

# Catch-up and fall-back through innovation and imitation

Jess Benhabib · Jesse Perla · Christopher Tonetti

Published online: 18 October 2013  
© Springer Science+Business Media New York 2013

**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 “fall-back” 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.

**Keywords** Endogenous growth · Technology diffusion · Innovation · Imitation · Catch-up · Fall-back

**JEL Classification** O14 · O30 · O31 · O33 · O40

---

**Electronic supplementary material** The online version of this article (doi:[10.1007/s10887-013-9095-z](https://doi.org/10.1007/s10887-013-9095-z)) contains supplementary material, which is available to authorized users.

---

J. Benhabib (✉)  
Department of Economics, New York University,  
19 W 4th St, 6th Floor, New York, NY 10012, USA  
e-mail: [jess.benhabib@nyu.edu](mailto:jess.benhabib@nyu.edu)

J. Perla  
Vancouver School of Economics, University of British Columbia,  
997-1873 East Mall, Vancouver, BC V6T 1Z1, Canada  
e-mail: [jesse.perla@ubc.ca](mailto:jesse.perla@ubc.ca)

C. Tonetti  
Graduate School of Business, Stanford University,  
655 Knight Way, Stanford, CA 94305, USA  
e-mail: [tonetti@stanford.edu](mailto:tonetti@stanford.edu)

## 1 Introduction

Over the last decade, a number of newly industrialized countries, most prominently Russia, India, and China, have exhibited rapid sustained growth, often at rates of 5–10 % 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?<sup>1</sup> While there may be other explanations for this behavior, in this paper we provide a simple theory that highlights this fall-back motive to better understand incentives to invest in technology diffusion and innovation and the corresponding output distribution and growth implications.

There already exists a large literature studying endogenous innovation and imitation, starting with the early stages of the endogenous growth literature, and a survey or summary of this literature is beyond the scope of this paper. Some of the important contributions are [Jovanovic and Rob \(1989\)](#), [Grossman and Helpman \(1991\)](#), [Rustichini and Schmitz \(1991\)](#), [Segerstrom \(1991\)](#), [Aghion et al. \(2001\)](#), and [Mukoyama \(2003\)](#).

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

---

<sup>1</sup> 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 Golden Growth 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?” Indeed, market failures and political pressures may well prevent optimal investments in technology adoption and innovation. By contrast, from the perspective of this theory, imitation and fall-back may be an optimal outcome from the perspective of the social planner or representative consumer. We show that countries close to the technology frontier may in fact optimally choose to limit their investments in innovation and technology adoption in favor of higher immediate consumption, and rely on technology diffusion externalities to sustain their growth rate, even if their distance to the technology frontier initially slips a bit. This force is discussed in the context of the “benefits to backwardness” in [Klenow and Rodriguez-Clare \(2005\)](#).

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 et al. \(2005\)](#) 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.<sup>2,3</sup> More recently in a related approach, [Acemoglu et al. \(2012\)](#) explore the choice of social institutions that may limit inequality via a redistributive social welfare state and thereby reduce the incentives to innovate, compared to a social system that encourages “cutthroat capitalism” and innovation that advances the world technology frontier. They show that in the presence of technology spillovers from the world technology frontier both social systems will coexist in a best response equilibrium. [Damsgaard and Krusell \(2010\)](#) also explore the dynamics of innovation across countries where TFP growth can be driven by spillovers from the TFP technology frontier, which itself is an average of country TFPs and therefore endogenous.

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-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.

In our baseline model, agents 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.<sup>4</sup> This relationship is modeled

---

<sup>2</sup> Along similar lines, see [Howitt and Mayer-Foulkes \(2005\)](#). For an exposition see [Aghion and Howitt \(2009\)](#), pp. 152–158).

<sup>3</sup> [Acemoglu et al. \(2006\)](#) 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.

<sup>4</sup> 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 \(1997\)](#) 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.

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 of 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.<sup>5,6</sup>

Engaging in imitation has a trade-off: the agents gain 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.

There is an equilibrium ratio of productivity to the technology frontier defining the innovator threshold. All agents above this threshold choose to invest only in innovation, while all agents below the threshold entirely focus investment expenditures on facilitating technology diffusion. The innovator 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.<sup>7</sup>

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 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 value that lies below the threshold. All imitators grow at the same rate as the frontier technology with growth only

---

<sup>5</sup> 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](#), p. 24).

<sup>6</sup> [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. Interpretations of our model with, and without, scale effects are discussed in Sect. 6.

<sup>7</sup> [Segerstrom \(1991\)](#) allows monopolistically competitive ex-ante identical firms to choose investment in innovation and imitation. Given the structure of costs and competition in [Segerstrom \(1991\)](#), in contrast to our model, the firms on the technology frontier choose not to invest in innovation and are eventually overtaken by firms who are behind the frontier and invest in innovation.

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.

In Sect. 6, we extend the model to allow for heterogeneity and scale effects. In this setup, countries with the same productivity level may produce different levels of output due to scale effects and differing factor endowments. We show that a non-degenerate stationary distribution still exists. Countries of different scale can nevertheless grow at the same rate precisely because technology diffusion acts as a counterbalance to scale effects, allowing growth at the frontier to pull along the followers.

An alternative approach to modeling technology diffusion is to consider how agents might adopt the technology of other agents spread throughout the productivity distribution, typically adopting technologies much closer to the mean productivity than the frontier. The recent works of [Lucas and Moll \(2013\)](#) and [Perla and Tonetti \(2013\)](#) endogenize technology adoption as an optimal search decision made by agents 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 \(2013\)](#) and [Perla and Tonetti \(2013\)](#) 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. 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.<sup>8</sup> In a recent paper [König et al. \(2012\)](#) obtain rich dynamic results that are similar to ours in that agents optimally choose both to imitate and innovate. In their model innovation is stochastic along a quality ladder while imitation generates stochastic productivity gains through random matching of firms. They characterize the long-run distribution of productivities resulting from innovation and imitation. Our deterministic framework is simpler and features a portfolio choice where agents choose between imitation and innovation. Studying this new mechanism in isolation allows a clear understanding of the incentives to invest in technology diffusion. The endogenous evolution of the frontier generates the key economic incentives for both catch-up as well as fall-back. Both innovation and imitation persist in the limit, and we analytically characterize the full dynamics of catch-up and fall-back, together with the limiting distribution of relative productivities.

---

<sup>8</sup> Other significant works that address the size distribution of innovations and/or productivity across sectors/firms are in [Aghion and Howitt \(1998\)](#), [Thompson \(2001\)](#) and [Laincz \(2009\)](#). [Aghion and Howitt \(1998\)](#) show that under some conditions there exists a stationary distribution of productivity levels in an otherwise standard quality-ladder endogenous growth model. [Thompson \(2001\)](#) derives in closed-form the steady state distribution of firms size in a model that builds on [Thompson and Waldo \(1994\)](#). [Laincz \(2009\)](#) studies an ambitious computational model that integrates into growth economics the approach of [Ericson and Pakes \(1995\)](#) to industry dynamics.

## 2 The model

Agents are heterogeneous over their productivity level,  $z$ , and choose to invest in improving productivity through imitation and innovation. Productivity growth depends linearly on optimally chosen expenditures on innovation,  $\gamma(t)$ , and adoption,  $s(t)$ . The benefits of innovation,  $\sigma\gamma(t)$ , the costs of innovation,  $\gamma(t)z(t)$ , the benefits of diffusion,  $\tilde{D}(t, z)s(t)$ , and the costs of diffusion,  $s(t)z(t)$ , are flow variables in continuous time. Consumption is  $Bz(t) - s(t)z(t) - \gamma(t)z(t)$ , utility is logarithmic, and future consumption is discounted at rate  $r$ . The distribution of productivities at time  $t$  is  $\tilde{\Phi}(t, z)$  on  $(0, F(t)]$ , where  $F(t)$  is the frontier. This distribution fully characterizes the aggregate state of the economy. 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:

$$\max_{s(t), \gamma(t)} \int_0^{\infty} e^{-rt} \ln (Bz(t) - s(t)z(t) - \gamma(t)z(t)) dt, \quad (1)$$

$$s.t. \quad \frac{\dot{z}(t)}{z(t)} = \sigma\gamma(t) + \tilde{D}(t, z)s(t), \quad (2a)$$

$$s(t) \geq 0, \quad \gamma(t) \geq 0. \quad (2b)$$

When convenient for exposition, we will alter notation to drop explicit dependence of functions on  $t$ .

**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),  $D$ , 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 no other statistics of the productivity distribution, the evolution of each initial productivity relative to the frontier can be studied independently.<sup>9</sup> This represents agents targeting their imitation efforts towards the best technologies, and in this sense it is the choices agents make that delivers the tractability.

All agents choose their expenditures in innovation and in technology adoption optimally, both of which promote the rate of productivity growth. Spending  $\gamma z$  on innovation produces growth of  $z$  at rate  $\gamma\sigma$ . For an investment of  $s z$  in imitation, the agent produces growth of  $z$  at 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

$$\tilde{D}(t, z) = \frac{c}{m} \left( 1 - \left( \frac{z}{F(t)} \right)^m \right), \quad (3)$$

<sup>9</sup> Using the frontier (the maximum statistic) to summarize the distribution provides a great deal of analytical tractability. [Perla and Tonetti \(2013\)](#) and [Lucas and Moll \(2013\)](#) use random search models to focus on the feedback relationship between the entire endogenously determined productivity distribution and optimal technology adoption decisions. [Alvarez et al. \(2012\)](#) 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.

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.<sup>10</sup> Letting  $m \rightarrow 0$  yields, after using L’Hopital’s rule, the Gompertz diffusion model:

$$\tilde{D}(t, z) = -c \ln \left( \frac{z}{F(t)} \right).$$

Note that  $\tilde{D}(t, F(t)) = 0$  for all  $t, c, m$ . Thus, the returns to imitation at the frontier are always 0.

**The Hamiltonian** 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 the  $\tilde{D}$  function yields

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

The first order necessary conditions of this Hamiltonian are augmented with complementary slackness conditions to ensure  $s \geq 0$  and  $\gamma \geq 0$ .  $H$  is concave in  $s$  and  $\gamma$ , but the full second order conditions are addressed in Sect. 4. Note that given log utility, it will not be optimal to set  $(B - s - \gamma) = 0$ .

For the moment the constant  $B$  is common across agents. In general,  $B$  may differ across agents, if for example  $B$  represents size and scale effects or other factors affecting output productivity of a country. In Sect. 6 we generalize our analysis by allowing  $B$  to be heterogeneous.

### 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

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

with

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

and

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

In Appendix 1.1 it is shown that optimal investment in innovation is

$$\gamma = B - \frac{r}{\sigma}. \tag{8}$$

<sup>10</sup> Note that in the logistic case ( $m = 1$ ) the increase in productivity  $\dot{z}$  from diffusion,  $\tilde{D}(t, z)z = cz \left( 1 - \left( \frac{z}{F(t)} \right) \right)$ , is zero when  $z = 0$  or when  $z = F(t)$ , and it is maximized at  $z = F(t)/2$ . By contrast in the Nelson–Phelps case ( $m = -1$ ),  $\tilde{D}(t, z)z = cz \left( \left( \frac{F(t)}{z} \right) - 1 \right) = c(F(t) - z)$ , the larger the distance to the frontier, the larger is the diffusion flow.

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

From Eq. (8), under Assumption 1 it is clear that the optimal investment in innovation at the frontier is strictly positive. 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 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  $z < F(t)$ , 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  $g = \sigma B - r$ . 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. Any catch-up function that can be written as a function only of the relative productivity to the frontier will enable this change of variables. Using the ratio of productivity to the frontier,

$$x(t) \equiv \frac{z}{F(t)}, \quad (9)$$

examination of the specific catch-up function in Eq. (3) shows that there is no dependence on time except through  $x(t)$

$$\tilde{D}(t, x(t)F(t)) = \frac{c}{m} \left( 1 - \left( \frac{x(t)F(t)}{F(t)} \right)^m \right) = \frac{c}{m} (1 - x(t)^m), \quad (10)$$

$$\equiv D(x(t)). \quad (11)$$

The growth of  $x(t)$  is

$$\frac{\dot{x}(t)}{x(t)} = \sigma\gamma(t) + D(x(t))s(t) - \frac{\dot{F}(t)}{F(t)}. \quad (12)$$

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

$$\frac{\dot{x}(t)}{x(t)} = \sigma\gamma(t) + D(x(t))s(t) - g. \quad (13)$$

Transforming the objective function in Eq. (1) using the change of variables yields

$$\int_0^{\infty} e^{-rt} (\ln(x(t)) + \ln(B - s(t) - \gamma(t))) dt + \int_0^{\infty} e^{-rt} \ln(F(t)) dt.$$

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



$$\max_{s(t), \gamma(t)} \int_0^\infty e^{-rt} (\ln x(t) + \ln (B - s(t) - \gamma(t))) dt + \frac{g + r \ln(F(0))}{r^2}, \tag{14}$$

$$\frac{\dot{x}(t)}{x(t)} = \sigma \gamma(t) + D(x(t))s(t) - g, \tag{15a}$$

$$s(t) \geq 0, \gamma(t) \geq 0. \tag{15b}$$

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

$$H(x, s, \gamma, \lambda) = \ln(x) + \ln(B - s - \gamma) + \lambda x \left( \sigma \gamma + \frac{c}{m}(1 - x^m)s - g \right). \tag{16}$$

The Lagrange multiplier on the law of motion in  $x$  is  $\lambda$ . As previously mentioned, this problem is now autonomous when compared to Eq. (4).

**Marginal productivity of innovation and imitation** Since Eq. (15a) is linear in both  $s$  and  $\gamma$ , the marginal productivity for an investment in either  $s$  or  $\gamma$  is constant. That is, an investment in  $\gamma$  has a marginal productivity of innovation  $\sigma$ . Given a fixed value of  $x$ , an investment in  $s$  has a marginal productivity of imitation  $D(x)$ . Since  $\frac{\partial D(x)}{\partial x} < 0$ , 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 an increase in  $x$  has on the marginal productivity of imitation.

#### 4.1 First order necessary conditions

The first order necessary conditions can be derived using the Hamiltonian in  $x$  space (Eq. 16).

$$H = \ln(x) + \ln(B - s - \gamma) + \lambda x (\sigma \gamma + D(x)s - g). \tag{17}$$

The Euler equation is

$$\begin{aligned} \dot{\lambda} &= -\frac{\partial H}{\partial x} + r\lambda, \\ &= -\frac{1}{x} - \lambda (\sigma \gamma + sD(x) + sx D'(x) - g) + \lambda r, \end{aligned} \tag{18}$$

$$= -\frac{1}{x} - \lambda \frac{\dot{x}}{x} - \lambda sx D'(x) + \lambda r. \tag{19}$$

Using the catch-up function defined in Eq. (10) yields

$$= -\frac{1}{x} - \lambda \frac{\dot{x}}{x} + \lambda csx^m + \lambda r. \tag{20}$$

Note that  $D'(x) = -cx^{m-1}$  is embedded in Eq. (18), 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

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

$$\frac{1}{B - s - \gamma} \geq \lambda x D(x), \tag{21}$$

$$\frac{1}{B - s - \gamma} \geq \lambda x \frac{c}{m} (1 - x^m). \tag{22}$$

The first-order conditions for  $\gamma$ , including the complementarity condition, are

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

$$\frac{1}{B - s - \gamma} \geq \lambda x \sigma. \tag{23}$$

The transversality condition is

$$\lim_{t \rightarrow \infty} e^{-rt} \lambda(t)x(t) = 0. \tag{24}$$

It can be seen from the the first-order conditions in Eqs. (21) and (23) that an agent considers the marginal productivity of innovation or imitation when optimizing its investment portfolio, where  $\lambda$  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 a bar over a variable denote steady state values (e.g.  $\bar{s}$  and  $\bar{x}$ ). The Euler equation (Eq. 20) can be rewritten as

$$\begin{aligned} \dot{\lambda}x &= -1 - \lambda\dot{x} - \lambda xs D(x) + \lambda xr, \\ \dot{\mu} &= -1 + \mu(r - sx D'(x)) = -1 + \mu(r + csx^m). \end{aligned} \tag{25}$$

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

$$\bar{\mu} = \frac{1}{r - \bar{s}\bar{x} D'(\bar{x})} = \frac{1}{r + c\bar{s}\bar{x}^m}. \tag{26}$$

#### 4.2 Solution to follower’s problem

The follower’s problem can be solved by analyzing all cases of the complementary slackness conditions for  $s$  and  $\gamma$ .

##### 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  $\gamma$  to be interior, the first order conditions, Eqs. (22) and (23), must be equal.

$$\begin{aligned} \lambda x^* \sigma &= \lambda x^* \frac{c}{m} (1 - x^{*m}), \\ \sigma &= D(x^*), \end{aligned} \tag{27}$$

$$\begin{aligned} \sigma &= \frac{c}{m} (1 - x^{*m}), \\ x^* &\equiv \left(1 - \frac{\sigma m}{c}\right)^{1/m}. \end{aligned} \tag{28}$$

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  $\sigma$ ,  $m$ , and  $c$ . At any  $x$  other than  $x^*$  both Eqs. (22) and (23) cannot hold with equality simultaneously, and therefore both  $s$  and  $\gamma$  cannot be simultaneously strictly positive except at  $x^*$ .<sup>11</sup>

When making the portfolio choice on the extensive margin, Eq. (27) 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:*

$$x^* = \left(1 - \frac{\sigma m}{c}\right)^{1/m}, \tag{29}$$

$$x < x^* \implies s(x) \geq 0, \quad \gamma(x) = 0, \tag{30}$$

$$x > x^* \implies s(x) = 0, \quad \gamma(x) > 0. \tag{31}$$

*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  $\gamma$  is strictly positive in the innovation region is proved in Sect. 4.2.2.<sup>12</sup> □

When the marginal productivity of innovation is high enough to violate Assumption 2, it is possible for  $x^*$  to equal 0, and no imitation will take place, as can be seen from Eq. (29). 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. Using Eq. (23) and  $s = 0$  yields:

$$\frac{1}{B - \gamma} = \lambda x \sigma. \tag{32}$$

<sup>11</sup> 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 Sect. 4.2.2.

<sup>12</sup> Since  $x$  and  $\lambda$  are continuous on an optimal path, at the threshold  $x^* = (1 - \frac{\sigma m}{c})^{1/m}$ , 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,

$$B - s = \left(\lambda x \frac{c}{m} (1 - x^m)\right)^{-1} = \left(\frac{1}{r} \frac{c}{m} \left(\frac{m\sigma}{c}\right)\right)^{-1},$$

$$s = B - \frac{r}{\sigma} > 0,$$

$$\frac{\dot{x}}{x} = s\sigma - g = B\sigma - r - g = 0.$$

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

Assuming that an agent with  $s = 0$  will never enter the region  $x < x^*$  where  $s \geq 0$ , as is later verified, Eq. (57) can be substituted into Eq. (32) to find

$$\gamma = \frac{B\sigma - r}{\sigma}. \tag{33}$$

This is the same level of investment as that of the frontier agent, given in Eq. (8). Plugging this into the LOM for  $x$  shows

$$\frac{\dot{x}}{x} = \sigma \frac{B\sigma - r}{\sigma} - g = 0. \tag{34}$$

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 Eq. (26) with  $s = 0$  yields

$$\lim_{t \rightarrow \infty} e^{-rt} \lambda(t)x(t) = \frac{1}{r} \lim_{t \rightarrow \infty} e^{-rt} = 0. \tag{35}$$

#### 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  $m$ :  $q(t) \equiv x(t)^m$ . In Appendix 1.2 it is shown that the joint dynamics of productivity and imitation investment are characterized by

$$\begin{aligned} \frac{d}{dt} \begin{pmatrix} s(t) \\ q(t) \end{pmatrix} &\equiv \Psi(s(t), q(t)) \\ &= \begin{pmatrix} \frac{(B-s(t))(Bc-mr+q(t)(-2Bc+m(-gm+r)+Bcq(t))-c(-1+q(t))^2s(t))}{mq(t) \left(\frac{c}{m}(1-q(t))s(t) - g\right)} \end{pmatrix}. \end{aligned} \tag{36}$$

**Stationary solution for  $s \geq 0, \gamma = 0$**  Define the stationary points of this system as  $\bar{s}$  and  $\bar{q}$  (or equivalently  $\bar{x}$ ).<sup>13</sup> In Appendix 1.2 it is shown that the unique stationary imitation policy,  $\bar{s}$ , and the unique interior stationary relative productivity level,  $\bar{x} \in (0, x^*)$ , are

$$\bar{s} = \frac{B\frac{1}{m}(1-\bar{q}) - \frac{r}{c}}{\frac{1}{m}(1-\bar{q}) + \bar{q}}, \tag{37}$$

$$\bar{x} = \left[ \frac{1}{2} \left( \left( 2 - \frac{m\sigma}{c} + \frac{m^2}{c} \left( \sigma - \frac{r}{B} \right) \right) - \text{sign}(m) \left( \left( 2 - \frac{m\sigma}{c} + \frac{m^2}{c} \left( \sigma - \frac{r}{B} \right) \right)^2 - 4 \left( 1 - \frac{m\sigma}{c} \right) \right)^{1/2} \right) \right]^{1/m}. \tag{38}$$

<sup>13</sup> For a general strictly monotone  $D(x)$  function, from Eq. (27),  $x^* = D^{-1}(\sigma)$ . Using Eqs. (26) and (21) and solving for the optimal control in the stationary equilibrium,  $\bar{s} = \frac{BD(\bar{x})-r}{D(\bar{x})-\bar{x}D'(\bar{x})}$ . Assuming the existence of a stationary equilibrium with imitation and using the law of motion in Eq. (15a) gives the implicit equation defining  $\bar{x}$ ,

$$B\sigma - r = \frac{BD(\bar{x}) - r}{1 - \bar{x}D'(\bar{x})/D(\bar{x})}.$$

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.

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 Eq. (26) yields

$$\lim_{t \rightarrow \infty} e^{-rt} \lambda(t)x(t) = \lim_{t \rightarrow \infty} e^{-rt} \bar{\lambda} \bar{x} = \frac{1}{r + c\bar{s}\bar{x}^m} \lim_{t \rightarrow \infty} e^{-rt} = 0. \tag{39}$$

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 4 in Supplementary Material for details).<sup>14</sup>

**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 Sect. 4.2.2, for  $x > x^*$  all agents choose to innovate at the same rate and grow at the same rate. As proven in Sect. 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 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.<sup>15</sup>

### 5 The dynamic solution

Proposition 1 characterizes the main results of the paper including the existence, uniqueness, and stability of the stationary solution.

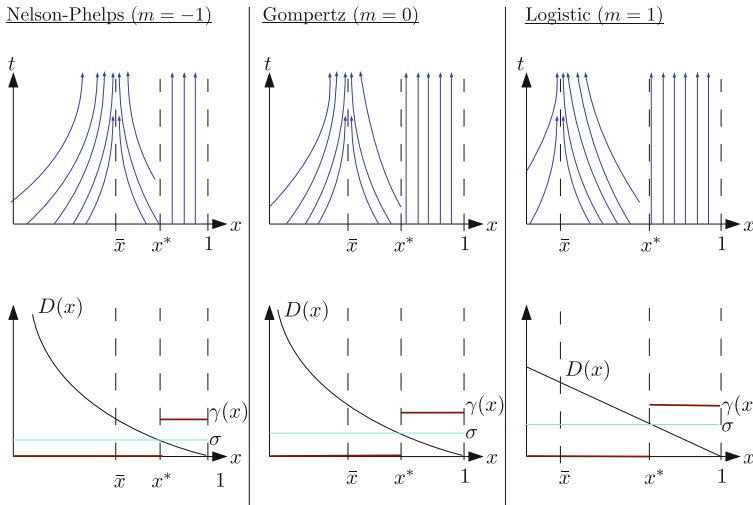
**Proposition 1** *Let Assumptions 1, 2, and 3 hold. For an arbitrary initial distribution  $\Phi(0, x)$ :*

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

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

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  $m = 0$ .

<sup>15</sup> The productivity improvement technology does not generate mixing, 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.



**Fig. 1** Equilibrium convergence and innovation policies

**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 Sect. 4 and the conditions in Appendix 2.

Denoting the cases where  $m > 0$  as logistic diffusions and the cases where  $m < 0$  as Nelson–Phelps diffusions, Fig. 1 provides examples illustrating system dynamics and Proposition 1. 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,  $\sigma$ , is greater than the marginal productivity of imitation,  $D(x)$ , the agent chooses to invest entirely in innovation,  $\gamma$ .<sup>16</sup>

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  $\Phi(0, x)$ ,

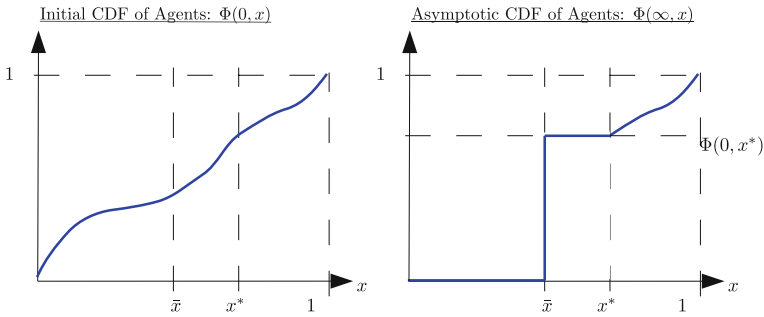
$$\Phi(\infty, x) = \begin{cases} 0 & \text{for } 0 \leq x < \bar{x} \\ \Phi(0, x^*) & \text{for } \bar{x} \leq x < x^* \\ \Phi(0, x) & \text{for } x^* \leq x \leq 1. \end{cases}$$

<sup>16</sup> 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  $\sigma = .9$ ;  $c = 1$ ;  $B = 2$ ;  $r = 0.05$ .



**Fig. 2** Balanced growth path distribution

*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 Fig. 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  $F(0)$ . 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 for some parameter values in the logistic case, if its 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 investment expenditures in imitation are endogenized, this finding is reversed. Agents optimally invest in technology diffusion such that they converge to a positive BGP productivity ratio,  $\bar{x}$ , and then grow at the same rate as the leader.

This can be seen directly from our Eq. (15a), which is equivalent to that of Eq. (4.1) in Benhabib and Spiegel (2005), expressed in our current notation. In Benhabib and Spiegel (2005),  $\gamma(t)$  and  $s(t)$  are treated as constants that can depend on country specific fixed human capital levels, not as choice variables. When  $\gamma(t)$  and  $s(t)$  are taken as constants, the solution to the law of motion equation, Eq. (4.4) in Benhabib and Spiegel (2005), is:

$$\lim_{t \rightarrow \infty} x(t) = \begin{cases} \left(1 + \frac{m(\sigma\gamma - g)}{cs}\right)^{\frac{1}{m}} & \text{if } cs - m(\sigma\gamma - g) > 0 \text{ or } m < 0 \\ 0 & \text{if } cs - m(\sigma\gamma - g) < 0 \text{ and } m > 0. \end{cases} \quad (40)$$

When  $\gamma(t)$  and  $s(t)$  are chosen optimally, the second case of  $\bar{x} = 0$  is ruled out, as shown in Eq. (38) and Appendix 2.1.

However we will show in Sect. 6 that if we allow the output productivity parameters  $B$  to differ across countries and if  $m > 0$ , the ratio of the productivity level of a country to that of the frontier may go to zero if its  $B$  is relatively small, i.e.,  $\lim_{t \rightarrow \infty} x(t) = 0$  (see Proposition 3). For similar reasons as in Benhabib and Spiegel (2005), this depends crucially on the curvature of the catch-up function. This can only happen for  $m > 0$ , including the logistic case  $m = 1$ , and may happen even if investment rates are optimally chosen, because the diffusion rates for  $m > 0$  eventually decrease as the distance to the frontier increases.

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 catch-up and fall-back forces that lead to the BGP.

From the Euler equation in (18),

$$r = \frac{\dot{\lambda}}{\lambda} + \frac{1}{x} + (\sigma\gamma + sD(x) - g) + sxD'(x). \tag{41}$$

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 the 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  $D'(x) < 0$ . 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 price  $\lambda(t)$ . The Euler equation at steady state can be arranged as

$$\bar{\mu} = \frac{1}{r - \bar{s}\bar{x}D'(\bar{x})} = \frac{1}{r + c\bar{s}\bar{x}^m} \leq \frac{1}{r}, \tag{42}$$

$$\bar{\lambda} = \frac{1}{\bar{x}} \left( \frac{1}{r - \bar{s}\bar{x}D'(\bar{x})} \right). \tag{43}$$

Equations (42) and (43) 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 Eq. 42). 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 Heterogeneous agents and scale effects

Thus far, agents in the model have only differed in their initial conditions. However, we know from previous growth research that many theories of growth exhibit scale effects and productivity differences when country size is taken into account. Output is bounded above by  $Bz$  and  $z$  is always interpreted as per-capita productivity. Thus, one possible interpretation is that  $B$  is related to population size or other fixed factors.<sup>17</sup> In such a case, as analyzed in Sect. 6.1, there are scale effects with heterogeneous  $B$ . If there were no imitation in the model, then all countries would innovate, and the largest country (with the highest  $B$ ) would grow the fastest. The model would then exhibit strong scale effects and the stationary distribution would collapse such that the leader would grow large and all other countries would disappear in the limit. However, with technology diffusion we still obtain balanced growth path equilibria with non-degenerate stationary distributions. This is because the diffusion of technology from the frontier will have a locomotive effect and pull the followers along. Thus, technology diffusion can act as a counterbalance to scale effects, and their interaction is essential to determining the limiting dynamics of the model.

We should note that even with heterogeneous  $B$ , the country with the largest  $B$  will not necessarily be the leader in the stationary distribution. Furthermore, with heterogeneous  $B$ , the stationary distribution is not ergodic and initial conditions matter. For example, if a large country like China started with a low productivity level, it may choose to never overtake the US to become the global leader, even if the US is smaller in population. The catch-up gains from technology diffusion can be large enough to prevent the scale effect from dominating the equilibrium stationary distribution.

Additionally, with heterogeneous  $B$ , the sign of  $m$  determines if countries can be left behind. When technology diffusion is efficient, as in the Nelson–Phelps case of  $m = -1$ , no country, even those with small  $B$ , will shrink relative to the leader in the limit. However, when technology diffusion is inefficient, countries with  $B$  below a threshold can be left behind and shrink in relative productivity in the limit. Analysis and results on heterogeneous  $B$  are presented in the next section.

### 6.1 Stationary solution with heterogeneous $B$

Although the stationary distribution depends on initial conditions, Proposition 3 characterizes properties of potential stationary distributions of BGP equilibria when agents have heterogeneous  $B$ .

Assume that the agents have idiosyncratic  $B$ , s.t.  $B_i > 0 \quad \forall i$ . Define the  $B$  of the frontier agent at time  $t$  as  $\hat{B}(t) \equiv \{B_i | z_i(t) = \max_{i'} \{z_{i'}(t)\}\}$ .

**Proposition 3** *The stationary distribution depends on the initial condition,  $z_i(0) \forall i$ . Given an initial condition, for  $m \neq 0$ , the following properties hold for the stationary solution,  $x^s(B)$ :*

- (1) *Eventually, the frontier agent’s type is constant. That is, there exists a  $\hat{t}$  such that for all  $t > \hat{t}$ ,  $\hat{B}(t) = \hat{B}$ .*

<sup>17</sup> We may alternatively consider  $B$  as an intrinsic parameter of per capita productivity. In that case,  $s$  and  $\gamma$  would be interpreted as per capita investment in intensity in imitation and innovation. If every agent within a country is identical, then the number of agents within a country would not affect productivity dynamics and there would be no size or scale effects. There may, however, be other factors that affect per capita productivity and induce heterogeneous  $B$  across countries. Another alternative interpretation is that scale effects may be avoided by subdividing countries with different factor endowments into economic regions that equalize relative factor endowments across regions.

(2) Given a  $\hat{B}$ , the stationary equilibrium as a function of  $B$  is

$$x^s(B) = \begin{cases} 0 & 0 \leq B \leq \underline{B} \\ \bar{x}(B) & \underline{B} < B < \hat{B} \\ \in \bar{x}(\hat{B}) \cup [x^*, 1] & B = \hat{B} \\ \bar{x}(B) & \hat{B} < B < \bar{B}, \end{cases}$$

where  $\bar{x}(B) \in \left(0, 1 - \frac{m\sigma \hat{B}}{c}\right)$ , with

$$\begin{aligned} \bar{x}(B) = & \left[ \frac{1}{2} \left( \left( 2 - \frac{m\sigma \hat{B}}{c} + \frac{m^2}{c} \left( \sigma \frac{\hat{B}}{B} - \frac{r}{B} \right) \right) - \text{sign}(m) \right. \right. \\ & \left. \left. \times \left( \left( 2 - \frac{m\sigma \hat{B}}{c} + \frac{m^2}{c} \left( \sigma \frac{\hat{B}}{B} - \frac{r}{B} \right) \right)^2 - 4 \left( 1 - \frac{m\sigma \hat{B}}{c} \right) \right)^{1/2} \right) \right]^{1/m}, \end{aligned} \tag{44}$$

$$\underline{B} = \max \left\{ \hat{B} \frac{m\sigma}{c}, 0 \right\}, \tag{45}$$

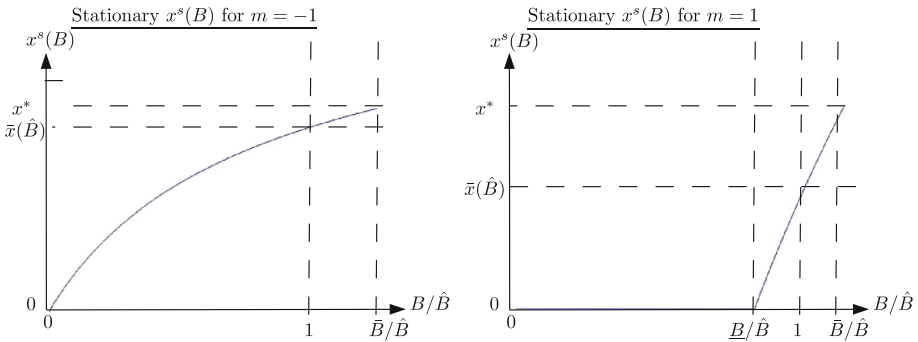
$$\bar{B} = \hat{B} \left( 1 - m + \frac{c}{\sigma} \right) + \frac{r(-c + m\sigma)}{\sigma^2}. \tag{46}$$

- (3) There may exist equilibria where the frontier agent does not have the highest  $B$  (i.e.,  $\hat{B} < \max \{B_i\}$ ) and  $x^s(B_i) < x^*$  for  $B_i > \hat{B}$ .
- (4) The only agents who choose to innovate are those with  $B = \hat{B}$  (i.e.,  $\gamma(B) > 0 \implies B = \hat{B}$ ). Alternatively,  $x^s(B) \in (x^*, 1] \implies B = \hat{B}$ .
- (5) If  $m < 0$ , then agents cannot be left behind (i.e.,  $\underline{B} = 0$ ).
- (6) The maximum  $B$  of the imitator is always finite (i.e.,  $\bar{B} < \bar{B} < \infty$ ).
- (7) If at time zero,  $x(\max \{B_i\}) > x^*$ , then the agent with the highest  $B_i$  eventually becomes the frontier agent.

*Proof* See Appendix 3.

The essential points in Proposition 3 are illustrated in Fig. 3. In the stationary equilibrium there is a unique  $B$ , denoted as  $\hat{B}$ , at the frontier. Figure 3 graphs the stationary  $x^s(B)$  for different  $B$ 's relative to the equilibrium frontier  $\hat{B}$ . Since all agents with  $B \neq \hat{B}$  have  $x^s(B) < x^*$ , only agents with  $B = \hat{B}$  choose to innovate, and all other agents choose to invest in technology diffusion. Note that this is true even for  $B > \hat{B}$ , illustrating that the largest  $B$  is not always the leader. There is, however, a limit as to how much larger a  $B$  can be compared to the frontier  $\hat{B}$ , given by  $\bar{B}$ .<sup>18</sup> Figure 3 also depicts the importance of the sign of  $m$  in determining whether agents can be left behind. When  $m < 0$ , imitation is very productive, so there is never a region of agents whose productivity goes to zero relative to the frontier agents in the limit. However, when  $m > 0$  imitation is not very productive,

<sup>18</sup> If  $B > \bar{B}/\hat{B}$ , then imitating is not optimal and the agent would want to innovate. However, if the agent innovates, it does so at a faster rate since  $B > \hat{B}$ , and hence they become the frontier, invalidating equilibria with  $B > \bar{B}(\hat{B})$ .



**Fig. 3** Stationary  $x$  for heterogeneous  $B$  relative to  $\hat{B}$

and agents can be left behind. There exists a lower threshold,  $\underline{B}$ , such that any agent with  $B < \underline{B}$  would choose to imitate and would still grow, but would be at such a disadvantage that ultimately it would fall behind in the limit.<sup>19</sup>

### 7 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 returns to the homogeneous  $B$  case to study how the balanced growth path equilibrium changes as parameters modulating the productivity of investment— $\sigma$ ,  $m$ , and  $c$ —are varied.

#### 7.1 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 Eq. (29),

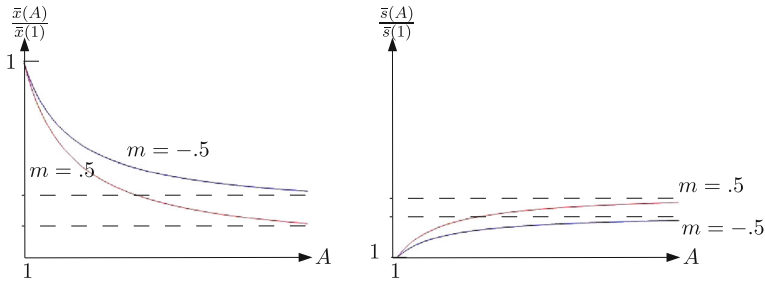
$$x^* = (1 - m\theta)^{1/m}, \tag{47}$$

the threshold for imitation versus innovation depends on the ratio  $\theta$  rather than on the levels of  $\sigma$  and  $c$  independently. This is because  $x^*$  is determined by the ratio of instantaneous marginal productivities that are linear in  $\sigma$  and  $c$ , as captured in Eq. (27). 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  $\sigma$  and  $c$  that keep  $\theta$  constant may act like a Hicks-neutral technical change, since they keep this measure of relative productivity constant.

However, Eq. (38) shows the lack of neutrality on  $\bar{x}$ .

$$\bar{x} = \left[ \frac{1}{2} \left( 2 + (m^2 - m)\theta - m^2 \frac{r}{cB} - \text{sign}(m) \right) \times \left( \left( 2 + (m^2 - m)\theta - m^2 \frac{r}{cB} \right)^2 - 4(1 - m\theta) \right)^{1/2} \right]^{1/m}. \tag{48}$$

<sup>19</sup> This result is similar to the force in Benhabib and Spiegel (2005).



**Fig. 4** Equilibrium with Hicks-neutral technical change

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 4** *An increase in the TFP of growth technologies*

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

*Proof*  $\frac{dx^*}{dA} = 0$  follows from the definition of  $x^*$ . For  $\frac{d\bar{x}}{dA} < 0$  and  $\frac{d\bar{s}}{dA} > 0$  (see Appendix 5 in Supplementary Material).

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  $A \rightarrow \infty$ . Clearly  $x^*$  is unchanged. For  $\bar{x}$ , the terms with a  $c$  in Eq. (48) go to zero. Hence,

$$\lim_{A \rightarrow \infty} \bar{x}(A) = \left[ \frac{1}{2} \left( 2 + (m^2 - m)\bar{\theta} - \text{sign}(m) \left( (2 + (m^2 - m)\bar{\theta})^2 - 4(1 - m\bar{\theta}) \right)^{1/2} \right) \right]^{1/m} > 0. \tag{49}$$

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  $\infty$ , where both  $x^*$  and  $\bar{x}$  converge to 1, as discussed in Sect. 7.4.

Figure 4 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  $A = 1$ .<sup>20</sup>

<sup>20</sup> The parameters used in this simulation are  $\bar{\sigma} = 1.25$ ,  $\bar{c} = 1$ ,  $B = 2$  and  $r = 0.05$ . Using these parameters, Fig. 4 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.

### 7.2 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. The notation  $D(x; m)$  is intended to emphasize this dependence.

As discussed in Sect. 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.  $\sigma$ ). 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.<sup>21</sup> 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 5.

**Proposition 5** *As curvature in the catch-up function changes to make imitation less efficient, the innovator threshold decreases. That is  $\frac{dx^*}{dm} < 0$ .*

*Proof* Taking the derivative of  $x^*$  from Eq. (29) with respect to  $m$  yields

$$\frac{dx^*}{dm} = \frac{\left(1 - \frac{m\sigma}{c}\right)^{\frac{1}{m}} \left(m\sigma + (c - m\sigma) \log\left(1 - \frac{m\sigma}{c}\right)\right)}{m^2(-c + m\sigma)}.$$

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

$$m\sigma + (c - m\sigma) \ln\left(1 - \frac{m\sigma}{c}\right) = c \left[\frac{m\sigma}{c} + \left(1 - \frac{m\sigma}{c}\right) \ln\left(1 - \frac{m\sigma}{c}\right)\right] > 0. \quad (50)$$

This term as a function of  $\frac{m\sigma}{c}$  achieves a unique global minimum of 0 at  $\frac{\sigma m}{c} = 0$ .<sup>22</sup> Since Assumption 2 restricts  $\frac{\sigma m}{c} > 0$ , this term is strictly positive, and thus  $\frac{dx^*}{dm} < 0$ .<sup>23</sup>  $\square$

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.

Intuitively, decreasing the efficiency of technology diffusion by increasing  $m$  also decreases the incentive to fall-back. Formally,  $\frac{d}{dm}(D'(x; m)) = \frac{\partial^2 D(x; m)}{\partial x \partial m} = -cx^{-1+m} \ln(x) > 0$ . 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 6** *As curvature in the catch-up function changes to make imitation less efficient, the BGP productivity ratio of imitators decreases if  $\frac{m}{q} \frac{d\bar{q}}{dm} < \ln \bar{q}$ . That is  $\frac{d\bar{x}}{dm} < 0$ .*

*Proof* See Appendix 6.2 (Supplementary Material). In Appendix 6.1 (Supplementary Material),  $\frac{d\bar{q}}{dm} < 0$  is shown using just Assumptions 1, 2, and 3.

<sup>21</sup> 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,  $\frac{d\left(\frac{c}{m}(1-x^m)\right)}{dm} = \frac{c}{m^2}(x^m - 1 - mx^m \ln x) < \frac{c}{m^2}(x^m - 1 + (1 - x^m)) = 0$ .

<sup>22</sup> The first derivative of the term is  $\ln\left(1 - \frac{\sigma m}{c}\right)$  and the second derivative is positive:  $\frac{1}{1 - \frac{\sigma m}{c}} > 0$  for  $1 - \frac{\sigma m}{c} > 0$ .

<sup>23</sup> Note:  $\lim_{m \rightarrow 0^+} \frac{dx^*}{dm} = \lim_{m \rightarrow 0^-} \frac{dx^*}{dm} = \frac{-\sigma^2}{2c^2} e^{-\frac{\sigma}{c}} < 0$ .

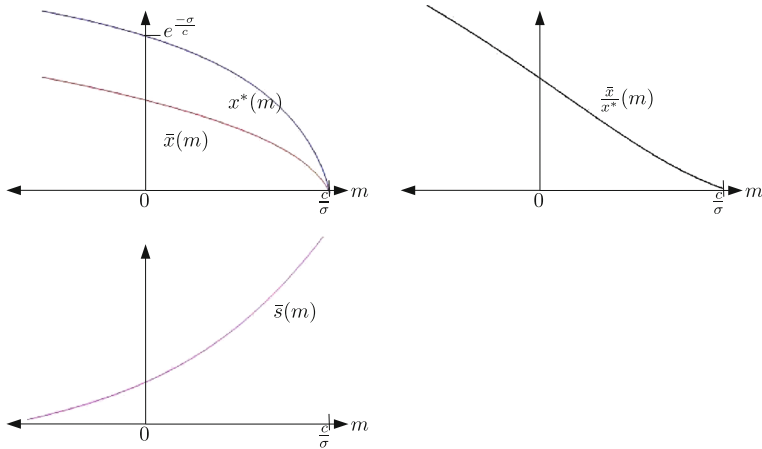


Fig. 5 Equilibrium indexed by  $m$

**Proposition 7** For the parameter restrictions in Appendix 6.3 (Supplementary Material), as curvature in the catch-up function changes to make imitation less efficient, the BGP level of expenditure on imitation increases. That is  $\frac{d\bar{s}}{dm} > 0$ .

*Proof* See Appendix 6.3 (Supplementary Material).

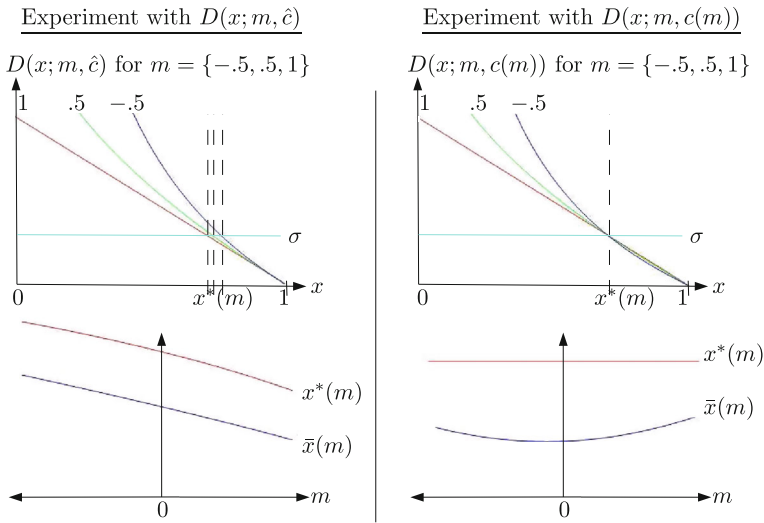
The parameter restrictions in Propositions 6 and 7 are in addition to those in Assumptions 1, 2, and 3. However, in all numerical exercises we have examined, Assumptions 1, 2, and 3 alone have ensured  $\frac{d\bar{x}}{dm} < 0$  and  $\frac{d\bar{s}}{dm} > 0$ .

Figure 5 illustrates the comparative dynamics of the economy with respect to changes in the efficiency of technology diffusion,  $m$ . Proposition 5 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 6 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.

### 7.3 Does less efficient diffusion technology always lower $\bar{x}$ ?

Proposition 6 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  $\bar{x}$ ?



**Fig. 6** Non-monotonicity of  $\bar{x}$

To show that such a non-monotonicity can occur, Fig. 6 conducts a variation of the experiment described in Sect. 7.2. 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 6 compares the results of changing both  $c$  and  $m$  to those of Sect. 7.2 where only  $m$  changes.<sup>24</sup> The left column shows results for an experiment where  $c$  is held constant at  $\hat{c}$  and only  $m$  varies—similar to Fig. 5. 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  $\bar{x}$ .

### 7.4 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  $x$ .

<sup>24</sup> 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.

From Eq. (72) 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.<sup>25</sup>

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 \(2013\)](#) or [Lucas and Moll \(2013\)](#). 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 \(2013\)](#) or [Lucas and Moll \(2013\)](#).

## 8 Conclusion

In our baseline model, 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. When agents with the same productivity differ in their production due to scale effects or factor endowments, there still exists a non-degenerate stationary distribution. Agents of different scale can nevertheless grow at the same rate precisely because technology diffusion acts as a counterbalance to scale effects, allowing growth at the frontier to pull along the followers.

**Acknowledgments** We would like to thank Boyan Jovanovic and Tom Sargent for useful feedback.

<sup>25</sup> 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  $c \rightarrow \infty$ . 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.



### Appendix 1: Solving for policy functions

#### 1.1 Optimal investment in innovation by the leader

The leader must choose optimal investment in innovation by solving

$$\tilde{H} = \ln((B - \gamma)z) + \tilde{\lambda}\sigma\gamma z, \tag{51}$$

with

$$\dot{z} = \sigma\gamma z, \tag{52}$$

and

$$\dot{\tilde{\lambda}} = -\frac{\partial \tilde{H}}{\partial z} + r\tilde{\lambda} = -z^{-1} + \tilde{\lambda}(r - \sigma\gamma). \tag{53}$$

Defining  $\tilde{\mu} = \tilde{\lambda}z$  gives

$$\dot{\tilde{\mu}} = \dot{\tilde{\lambda}}z + \dot{z}\tilde{\lambda} = -1 + \tilde{\mu}(r - \sigma\gamma) + \sigma\gamma\tilde{\mu} = -1 + r\tilde{\mu}. \tag{54}$$

The solution is

$$\tilde{\mu}(t) = e^{rt} \left( \tilde{\mu}(0) - \frac{1}{r} \right) + \frac{1}{r}, \tag{55}$$

and the transversality condition,

$$\lim_{t \rightarrow \infty} e^{-rt} \tilde{\mu}(t) = 0, \tag{56}$$

immediately requires

$$\tilde{\mu}(t) = \frac{1}{r}. \tag{57}$$

Maximizing the Hamiltonian with respect to  $\gamma$  yields

$$\frac{-1}{B - \gamma} + \tilde{\lambda}z\sigma \leq 0. \tag{58}$$

Using Eq. (57), optimal investment in innovation satisfies

$$\gamma \geq B - \frac{r}{\sigma}, \tag{59}$$

which holds with equality since  $\gamma > 0$  under Assumption 1. Finally, it is straightforward to show that the maximized Hamiltonian is concave, so that this is indeed an optimal solution. See Appendix 4 (Supplementary Material) for details.

#### 1.2 Optimal investment in imitation

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  $x$  (Eq. 15a) when  $\gamma = 0$ :

$$\frac{\dot{q}}{q} = m \left( \frac{c}{m}(1 - q)s - g \right). \tag{60}$$

The first order necessary condition in Eq. (22), with  $\gamma = 0$ , gives

$$\mu = \frac{m}{c(B - s)(1 - q)}. \tag{61}$$

The law of motion for  $\mu$  in terms of  $q$  is obtained from Eq. (25)

$$\dot{\mu} = -1 + \mu(r + csq). \tag{62}$$

Equations (60), (61), and (62) form a system in the variables  $\mu, q, s$ . Eliminating  $\mu$  via substitution will create a 2x2 system. First, differentiating Eq. (61) provides

$$\dot{\mu} = \frac{m((B - s)\dot{q} + (1 - q)\dot{s})}{c(1 - q)^2(B - s)^2}. \tag{63}$$

Substituting Eqs. (61) and (63) into (62), and rearranging for  $\dot{s}$  gives

$$\dot{s} = \frac{(B - s)((-1 + q)(-Bc + mr + cs + cq(B + (-1 + m)s)) + m\dot{q})}{m(-1 + q)}. \tag{64}$$

Substituting for  $\dot{q}$  from the law of motion in Eq. (60), generates a first order ODE in  $s$  and  $q$ :

$$\dot{s} = \frac{(B - s)(Bc - mr + q(-2Bc + m(-gm + r) + Bcq) - c(-1 + q)^2s)}{m(-1 + q)}. \tag{65}$$

Equations (60) and (65) are a system of 2 nonlinear first-order ODEs in  $q$  and  $s$ . Stacking the equations and defining the vector valued function  $\Psi$  yields

$$\frac{d}{dt} \begin{pmatrix} s(t) \\ q(t) \end{pmatrix} \equiv \Psi(s(t), q(t)) \tag{66}$$

$$= \begin{pmatrix} \frac{(B-s(t))(Bc-mr+q(t)(-2Bc+m(-gm+r)+Bcq(t))-c(-1+q(t))^2s(t))}{m(-1+q(t))} \\ mq(t) \left( \frac{c}{m}(1 - q(t))s(t) - g \right) \end{pmatrix}. \tag{67}$$

To find the stationary solution, evaluate the first order condition in Eq. (61) at the stationary value  $\bar{s}$ ,

$$\bar{s} = B - \frac{1}{c/m\bar{\mu}(1 - \bar{q})}. \tag{68}$$

Substituting for  $\bar{\mu}$  from Eq. (26) and solving for  $\bar{s}$  gives

$$\bar{s} = \frac{B \frac{1}{m}(1 - \bar{q}) - \frac{r}{c}}{\frac{1}{m}(1 - \bar{q}) + \bar{q}}. \tag{69}$$

From Assumption 1 and Lemma 1 it follows that  $\bar{s} > 0$ .<sup>26</sup> To determine the stationary solution for  $q(t)$ ,  $\bar{q}$ , substitute this expression into the stationary law of motion in Eq. (60) to find

$$\frac{\dot{q}}{q} = 0 = m \left( \frac{c}{m}(1 - \bar{q}) \frac{B \frac{1}{m}(1 - \bar{q}) - \frac{r}{c}}{\frac{1}{m}(1 - \bar{q}) + \bar{q}} - g \right). \tag{70}$$

<sup>26</sup> 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  $\frac{1}{m}(1 - \bar{q}) > \frac{\sigma}{c}$ .

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

$$\frac{\sigma B - r}{B} = \left( \frac{\frac{c}{m}(1 - \bar{q}) - \frac{r}{B}}{1 + \bar{q}(m - 1)} \right) (1 - \bar{q}), \tag{71}$$

$$0 = \bar{q}^2 - \left( 2 - \left( \frac{m\sigma}{c} \right) + \frac{m^2}{c} \left( \sigma - \frac{r}{B} \right) \right) \bar{q} + \left( 1 - \frac{m\sigma}{c} \right). \tag{72}$$

The roots of this quadratic are

$$\bar{q}_{1,2} = \frac{1}{2} \left( \left( 2 - \frac{m\sigma}{c} + \frac{m^2}{c} \left( \sigma - \frac{r}{B} \right) \right) \pm \left( \left( 2 - \frac{m\sigma}{c} + \frac{m^2}{c} \left( \sigma - \frac{r}{B} \right) \right)^2 - 4 \left( 1 - \frac{m\sigma}{c} \right) \right)^{1/2} \right). \tag{73}$$

As is proven in Appendix 2.1, the unique interior stationary solution,  $\bar{x} \in (0, x^*)$ , is

$$\bar{x} = \left[ \frac{1}{2} \left( \left( 2 - \frac{m\sigma}{c} + \frac{m^2}{c} \left( \sigma - \frac{r}{B} \right) \right) - \text{sign}(m) \times \left( \left( 2 - \frac{m\sigma}{c} + \frac{m^2}{c} \left( \sigma - \frac{r}{B} \right) \right)^2 - 4 \left( 1 - \frac{m\sigma}{c} \right) \right)^{1/2} \right) \right]^{1/m}. \tag{74}$$

### Appendix 2: Stability and uniqueness

#### 2.1 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 Eq. (72) in the form of  $\bar{q}^2 - a\bar{q} + b$ , where

$$a = 2 - \left( \frac{m\sigma}{c} \right) + \frac{m^2}{c} \left( \sigma - \frac{r}{B} \right) \text{ and } b = 1 - \frac{m\sigma}{c}.$$

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

$$D = \left( \left( 2 - \left( \frac{m\sigma}{c} \right) + \frac{m^2}{c} \left( \sigma - \frac{r}{B} \right) \right) \right)^2 - 4 \left( 1 - \frac{m\sigma}{c} \right) > \left( 2 - \left( \frac{m\sigma}{c} \right) \right)^2 - 4 \left( 1 - \frac{m\sigma}{c} \right), \tag{75}$$

$$= 4 - 4 \left( \frac{m\sigma}{c} \right) + \left( \frac{m\sigma}{c} \right)^2 - 4 + 4 \left( \frac{m\sigma}{c} \right) = \left( \frac{m\sigma}{c} \right)^2 \geq 0. \tag{76}$$

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.,  $a > 1 + b$ :

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

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  $\bar{x} \in (0, x^*)$ .

### 2.2 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 Eq. (36),

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

Taking the Jacobian of this and simplifying yields

$$\begin{aligned}
 J_\Psi(s, q) &= \begin{pmatrix} -2Bc+mr-q(m(1+m)r-B(4c+m^2\sigma)+2Bcq)+2c(-1+q)^2s & (B-s)(-m^2r+B(c+m^2\sigma)+Bc(-2+q)q-c(-1+q)^2s) \\ \frac{m(-1+q)}{-c(-1+q)q} & \frac{m(-1+q)^2}{m(r-B\sigma)+(c-2cq)s} \end{pmatrix}. \tag{77}
 \end{aligned}$$

For the starting point of the homotopy argument, take the limit as  $r \rightarrow \sigma B$ , where  $g \rightarrow 0$ . As proven in Appendix 2.1,  $\bar{q}_1$  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  $r \rightarrow \sigma B$  is

$$\bar{s} = 0, \quad \bar{q} = 1 - \frac{m\sigma}{c}. \tag{78}$$

Substituting these into the the Jacobian in Eq. (77) generates

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

Taking  $r \rightarrow \sigma B$  and simplifying further gives

$$= \begin{pmatrix} B\sigma & (B^2c)/m \\ m\sigma \left( 1 - \frac{m\sigma}{c} \right) & 0 \end{pmatrix}. \tag{79}$$

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

$$\det (J_{\bar{\psi}}) \equiv \det (J_{\Psi}(\bar{s}, \bar{q})) = B^2\sigma(m\sigma - c). \tag{80}$$

The parameter restriction in Assumption 2 ensures that  $(m\sigma - c) < 0$ , and thus  $\det (J_{\bar{\psi}}) < 0$ . 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  $\det (J_{\bar{\psi}}) < 0$  near  $r = B\sigma$ , then  $\det (J_{\bar{\Gamma}}) < 0$  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 any small  $\varepsilon > 0$ . Since  $\det (J_{\bar{\Gamma}_r}) < 0$  near  $r = B\sigma$ , it is negative at  $r = B\sigma - \varepsilon$  for small  $\varepsilon > 0$ .

Let  $\{\bar{y}(r)\} = \{\Gamma_r^{-1}(0)\}$  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 Poincaré–Hopf degree on the boundary  $\partial\mathcal{D}$  of  $\mathcal{D}$  is  $\sum_{y \in \{\bar{y}(r)\}} \text{sign det} (J_{\Gamma_r}(y)) = 1$  or  $-1$ . See Milnor (1965). Since, from Eq. (38), the stationary point  $\bar{x}(r)$ , and therefore  $\bar{y}(r)$ , is always unique for  $r \in S$ , then  $\text{sign det} (J_{\Gamma_r}(\bar{y}(r))) = \text{sign det} J_{\Gamma_r}(\bar{y}(B\sigma - \varepsilon)) = -1$ .

If the roots of  $J_{\Gamma_r}(\bar{y}(r))$  are real, then they must be of opposite sign. Therefore  $\det (J_{\Gamma_r}(\bar{y}(r))) < 0$  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 3: Agents with heterogeneous $B$

Assume that the agents have idiosyncratic  $B$ , s.t.  $B_i > 0 \ \forall i$ . Define the  $B$  of the frontier agent at time  $t$  as  $\hat{B}(t) \equiv \{B_i | z_i = \max_{i'} \{z_{i'}\}\}$ . In this Appendix section, we prove Proposition 3.

#### 3.1 Eventually, the frontier agent’s type is constant

To show that there are no oscillations at the frontier in the stationary equilibrium, assume there was an oscillating set of frontier agents with different values of  $B$ . Compare pairwise any two values for  $B$  within this set. In the instance where the higher  $B$  becomes the leader, the first order conditions show that it will continue to innovate at a rate faster than that of a lower  $B$  innovator, and thus never fall behind. By process of pairwise iteration through the set of potential frontier agents, the highest  $B$  within the set of potential frontier agents must eventually become the leader for all time. Thus, there exists a  $\hat{t}$  such that for all  $t > \hat{t}$ ,  $\hat{B}(t) = \hat{B}$ . Then, following the solution for the frontier agent in Sect. 3,  $g \equiv \hat{B}\sigma - r$  for all  $t > \hat{t}$ .

To solve for full dynamics, the problem is no longer autonomous, as  $g$  depends on the growth rate of the current frontier agent. However, for  $t > \hat{t}$  the Hamiltonian for an agent with state  $x$  and parameter  $B$  is identical to Eq. (16).

### 3.2 Regions of innovation and imitation

Following identical algebra to Sect. 4.2.1, the indifference point between innovating and imitating is  $x^* \equiv (1 - \frac{\sigma m}{c})^{1/m}$ . Hence, the innovation threshold is independent of the particular  $B$  of an agent and the  $\hat{B}$  of the frontier agent.

### 3.3 Innovation only: $x > x^*$

Consider the stationary solution where an agent with parameter  $B$  innovates forever. The first order condition is identical to Eq. (32). As they innovate forever, the Lagrange multiplier is identical to Eq. (57). Combining these equations yields the optimal policy,  $\gamma(B) = B - r/\sigma$ . Substituting into the law of motion generates,

$$\frac{\dot{x}}{x} = \gamma\sigma - g = \sigma(B - \hat{B}). \tag{81}$$

If  $B \neq \hat{B}$ ,  $x$  cannot be stationary as it grows at a constant rate towards either the frontier or the imitation region. Hence, the only agents in the innovation region of the stationary solution are those with  $B = \hat{B}$ .

### 3.4 Imitation only: $x < x^*$

For the stationary solution, the algebra follows identically to Eq. (70). The growth rate of the frontier is  $g(\hat{B}) = \sigma \hat{B} - r$ . In Eq. (70), define  $\hat{\sigma} \equiv \sigma \frac{\hat{B}}{B}$ , then Eq. (71) becomes

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

The derivation is then identical to Appendix 1.2, with the new  $\hat{\sigma}(\hat{B})$ . This yields the quadratic

$$0 = \bar{q}^2 - \left( 2 + \frac{m^2}{c} \left( \hat{\sigma} - \frac{r}{B} \right) - \frac{m}{c} \hat{\sigma} \right) \bar{q} + \left( 1 - \frac{m}{c} \hat{\sigma} \right), \tag{82}$$

$$= \bar{q}^2 - \left( 2 + \frac{m^2}{c} \left( \frac{\hat{B}}{B} \sigma - \frac{r}{B} \right) - \frac{m}{c} \frac{\hat{B}}{B} \sigma \right) \bar{q} + \left( 1 - \frac{m}{c} \frac{\hat{B}}{B} \sigma \right). \tag{83}$$

The solution for the interior root is then  $\bar{x}(B)$ , where

$$\begin{aligned} \bar{x} &= \left[ \frac{1}{2} \left( \left( 2 - \frac{m\hat{\sigma}}{c} + \frac{m^2}{c} \left( \hat{\sigma} - \frac{r}{B} \right) \right) - \text{sign}(m) \right. \right. \\ &\quad \left. \left. \times \left( \left( 2 - \frac{m\hat{\sigma}}{c} + \frac{m^2}{c} \left( \hat{\sigma} - \frac{r}{B} \right) \right)^2 - 4 \left( 1 - \frac{m\hat{\sigma}}{c} \right) \right)^{1/2} \right) \right]^{1/m}. \end{aligned} \tag{84}$$

$$\begin{aligned} &= \left[ \frac{1}{2} \left( \left( 2 - \frac{m\sigma}{c} \frac{\hat{B}}{B} + \frac{m^2}{c} \left( \sigma \frac{\hat{B}}{B} - \frac{r}{B} \right) \right) - \text{sign}(m) \right. \right. \\ &\quad \left. \left. \times \left( \left( 2 - \frac{m\sigma}{c} \frac{\hat{B}}{B} + \frac{m^2}{c} \left( \sigma \frac{\hat{B}}{B} - \frac{r}{B} \right) \right)^2 - 4 \left( 1 - \frac{m\sigma}{c} \frac{\hat{B}}{B} \right) \right)^{1/2} \right) \right]^{1/m}. \end{aligned} \tag{85}$$

### 3.5 Existence of interior $\bar{x}(B) \in \left(0, 1 - \frac{m\sigma \frac{\hat{B}}{B}}{c}\right)$

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

Now note that if  $m > 0$ , then  $\bar{x} < (1 - \frac{m\sigma \frac{\hat{B}}{B}}{c})^{1/m}$  iff  $\bar{q} < (1 - \frac{m\sigma \frac{\hat{B}}{B}}{c})$  and if  $m < 0$ ,  $\bar{x} < (1 - \frac{m\sigma \frac{\hat{B}}{B}}{c})^{1/m}$  iff  $\bar{q} > (1 - \frac{m\sigma \frac{\hat{B}}{B}}{c})$ . 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  $\bar{x}(B) \in \left(0, 1 - \frac{m\sigma \frac{\hat{B}}{B}}{c}\right)$ .

We can now derive  $\underline{B}$  and  $\bar{B}$ .

### 3.6 Derivation of $\underline{B}$

From Eq. (85), solve for the  $B$  such that  $\bar{x} = 0$ , which determines the lower crossing point. Solving for this root gives,

$$\underline{B} = \frac{m\sigma}{c} \hat{B}.$$

From Assumption 2,  $\underline{B} < \hat{B}$ , and if the agent has the  $B$  of the frontier (i.e.,  $\hat{B} = B$ ), then this reduces directly to Assumption 2. Note, if  $m < 0$ , then  $\nexists B > 0$  s.t.  $\bar{x} = 0$ .

### 3.7 Derivation of $\bar{B}$

From Eq. (85), solve for the  $B$  such that  $\bar{x} = x^*$ , which determines the upper crossing point. Solving for this root gives,

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

Rearranging yields

$$\frac{\bar{B}}{\hat{B}} = 1 + \frac{c - m\sigma}{\sigma} \left(1 - \frac{r}{\sigma \hat{B}}\right).$$

From Assumption 1, which guarantees positive growth at the frontier,  $\frac{r}{\sigma \hat{B}} < 1$ , and from Assumption 2,  $c - m\sigma > 0$ . Thus,  $1 < \frac{\bar{B}}{\hat{B}} < \infty$ .

**Lemma 3** For all  $B < \hat{B}$ ,  $\bar{x}(B) < x^*$ .

*Proof* Remember that  $x^* = (1 - \frac{m\sigma}{c})^{1/m}$  and  $\bar{x} < (1 - \frac{m\sigma \frac{\hat{B}}{B}}{c})^{1/m}$ . To determine the relationship between  $x^*$  and  $\bar{x}$  there are four cases,  $m$  positive or negative and  $\frac{\hat{B}}{B}$  above or below one.

If  $B < \hat{B}$ , for both  $m > 0$  or  $m < 0$ , then

$$x^* = \left(1 - \frac{m\sigma}{c}\right)^{\frac{1}{m}} > \left(1 - \frac{m\sigma}{c} \frac{\hat{B}}{B}\right)^{\frac{1}{m}} > \bar{x}.$$

However, if  $B > \hat{B}$ , for both  $m > 0$  or  $m < 0$ , either

$$\left(1 - \frac{m\sigma}{c} \frac{\hat{B}}{B}\right)^{\frac{1}{m}} > \bar{x} \geq \left(1 - \frac{m\sigma}{c}\right)^{\frac{1}{m}} = x^* \quad \text{or} \quad \left(1 - \frac{m\sigma}{c} \frac{\hat{B}}{B}\right)^{\frac{1}{m}} > \left(1 - \frac{m\sigma}{c}\right)^{\frac{1}{m}} = x^* \geq \bar{x}.$$

In brief, if  $B < \hat{B}$ , then  $\bar{x} < x^*$ , but if  $B > \hat{B}$  then it is possible but not necessary for  $\bar{x} > x^*$ .

**Lemma 4** For all  $B < \bar{B}$ ,  $\bar{x}(B) < x^*$ .

*Proof* We have shown that given  $\hat{B}$  there is a unique  $\bar{B}$  such that  $\bar{x} = x^*$ . Now we can show that for  $B < \bar{B}$ ,  $\bar{x}(B) < x^*$ . From Lemma 3, for  $B < \hat{B}$ ,  $\bar{x}(B) < x^*$ . But there is a unique  $\bar{B} > \hat{B}$  at which  $\bar{x}(B) = x^*$ , so by continuity of  $\bar{x}$  in  $B$ ,  $\bar{x}(B) < x^*$  for  $B \in (\hat{B}, \bar{B})$ .

To prove that the solution is stable for  $t > t^*$ , parameterize  $r(\tau)$  and  $B(\tau)$  to be continuous functions for  $\tau \in [0, 1]$ , where  $\tilde{r}(0) = \sigma\hat{B}$ ,  $\tilde{r}(1) = r$ ,  $\tilde{B}(0) = \hat{B}$ , and  $\tilde{B}(1) = B$ . Then use the same homotopy argument as in Appendix 2.2, but over  $\tau$  instead of  $r$ . As the set of  $(B, r) \in [0, \bar{B}] \times [\sigma B, \infty)$  is connected, any point in the parameter space can be reached by appropriately chosen functions  $r(\tau)$  and  $B(\tau)$ .

Finally, note that for any  $B_1$  and  $B_2$  such that  $B_1 < \bar{B}(B_2)$  and  $B_2 < \bar{B}(B_1)$ , there exist initial conditions leading to stationary solutions,  $(\bar{x}(B_2), 1)$  and  $(1, \bar{x}(B_1))$ . Hence, given a set of  $B$ , there are several possible stationary solutions. Now, for each of these initial conditions, perturb an  $\epsilon$  such that  $(\bar{x}(B_2) + \epsilon, 1)$  and  $(1, \bar{x}(B_1) + \epsilon)$ . Since  $\bar{x}(B) < x^*$  and the solutions are stable, these perturbed initial conditions will approach the respective stationary solutions. Hence, the system is not ergodic, and the stationary solution depends on the initial conditions.

### 3.8 Non-existence of innovators with $B \neq \hat{B}$

By contradiction, separately consider the cases where  $B > \hat{B}$  and  $B < \hat{B}$ .

Assume that there exists a  $B$  such that  $B > \hat{B}$  with a stationary  $x^* \leq x^s(B) < 1$ . From Eq. (81), if the agent is an innovator and  $B > \hat{B}$ , then  $x(t)$  has a constant, positive growth rate and hence is not stationary (i.e., will eventually overtake the existing  $\hat{B}$  frontier).

Alternatively, assume that there exists a  $B < \hat{B}$  such that  $x^* \leq x^s(B) < 1$ . From Eq. (81), if the agent is an innovator, then  $x(t)$  has a constant, negative growth rate and is not stationary (i.e., will eventually drop to  $x(t) \leq x^*$ , and hence enter the imitation region).

### 3.9 Non-existence of equilibria with imitators $B > \bar{B}$

We want to show that there does not exist an equilibrium in which agents with  $B > \bar{B}$  imitates. There are two possible cases: the agent invests only in imitation or the agent invests in some combination of imitation and innovation.

For case one, assume by contradiction that an equilibrium exists in which an agent with arbitrarily large  $B/\hat{B}$  invests only in imitation. As proved in the derivation of  $\bar{B} > \hat{B}$ , since  $\bar{x}(B)$  crosses  $x^*$  at  $\bar{B}$ , any agent with  $B > \bar{B} > \hat{B}$  would optimally choose to invest in innovation, contradicting that it optimally invests only in imitation.



For case two, assume there exists a stationary equilibrium such that  $B > \hat{B}$  and the agent invests in both innovation and imitation simultaneously. By using the two first-order necessary conditions in Eq. (27), this is only possible if  $x^s(B) = x^*$ , where the agent optimally remains at the  $x^*$  threshold. Next, note that from the first order condition,  $\sigma = D(x^*)$ , so the law of motion is

$$0 = \frac{\dot{x}}{x} = \sigma \bar{\gamma} + D(x)\bar{s} - \sigma \hat{B} + r = \sigma(\bar{\gamma} + \bar{s}) - \sigma \hat{B} + r.$$

Since the returns to innovation and imitation are identical at  $x^*$ , the optimal investment composition between innovation and imitation is indeterminate. Nonetheless, total investment is  $\bar{\gamma} + \bar{s} = \hat{B} - \frac{r}{\sigma}$ .

Rephrasing the contradiction, assume that for some  $B > \hat{B}$ , the choice of  $x^s(B) = x^*$  and  $\bar{\gamma} + \bar{s} = \hat{B} - \frac{r}{\sigma}$  is the optimal solution. Then the optimal path of  $z$  is  $z(t) = z_0 e^{g(\hat{B})t}$ , where  $z_0$  is an initial condition and  $g(\hat{B}) \equiv \sigma \hat{B} - r$ . Substituting into the agent’s optimization problem in Eq. (1) yields,

$$U(z; B, \hat{B}) = \int_0^\infty (\ln(Bz - sz - \gamma z)) e^{-rt} dt, \tag{86}$$

$$= \int_0^\infty \left( \ln(z_0 e^{g(\hat{B})t}) + \ln\left(B - \hat{B} + \frac{r}{\sigma}\right) \right) e^{-rt} dt, \tag{87}$$

$$= \left( \ln(z_0) + \frac{1}{m} \ln\left(1 - \frac{\sigma m}{c}\right) + \ln\left(B - \hat{B} + \frac{r}{\sigma}\right) \right) \int_0^\infty e^{-rt} dt + (\sigma \hat{B} - r) \int_0^\infty t e^{-rt} dt, \tag{88}$$

$$= \frac{1}{r} \ln(z_0) + \frac{1}{r} \ln\left(B - \hat{B} + \frac{r}{\sigma}\right) + \frac{1}{r^2} (\sigma \hat{B} - r). \tag{89}$$

An alternative feasible policy would be for this agent to choose to become the frontier agent and grow at a rate  $g = \sigma B - r$ . In that case  $z(t) = z_0 e^{g(B)t}$  and the value is

$$U(z; B, B) = \int_0^\infty \left( \ln(z_0 e^{(\sigma B - r)t}) + \ln\left(\frac{r}{\sigma}\right) \right) e^{-rt} dt, \tag{90}$$

$$= \frac{1}{r} \ln(z_0) + \frac{1}{r} \ln\left(\frac{r}{\sigma}\right) + \frac{1}{r^2} (\sigma B - r). \tag{91}$$

Comparing the value of the two policies by subtracting Eq. (89) from Eq. (91),

$$U(z; B, B) - U(z; B, \hat{B}) = \frac{1}{r} \left( \frac{\sigma}{r} (B - \hat{B}) - \ln\left(\frac{\sigma}{r} (B - \hat{B}) + 1\right) \right). \tag{92}$$

Since  $B - \hat{B} > 0$ ,

$$U(z; B, B) - U(z; B, \hat{B}) \geq 0. \tag{93}$$

This expression holds with equality iff  $B = \hat{B}$ . Therefore, we have a contradiction as the choice to become the frontier agent dominates the policy of investing in innovation and imitation simultaneously to remain at  $x^*$  with leader  $\hat{B}$ .

## References

- Acemoglu, D., Aghion, P., & Zilibotti, F. (2006). Distance to frontier, selection and economic growth. *Journal of the European Economic Association*, 4(1), 37–74.
- Acemoglu, D., Robinson, J. A., & Verdier, T. (2012). Can't we all be more like Scandinavians? Asymmetric growth and institutions in an interdependent world. NBER Working Paper No. 18441.
- Aghion, P., Harris, C., Howitt, P., & Vickers, J. (2001). Competition, imitation and growth with step-by-step innovation. *Review of Economic Studies*, 68, 467–492.
- Aghion, P., & Howitt, P. (1998). *Endogenous growth theory*. Cambridge, MA: MIT Press.
- Aghion, P., & Howitt, P. (2009). *The economics of growth*. Cambridge, MA: MIT Press.
- Aghion, P., Howitt, P., & Mayer-Foulkes, D. (2005). The effect of financial development on convergence: Theory and evidence. *The Quarterly Journal of Economics*, 120, 173–222.
- Alvarez, F., Buera, F., & Lucas, R. E. (2012). *Idea flows, economic growth, and trade*. New York: mimeo.
- Barro, R. J., & Sala-i-Martin, X. (1997). Technological diffusion, convergence, and growth. *Journal of Economic Growth*, 2(1), 1–26.
- Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development: Evidence from aggregate cross-country data. *Journal of Monetary Economics*, 34(2), 143–173.
- Benhabib, J., & Spiegel, M. M. (2005). Human capital and technology diffusion. In P. Aghion & S. Durlauf (Eds.), *Handbook of economic growth, chap. 13* (Vol. 1, pp. 935–966). Amsterdam: Elsevier.
- Damsgaard, E. F., & Krusell, P. (2010). The world distribution of productivity: Country TFP choice in a Nelson–Phelps economy. NBER Working Paper No. w16375.
- Eeckhout, J., & Jovanovic, B. (2002). Knowledge spillovers and inequality. *American Economic Review*, 92(5), 1290–1307.
- Ericson, R., & Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, 62(1), 53–82.
- Gerschenkron, A. (1962). *Economic backwardness in historical perspective*. Cambridge, MA: Belknap Press of Harvard University Press.
- Gill, I. S., & Raiser, M. (2012). *Golden growth: Restoring the lustre of the European economic model*. Washington, DC: The World Bank.
- Grossman, G. M., & Helpman, E. (1991). Quality ladders in the theory of growth. *The Review of Economic Studies*, 58(1), 43–61.
- Howitt, P., & Mayer-Foulkes, D. (2005). R&D, implementation, and stagnation: A Schumpeterian theory of convergence clubs. *Journal of Money, Credit and Banking*, 37(1), 147–177.
- Jones, C. I. (2005). Growth and ideas. In P. Aghion & S. Durlauf (Eds.), *Handbook of economic growth, chap. 16* (Vol. 1, pp. 1063–1111). Amsterdam: Elsevier.
- Jovanovic, B., & Rob, R. (1989). The growth and diffusion of knowledge. *Review of Economic Studies*, 56(4), 569–582.
- Klenow, P. J., & Rodriguez-Clare, A. (2005). Externalities and growth. In P. Aghion & S. Durlauf (Eds.), *Handbook of economic growth, chap. 11* (pp. 817–861). Amsterdam: Elsevier.
- König, M., Lorenz, J., Zilibotti, F. (2012). Innovation vs imitation and the evolution of productivity distributions. CEPR Discussion Paper No. 8843.
- Kortum, S. S. (1997). Research, patenting, and technological change. *Econometrica*, 65(6), 1389–1420.
- Laincz, C. A. (2009). R&D subsidies in a model of growth with dynamic market structure. *Journal of Evolutionary Economic*, 19, 643–673.
- Lucas, R. E. (2009). Trade and the diffusion of the industrial revolution. *American Economic Journal: Macroeconomics*, 1(1), 1–25.
- Lucas, R. E., & Moll, B. (2013). Knowledge growth and the allocation of time. *Journal of Political Economy* (forthcoming).
- Luttmer, E. G. J. (2011). Technology diffusion and growth. *Journal of Economic Theory*, 147(2), 602–622.
- Magaziner, I., & Patinkin, M. (1990). *The silent war*. New York: Vintage Books.
- Milnor, J. W. (1965). *Topology from the differentiable viewpoint*. Charlottesville, VA: The University Press of Virginia.
- Mukoyama, T. (2003). Innovation, imitation, and growth with cumulative technology. *Journal of Monetary Economics*, 50(2), 361–380.

- Nelson, R. R., & Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. *American Economic Review*, 56(1/2), 69–75.
- Parente, S. L., & Prescott, E. C. (1994). Barriers to technology adoption and development. *Journal of Political Economy*, 102(2), 298–321.
- Perla, J., & Tonetti, C. (2013). Equilibrium imitation and growth. *Journal of Political Economy* (forthcoming).
- Rustichini, A., & Schmitz, J. A. (1991). Research and imitation in long-run growth. *Journal of Monetary Economics*, 27(2), 271–292.
- Segerstrom, P. S. (1991). Innovation, imitation, and economic growth. *Journal of Political Economy*, 99(4), 807–827.
- Seierstad, A., & Sydsaeter, K. (1977). Sufficient conditions in optimal control theory. *International Economic Review*, 18(2), 367–391.
- Thompson, P. (2001). The microeconomics of an R&D-based model of endogenous growth. *Journal of Economic Growth*, 6(4), 263–283.
- Thompson, P., & Waldo, D. (1994). Growth and trustified capitalism. *Journal of Monetary Economics*, 34(3), 445–462.