.2 Stability of BGP

The proof of stability for the entire parameter space is completed as follows: First, prove stability for a particular r by ensuring the determinant of the Jacobian of the dynamic system $\Psi(\cdot, \cdot)$, $J_\Psi(\cdot, \cdot)$, is negative. Then, use a homotopy argument to show that the system must remain stable as r is varied throughout the rest of the valid parameter space.

Repeating equation 46,

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} s \\ q \end{pmatrix} = \Psi(s,q) \equiv \begin{pmatrix} \frac{(B-s)\left(Bc-mr+q(-2Bc+m(-\sigma B-rm+r)+Bcq)-c(-1+q)^2 s\right)}{m(-1+q)} \\ mq\left(\frac{c}{m}(1-q)s-\sigma B-r\right) \end{pmatrix}$$

Taking the Jacobian of this and simplifying

$$J_\Psi(s,q) = \begin{pmatrix} \frac{-2Bc+mr-q\left(m(1+m)r-B\left(4c+m^2\sigma\right)+2Bcq\right)+2c(-1+q)^2 s}{m(-1+q)} & \frac{(B-s)\left(-m^2r+B\left(c+m^2\sigma\right)+Bc(-2+q)q-c(-1+q)^2 s\right)}{m(-1+q)^2} \\ -c(-1+q)q & m(r-B\sigma)+(c-2cq)s \end{pmatrix} \tag{A.3}$$

For the starting point of the homotopy argument, take the limit as $r \to \sigma B$, where $g \to 0$. As proven in Appendix A.1, $\bar{q}_1$ is the unique stationary root if $m > 0$, and $\bar{q}_2$ is the unique stationary root if $m < 0$. In either case, the stationary solution, $\Psi(\bar{s},\bar{q}) = 0$, as $r \to \sigma B$ is

$$\bar{s} = 0, \quad \bar{q} = 1 - \tfrac{m\sigma}{c} \tag{A.4}$$

Substituting these into the the Jacobian in equation A.3

$$J_\Psi(\bar{s},\bar{q}) = \begin{pmatrix} -Bc-(1+m)r+\frac{cr}{\sigma}+B(2+m)\sigma & \frac{Bc\left(-cr+Bc\sigma+B\sigma^2\right)}{m\sigma^2} \\ m\sigma\left(1-\frac{m\sigma}{c}\right) & m(r-B\sigma) \end{pmatrix}$$

Taking $r \to \sigma B$ and simplifying further

$$= \begin{pmatrix} B\sigma & (B^2c)/m \\ m\sigma\left(1-\frac{m\sigma}{c}\right) & 0 \end{pmatrix} \tag{A.5}$$

The determinant of this Jacobian evaluated at the steady state in the limit as $r \to \sigma B$ is

$$\det\left(J_{\bar{\Psi}}\right) \equiv \det\left(J_\Psi(\bar{s},\bar{q})\right) = B^2\sigma(m\sigma-c) \tag{A.6}$$

The parameter restriction in Assumption 2 ensures that $(m\sigma - c) < 0$, and thus $\det\left(J_{\bar{\Psi}}\right) < 0$. Thus for a small neighborhood around $r = \sigma B$ the system has local saddle-point stability. For clarity of exposition, let $\Gamma(s,x)$ be the equivalent system to $\Psi(s,q)$ in x space. Since $\det\left(J_{\bar{\Psi}}\right) < 0$ near $r = B\sigma$, then $\det\left(J_{\bar{\Gamma}}\right) < 0$ near $r = B\sigma$.

To complete the proof, use a homotopy argument for varying r, by parameterizing the system of equations by r: $\Gamma_r(s(r),x(r))$. Let $y(r) = (s(r),x(r))$ and the steady state, $\bar{y}(r) = (\bar{s}(r),\bar{x}(r))$, be defined by $\bar{\Gamma}_r \equiv \Gamma_r(\bar{y}) = 0$ for $r \in S$ where $S = \{\varepsilon, B\sigma - \varepsilon\}$ for small $\varepsilon > 0$. Since $\det\left(J_{\bar{\Gamma}_r}\right) < 0$ near $r = B\sigma$, it is negative at $r = B\sigma - \varepsilon$ for small $\varepsilon > 0$.

Let $\{\bar{y}(r)\} = \left\{\Gamma_r^{-1}(0)\right\}$ for isolated zeros of $\Gamma_r(\bar{y})$. For $r \in S$ let $\bar{y}(r) \in \mathcal{D} \subset (0,B) \times (0,1)$, where $\mathcal{D}$ is chosen so that there are no $\bar{y}(r) \in \partial\mathcal{D}$. Then the homotopy invariant Poincare-Hopf degree on the boundary $\partial\mathcal{D}$ of $\mathcal{D}$ is $\sum_{y \in \{\bar{y}(r)\}} \operatorname{sign}\det\left(J_{\Gamma_r}(y)\right) = 1$ or -1. See Milnor (1965). Since the stationary point $\bar{x}(r)$, and therefore $\bar{y}(r)$, is always unique for $r \in S$, then $\operatorname{sign}\det\left(J_{\Gamma_r}(\bar{y}(r))\right) = \operatorname{sign}\det J_{\Gamma_r}(\bar{y}(B\sigma-\varepsilon)) = -1$.

If the roots of $J_{\Gamma_r}(\bar{y}(r))$ are real, then they must be of opposite sign. Therefore $\det\left(J_{\Gamma_r}(\bar{y}(r))\right) < 0$ for

$r \in S$ and the system is locally saddle-point stable: given $x(0)$ we can choose $s(0)$ on the stable manifold converging to $\bar{y}(r)$. Note that the roots of $\det(J_{\Gamma_r}(\bar{y}(r)))$ cannot be complex since the determinant, the product of the roots, would have to be positive in that case. Therefore x is a continuous function of time described by the solution of $\dot{x} = \sigma\gamma x + s(x)\,x\frac{c}{m}(1-x^m)$ with $s(x)$ optimally chosen for $r \in S$. Therefore the BGP has to be globally stable since $\frac{dx}{dt} \neq 0$ at points other than at the unique $\bar{x}(r)$, or the BGP would not be unique. Thus, it must be that $\frac{dx}{dt} > 0$ to the left of $\bar{x}$ and $\frac{dx}{dt} < 0$ to the right of $\bar{x}$, so the unique BGP is stable: all agents with productivities $x(t) < x^*$ eventually end up at $\bar{x}(r)$.

Appendix B Concavity and Sufficient Parameter Restrictions

For the Hamiltonian

$$H(x,s,\gamma,\lambda) = \ln(x) + \lambda x\sigma\gamma + \ln(B-s-\gamma) + \lambda x\left(\tfrac{c}{m}(1-x^m)s - g\right)$$

define $\hat{H}(x,\lambda) = \max_{\gamma,s} H(x,s,\gamma,\lambda)$.

Lemma 3. *The maximized Hamiltonian, $\hat{H}(x,\lambda)$ is concave in x if $m \geq -1$ and $\frac{c}{\sigma} < m+2$.*

Note that both conditions in the Lemma above restrict the efficiency of technology diffusion.

Proof. Differentiating $H(x,s,\gamma,\lambda)$

$$H_{ss} = -\frac{1}{(B-s-\gamma)^2};\ \gamma = 0 \tag{B.1}$$

$$H_{sx} = \tfrac{c}{m}\lambda\left(1-(1+m)x^m\right);\quad \gamma = 0 \tag{B.2}$$

$$H_{xx} = -\frac{1+c(1+m)\lambda s x^{1+m}}{x^2};\ \gamma = 0 \tag{B.3}$$

$$H_{\gamma\gamma} = -\frac{1}{(B-s-\gamma)^2};\ s = 0 \tag{B.4}$$

$$H_{\gamma x} = \lambda\sigma;\quad s = 0 \tag{B.5}$$

$$H_{xx} = -x^{-2}\quad;\quad s = 0 \tag{B.6}$$

Over the innovation only region $x > x^*$, $\gamma(x,\lambda) > 0$, so we can compute $\frac{\partial\gamma}{\partial x}$ from the first order condition $H_\gamma = 0$. Then by the envelope condition $\frac{d\hat{H}}{dx} = H_x + H_\gamma\frac{\partial\gamma}{\partial s}$ and $\frac{d^2\hat{H}}{dx^2} = H_{xx} + H_{x\gamma}\frac{\partial\gamma}{\partial s}$, which must be non-positive to assert the concavity $\hat{H}(x,\lambda)$ in x. The concavity of $\hat{H}(x,\lambda)$ in that region will then follow if $\frac{d^2\hat{H}}{dx^2} = H_{xx} + H_{x\gamma}\frac{\partial\gamma}{\partial x} = H_{xx} - \frac{(H_{x\gamma})^2}{H_{\gamma\gamma}} \leq 0$ since $\gamma > 0$ and thus $H_\gamma = 0$. Therefore in the region $\frac{c}{m\sigma}(1-x^m) < 1$ where $s = 0$, $\gamma > 0$ we have $H_{\gamma\gamma} = -(B-\gamma)^{-2} < 0$, $H_{xx} = -x^{-2} < 0$, $H_{\gamma z} = \sigma\lambda > 0$,

$$H_{xx} - \frac{(H_{\gamma x})^2}{H_{\gamma\gamma}} = -x^{-2} + (B-\gamma)^2(\sigma\lambda)^2 \tag{B.7}$$

$$= -x^{-2} + \frac{(\sigma\lambda)^2}{(z\sigma\lambda)^2} = 0 \tag{B.8}$$

This is enough to establish the required concavity over the range $(x^*, 1)$ where $s = 0$, and together with the first order optimality conditions establishes the optimal solution for the region where $\gamma(x,\lambda) > 0$ and $s = 0$.

The range $x < x^*$, where $s \geq 0$, and $\gamma = 0$, is more complicated. Since $\gamma = 0$ over the latter region, either $s = 0$ over the interior of a sub-interval where $\frac{\partial s}{\partial x} = 0$, or $s > 0$ so that $\frac{\partial s}{\partial x}$ can be computed from the first order condition $H_s = 0$. So either $H_s = 0$ or $\frac{\partial s}{\partial x} = 0$. Then by the envelope condition $\frac{d\hat{H}}{dx} = H_x + H_s\frac{\partial s}{\partial x}$

and $\frac{d^2\hat{H}}{dx^2} = H_{xx} + H_{xs}\frac{\partial s}{\partial x}$, which must be non-positive to assert the concavity $\hat{H}(x,\lambda)$. The concavity of $\hat{H}(x,\lambda)$ in that region will then follow if $\frac{d^2\hat{H}}{dx^2} = H_{xx} + H_{xs}\frac{\partial s}{\partial x} = H_{xx} - \frac{(H_{xs})^2}{H_{ss}} < 0$ if $H_s = 0$ or, $\frac{d^2\hat{H}}{dx^2} = H_{xx}$ if s is in the interior of a sub-interval where $s = 0$. We have:

$$H_{ss} = -\frac{1}{(B-s)^2};\ \gamma = 0 \tag{B.9}$$

$$H_{sx} = \tfrac{c}{m}\lambda\left(1-(1+m)x^m\right);\quad \gamma = 0 \tag{B.10}$$

$$H_{xx} = -\frac{1+c(1+m)\lambda s x^{1+m}}{x^2};\ \gamma = 0 \tag{B.11}$$

Therefore we need $Q = H_{xx} - \frac{(H_{sx})^2}{H_{ss}} < 0$. Evaluating we obtain:

$$Q \quad = \quad -\frac{1+c(1+m)\lambda s x^{1+m}}{x^2} + \left(\tfrac{c}{m}\lambda\left(1-(1+m)x^m\right)\right)^2 (B-s)^2 \tag{B.12}$$

The first order condition for s in the range $x < x^*$ is:

$$\frac{1}{B-s} \quad \geq \quad \lambda x\frac{c}{m}\left(1-x^m\right) \tag{B.13}$$

$$B-s \quad \leq \quad \left(\lambda x\frac{c}{m}\left(1-x^m\right)\right)^{-1} \tag{B.14}$$

$$(B-s)^2 \quad \leq \quad \left(\lambda x\frac{c}{m}\left(1-x^m\right)\right)^{-2} \tag{B.15}$$

From equation B.12, $\hat{H}(x,\lambda)$ is concave if $Q < 0$ and since $x > 0$, $\hat{H}(x,\lambda)$ is concave if $Qx^2 < 0$. Using the FOC:

$$Qx^2 \quad = \quad -\left(1+c(1+m)\lambda s x^{1+m}\right) + x^2\left(\tfrac{c}{m}\lambda\left(1-(1+m)x^m\right)\right)^2 (B-s)^2 \tag{B.16}$$

$$\leq \quad -\left(1+c(1+m)\lambda s x^{1+m}\right) + x^2\left(\tfrac{c}{m}\lambda\left(1-(1+m)x^m\right)\right)^2 \left(\lambda x\frac{c}{m}\left(1-(x)^m\right)\right)^{-2} \tag{B.17}$$

Then, $\hat{H}(x,\lambda)$ is concave in x if

$$-\left(1+c(1+m)\lambda s x^{1+m}\right) + x^2\left(\tfrac{c}{m}\lambda\left(1-(1+m)x^m\right)\right)^2 \left(\lambda x\frac{c}{m}\left(1-(x)^m\right)\right)^{-2} < 0$$

$$x^2\left(\tfrac{c}{m}\lambda\left(1-(1+m)x^m\right)\right)^2 \left(\lambda x\frac{c}{m}\left(1-x^m\right)\right)^{-2} < 1+c(1+m)\lambda s x^{1+m} \tag{B.18}$$

Since $1+c(1+m)\lambda s x^{1+m} \geq 1$ for $m \geq -1$, we can obtain sufficient conditions for $Q < 0$ by checking for conditions under which

$$x^2\left(\tfrac{c}{m}\lambda\left(1-(1+m)x^m\right)\right)^2 \left(\lambda x\frac{c}{m}\left(1-x^m\right)\right)^{-2} < 1$$

for the range $x < x^*$. Reorganizing equation B.18,

$$\left(\frac{\left(\frac{c}{m}\lambda\left(1-(1+m)x^m\right)\right)}{\left(\lambda\frac{c}{m}\left(1-x^m\right)\right)}\right)^2 \quad < \quad 1 \leq 1+c(1+m)\lambda s x^{1+m} \tag{B.19}$$

$$\left(1-\frac{mx^m}{1-x^m}\right)^2 \quad < \quad 1 \tag{B.20}$$

Therefore, we need

$$-1 < 1-\frac{mx^m}{1-x^m} < 1 \tag{B.21}$$

The inequalities on the right holds since $\frac{mx^m}{1-x^m} > 0$ for any $m \neq 0$, so we focus on

$$-1 < 1 - \frac{mx^m}{1-x^m} \tag{B.22}$$

First, if $m < 0$ then

$$x^m - 1 > 1 - x^m - mx^m \tag{B.23}$$
$$(2+m)x^m > 2 \tag{B.24}$$

Note that $(2+m)x^m$ is decreasing in x for $0 > m > -2$. Since in this region x is bounded above by x^*, $(2+m)x^m$ is bounded below by $(2+m)x^{*m}$. Substituting $x^* = \left(1 - \frac{m\sigma}{c}\right)^{\frac{1}{m}}$ into the above yields

$$(2+m)\left(1 - \frac{m\sigma}{c}\right) > 2 \tag{B.25}$$
$$m\left(1 - 2\frac{\sigma}{c} - \frac{m\sigma}{c}\right) > 0 \tag{B.26}$$
$$1 - 2\frac{\sigma}{c} - \frac{m\sigma}{c} < 0 \tag{B.27}$$
$$\frac{c}{\sigma} < m + 2 \tag{B.28}$$

Similarly, if $m > 0$ then

$$x^m - 1 < 1 - x^m - mx^m \tag{B.29}$$
$$(2+m)x^m < 2 \tag{B.30}$$

Note that $(2+m)x^m$ is increasing in x for $m > 0$. Since in this region x is bounded above by x^*, $(2+m)x^m$ is bounded above by $(2+m)x^{*m}$. Substituting $x^* = \left(1 - \frac{m\sigma}{c}\right)^{\frac{1}{m}}$ into the above yields

$$(2+m)\left(1 - \frac{m\sigma}{c}\right) < 2 \tag{B.31}$$
$$m\left(1 - 2\frac{\sigma}{c} - \frac{m\sigma}{c}\right) < 0 \tag{B.32}$$
$$1 - 2\frac{\sigma}{c} - \frac{m\sigma}{c} < 0 \tag{B.33}$$
$$\frac{c}{\sigma} < m + 2 \tag{B.34}$$

Thus, $\hat{H}(x, \lambda)$ is concave in x if $m \geq -1$ and $\frac{c}{\sigma} < m + 2$. □

Appendix C Proofs for Hicks-Neutral Technical Change

Note that fixing $\bar{\theta} \equiv \frac{\bar{\sigma}}{\bar{c}}$, defining $\sigma = A\bar{\sigma}, c = A\bar{c}$, and taking derivatives with respect to A is mathematically equivalent to varying c and keeping $\bar{\theta}$ constant by varying σ. For mathematical convenience, we will use this alternative approach to the proofs.

C.1 Proof of $\frac{\mathrm{d}\bar{x}}{\mathrm{d}A} < 0$

Differentiating $\bar{x} = \bar{q}^{\frac{1}{m}}$,

$$\frac{d\bar{x}}{dc} = \frac{1}{m}\bar{q}^{\frac{1}{m}-1}\frac{d\bar{q}}{dc} \tag{C.1}$$

Differentiating $\bar{q}$ defined in equation 53:

For $m > 0$ the smaller root, $\bar{q}_1$, is the unique stationary root.

$$\begin{aligned}
\frac{d\bar{q}}{dc} &= \frac{1}{2}\frac{m^2 r}{Bc^2} - \frac{1}{4}\left(\left(2 + m\bar{\theta}\left(m-1\right) - \frac{m^2 r}{Bc}\right)^2 - 4\left(1 - m\bar{\theta}\right)\right)^{-.5}\left(2\left(2 + m\bar{\theta}\left(m-1\right) - \frac{m^2 r}{Bc}\right)\left(\frac{m^2 r}{Bc^2}\right)\right) \\
&= \frac{1}{2}\frac{m^2 r}{Bc^2}\left(1 - \frac{1}{2}\frac{\left(2\left(2 + m\bar{\theta}\left(m-1\right) - \frac{m^2 r}{Bc}\right)\right)}{\left(\left(2 + m\bar{\theta}\left(m-1\right) - \frac{m^2 r}{Bc}\right)^2 - 4\left(1 - m\bar{\theta}\right)\right)^{.5}}\right) \\
&< \frac{1}{2}\frac{m^2 r}{Bc^2}\left(1 - \frac{1}{2}\frac{\left(2\left(2 + m\bar{\theta}\left(m-1\right) - \frac{m^2 r}{Bc}\right)\right)}{\left(\left(2 + m\bar{\theta}\left(m-1\right) - \frac{m^2 r}{Bc}\right)^2\right)^{.5}}\right) \\
&= \frac{1}{2}\frac{m^2 r}{Bc^2}\left(1 - \frac{1}{2}\frac{\left(2\left(2 + m\bar{\theta}\left(m-1\right) - \frac{m^2 r}{Bc}\right)\right)}{\left(2 + m\bar{\theta}\left(m-1\right) - \frac{m^2 r}{Bc}\right)}\right) \\
&= \frac{1}{2}\frac{m^2 r}{Bc^2}\left(1 - 1\right) = 0
\end{aligned}$$

Hence for $m > 0$, $\frac{d\bar{q}}{dc} < 0$ which, together with equation C.1, proves $\frac{d\bar{x}}{dc} < 0$.

For $m < 0$ the larger root, $\bar{q}_2$, is the unique stationary root.

$$\begin{aligned}
\frac{d\bar{q}}{dc} &= \frac{1}{2}\frac{m^2 r}{Bc^2} + \frac{1}{4}\left(\left(2 + m\bar{\theta}\left(m-1\right) - \frac{m^2 r}{Bc}\right)^2 - 4\left(1 - m\bar{\theta}\right)\right)^{-.5}\left(2\left(2 + m\bar{\theta}\left(m-1\right) - \frac{m^2 r}{Bc}\right)\left(\frac{m^2 r}{Bc^2}\right)\right) \\
&= \frac{1}{2}\frac{m^2 r}{Bc^2} + \frac{1}{4}\left(\left(2\left(1 + \frac{m^2}{c}\left(\sigma - \frac{r}{B}\right) + \left(1 - \frac{m\sigma}{c}\right)\right)\right)^2 - 4\left(1 - \frac{m\sigma}{c}\right)\right)^{-.5} \\
&\quad \cdot \left(2\left(1 + \frac{m^2}{c}\left(\sigma - \frac{r}{B}\right) + \left(1 - \frac{m\sigma}{c}\right)\right)\left(\frac{m^2 r}{Bc^2}\right)\right)
\end{aligned}$$

Using results from Appendix A.1, $\left(1 + \frac{m^2}{c}\left(\sigma - \frac{r}{B}\right) + \left(1 - \frac{m\sigma}{c}\right)\right)$ is the sum of roots, which is positive and $\left(2\left(1 + \frac{m^2}{c}\left(\sigma - \frac{r}{B}\right) + \left(1 - \frac{m\sigma}{c}\right)\right)\right)^2 - 4\left(1 - \frac{m\sigma}{c}\right)$ is the discriminant, which is positive. Hence for $m < 0$, $\frac{d\bar{q}}{dc} > 0$ which, together with equation C.1, proves $\frac{d\bar{x}}{dc} < 0$.

Thus $\frac{d\bar{x}}{dc} < 0$ for constant θ, which implies $\frac{\mathrm{d}\bar{x}}{\mathrm{d}A} < 0$.

C.2 Proof of $\frac{d\bar{s}}{dc} > 0$

$$\frac{d\bar{s}}{dc} = \frac{\partial\bar{s}}{\partial c} + \frac{\partial\bar{s}}{\partial\bar{q}}\frac{d\bar{q}}{dc} \tag{C.2}$$

Solving first for $\frac{\partial\bar{s}}{\partial\bar{q}}$:

$$\bar{s} = \frac{B - \frac{rm}{c}\left(1-\bar{q}\right)^{-1}}{\left(1 + m\bar{q}\left(1-\bar{q}\right)^{-1}\right)} = \frac{B\left(\frac{1}{m}\left(1-\bar{q}\right)\right) - \frac{r}{c}}{\frac{1}{m}\left(1-\bar{q}\right) + \bar{q}} > 0 \tag{C.3}$$

$$\frac{\partial\bar{s}}{\partial\bar{q}} = -\frac{1}{c}\frac{m}{\left(m\bar{q} - \bar{q} + 1\right)^2}\left(r + Bc - mr\right) \tag{C.4}$$

Thus,

$$\frac{d\bar{s}}{dc} = \frac{\partial \bar{s}}{\partial c} + \frac{\partial \bar{s}}{\partial \bar{q}}\frac{d\bar{q}}{dc} \tag{C.5}$$

$$= \frac{\frac{r}{c^2}}{\frac{1}{m}(1-\bar{q})+\bar{q}} - \frac{1}{c}\frac{m}{(m\bar{q}-\bar{q}+1)^2}(r+Bc-mr)\frac{d\bar{q}}{dc} \tag{C.6}$$

$$= \frac{1}{c^2}\frac{rm}{(1-\bar{q})+m\bar{q}} - \frac{1}{c}\frac{m}{(m\bar{q}-\bar{q}+1)^2}(r+Bc-mr)\frac{d\bar{q}}{dc} \tag{C.7}$$

$$= \frac{1}{c}\frac{m}{(m\bar{q}-\bar{q}+1)}\left(\frac{r}{c} - \frac{(r+Bc-mr)}{(1-\bar{q})+m\bar{q}}\frac{d\bar{q}}{dc}\right) \tag{C.8}$$

There is no $\bar{\theta}$ in the equation for $\bar{s}$ (equation 48), but this productivity ratio is built into the $\frac{d\bar{q}}{dc}$ term by taking $\theta = \frac{\sigma}{c}$ constant.

From Appendix C.1, if $m > 0$ then $\bar{q} < 1$ and $\frac{d\bar{q}}{dc} < 0$. Then,

$$\frac{d\bar{s}}{dc} > 0 \text{ if } 0 < m \leq 1\text{, or better, if } r(1-m)+Bc > 0 \tag{C.9}$$

From Appendix C.1, if $m < 0$ then $\bar{q} > 1$ and $\frac{d\bar{q}}{dc} > 0$. Then $(1-\bar{q})+m\bar{q} < 0$. So,

$$\frac{d\bar{s}}{dc} > 0 \tag{C.10}$$

Appendix D Comparative Dynamics for m

D.1 $\bar{q}(m)$ is Decreasing

To find $\frac{\mathrm{d}\bar{q}}{\mathrm{d}m}$, take the total derivative of equation 52 with respect to $\bar{q}$ and m,

$$2\bar{q}d\bar{q} - \left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)d\bar{q} - \left(-\frac{\sigma}{c}+\frac{2m}{c}\left(\sigma-\frac{r}{B}\right)\right)\bar{q}dm - \frac{\sigma}{c}dm = 0$$

$$\left(2\bar{q} - \left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)\right)d\bar{q} - \left(\left(-\frac{\sigma}{c}+\frac{2m}{c}\left(\sigma-\frac{r}{B}\right)\right)q+\frac{\sigma}{c}\right)dm = 0$$

Reorganizing,

$$\frac{d\bar{q}}{dm} = \frac{\frac{2m}{c}\left(\sigma-\frac{r}{B}\right)\bar{q}+\frac{\sigma}{c}(1-\bar{q})}{2\bar{q}-\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)} \tag{D.1}$$

Lemma 4. $\frac{d\bar{q}}{dm} \leq 0$

Proof. Analyze the case of $m > 0$ and $m < 0$ separately.

Case $m > 0$: The unique stable root is $\bar{q}_1$.

$$\bar{q}_1 = \frac{1}{2}\left(\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right) - \left(\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)^2 - 4\left(1-\left(\frac{m\sigma}{c}\right)\right)\right)^{0.5}\right)$$

$$\frac{d\bar{q}}{dm} = \frac{\frac{2m}{c}\left(\sigma-\frac{r}{B}\right)\bar{q}+\frac{\sigma}{c}(1-\bar{q})}{\left(\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right) - \left(\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)^2 - 4\left(1-\left(\frac{m\sigma}{c}\right)\right)\right)^{0.5}\right) - \left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)}$$

$$\frac{d\bar{q}}{dm} = \frac{\frac{2m}{c}\left(\sigma - \frac{r}{B}\right)\bar{q} + \frac{\sigma}{c}\left(1-\bar{q}\right)}{-\left(\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)^2 - 4\left(1-\left(\frac{m\sigma}{c}\right)\right)\right)^{0.5}} < 0$$

since, as we have shown, the discriminant $\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)^2 - 4\left(1-\left(\frac{m\sigma}{c}\right)\right) > 0$.

Case $m < 0$: The unique stable root is $\bar{q}_2$.

$$\bar{q}_2 = \frac{1}{2}\left(\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right) + \left(\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)^2 - 4\left(1-\left(\frac{m\sigma}{c}\right)\right)\right)^{0.5}\right)$$

$$\frac{d\bar{q}}{dm} = \frac{\frac{2m}{c}\left(\sigma - \frac{r}{B}\right)\bar{q} + \frac{\sigma}{c}\left(1-\bar{q}\right)}{2\bar{q} - \left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)}$$

$$\frac{d\bar{q}}{dm} = \frac{\frac{2m}{c}\left(\sigma - \frac{r}{B}\right)\bar{q} + \frac{\sigma}{c}\left(1-\bar{q}\right)}{\left(\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right) + \left(\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)^2 - 4\left(1-\left(\frac{m\sigma}{c}\right)\right)\right)^{0.5}\right) - \left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)}$$

$$\frac{d\bar{q}}{dm} = \frac{\frac{2m}{c}\left(\sigma - \frac{r}{B}\right)\bar{q} + \frac{\sigma}{c}\left(1-\bar{q}\right)}{\left(\left(2-\left(\frac{m\sigma}{c}\right)+\frac{m^2}{c}\left(\sigma-\frac{r}{B}\right)\right)^2 - 4\left(1-\left(\frac{m\sigma}{c}\right)\right)\right)^{0.5}} < 0$$

since here $m < 0$ and $\bar{q} > 1$. $\square$

D.2 Parameter Restrictions for $\frac{\mathrm{d}\bar{x}}{\mathrm{d}m}$

Now consider $\bar{x}$,

$$\bar{x} = \bar{q}^{\frac{1}{m}}, \quad 0 < \bar{x} < 1$$

$$\frac{d\bar{x}}{dm} = -\frac{1}{m^2}\bar{q}^{\frac{1}{m}}\ln\bar{q} + \frac{1}{m}\bar{q}^{\frac{1}{m}-1}\frac{d\bar{q}}{dm}$$

$$\frac{d\bar{x}}{dm} = \frac{1}{m^2}\bar{q}^{\frac{1}{m}}\left(\frac{m}{\bar{q}}\frac{d\bar{q}}{dm} - \ln\bar{q}\right) \tag{D.2}$$

Using equation D.2, we can come up with an expression to verify that the $\bar{x}(m)$ is decreasing in m:

$$0 > \frac{m}{\bar{q}}\frac{d\bar{q}}{dm} - \ln\bar{q} \tag{D.3}$$

Here, we can use the analytic expressions for $\frac{d\bar{q}}{dm}$ from equation D.1 and $\bar{q}$ from equation 53.

We suspect, but have not been able to analytically show, that Assumptions 1, 2, and 3 imply this restriction always holds. Numerically, for all parameter values within Assumptions 1, 2, and 3 we have tried, this requirement has not been violated.

D.3 Parameter Restrictions for $\frac{\mathrm{d}\bar{s}}{\mathrm{d}m}$

Looking at the expression for $\bar{s}(m)$ in equation 48, we will need to take its derivative with respect to m. We need to find the derivative of $\bar{s}$ with respect to m. Since $\bar{s}$ is a function of other equilibrium values,

$$\frac{d\bar{s}}{dm} = \frac{\partial\bar{s}}{\partial m} + \frac{\partial\bar{s}}{\partial\bar{q}}\frac{d\bar{q}}{dm} \tag{D.4}$$

These partials are taken using the expression for $\bar{s}(m)$ in equation 48

$$\frac{\partial \bar{s}}{\partial m} = \frac{1}{c}\left(r + Bc\bar{q}\right)\frac{\bar{q}-1}{\left(m\bar{q}-\bar{q}+1\right)^2} \tag{D.5}$$

$$\frac{\partial \bar{s}}{\partial \bar{q}} = -\frac{1}{c}\frac{m}{\left(m\bar{q}-\bar{q}+1\right)^2}\left(r + Bc - mr\right) \tag{D.6}$$

An expression for $\frac{\mathrm{d}\bar{q}}{\mathrm{d}m} < 0$ is given in Lemma 4. Compiling the elements of the total derivative in equation D.4,

$$\frac{d\bar{s}}{dm} = \frac{1}{c}\left(r + Bc\bar{q}\right)\frac{\bar{q}-1}{\left(m\bar{q}-\bar{q}+1\right)^2} - \left(\frac{1}{c}\frac{m}{\left(m\bar{q}-\bar{q}+1\right)^2}\left(r\left(1-m\right)+Bc\right)\right)\frac{d\bar{q}}{dm} \tag{D.7}$$

Hence, we can get a parameter constraint on $\frac{d\bar{s}}{dm} < 0$ by substituting for $\bar{q}$ from equation 53 and for $\frac{\mathrm{d}\bar{q}}{\mathrm{d}m}$ from equation D.1 into the following inequality:

$$0 < \frac{1}{c}\left(r + Bc\bar{q}\right)\frac{\bar{q}-1}{\left(m\bar{q}-\bar{q}+1\right)^2} - \left(\frac{1}{c}\frac{m}{\left(m\bar{q}-\bar{q}+1\right)^2}\left(r\left(1-m\right)+Bc\right)\right)\frac{d\bar{q}}{dm} \tag{D.8}$$

We suspect, but have not been able to analytically show, that Assumptions 1, 2, and 3 imply this restriction always holds. Numerically, for all parameter values within Assumptions 1,2, and 3 we have tried, this requirement has not been violated.

Appendix E Simplification of Quadratic for $\bar{q}$

We spell out in detail the simplification for equation 51:

$$\begin{aligned}
0 &= \frac{c}{m}\left(1-\bar{q}\right)^2 - \frac{r}{B}\left(1-\bar{q}\right) - \frac{\left(\sigma B - r\right)}{B}\left(1+\bar{q}(m-1)\right) \\
0 &= \left(1-\bar{q}\right)^2 - \frac{m}{c}\frac{r}{B}\left(1-\bar{q}\right) - \frac{m}{c}\left(\frac{\sigma B - r}{B}\right)\left(1+\bar{q}(m-1)\right) \\
0 &= 1 - 2\bar{q} + \left(\bar{q}\right)^2 - \frac{m}{c}\frac{r}{B} + \frac{m}{c}\frac{r}{B}\bar{q} - \frac{m}{c}\left(\sigma - \frac{r}{B}\right) - \frac{m}{c}\left(\sigma - \frac{r}{B}\right)(m-1)\bar{q} \\
0 &= \bar{q}^2 - \left(2 - \frac{m}{c}\frac{r}{B} + \frac{m}{c}\left(\sigma - \frac{r}{B}\right)(m-1)\right)\bar{q} + 1 - \frac{m}{c}\frac{r}{B} + \frac{m}{c}\frac{r}{B} - \frac{m}{c}\sigma \\
0 &= \bar{q}^2 - \left(2 + \frac{m^2}{c}\left(\sigma - \frac{r}{B}\right) - \frac{m}{c}\sigma\right)\bar{q} + \left(1 - \frac{m}{c}\sigma\right)
\end{aligned}$$