

Federal Reserve Bank of Minneapolis
Research Department

Eventually, Noise and Imitation Implies Balanced Growth*

Erzo G.J. Luttmer

Working Paper 699

August 2012

ABSTRACT

This paper adds imitation by incumbent firms, and not just by new entrants, to the model of selection and growth developed in Luttmer [2007]. Noisy firm-level innovation and imitation give rise to a long-run growth rate that exceeds the average rate at which individual firms innovate. It can be shown, in simple examples, that the economy converges to a long-run balanced growth path from compactly supported initial productivity distributions. The right tail of the stationary distribution of de-trended productivity is approximately Pareto. The tail index of this distribution depends on the rate at which incumbents are able to imitate only indirectly, through general equilibrium effects of this parameter on the equilibrium growth rate.

JEL classification: L110, O330

Keywords: Technology diffusion; Size distribution of firms; Endogenous growth

*Luttmer: University of Minnesota and Federal Reserve Bank of Minneapolis. This is a first draft. Comments and suggestions welcome. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1. INTRODUCTION

Imitation can speed up growth if there is something to imitate, and if it does not do too much harm to the incentives of firms to innovate. But imitation makes a heterogeneous population more homogeneous, and this tends to make the effects of imitation peter out. Luttmer [2007] presents a model of growth driven by imitation and selection in which entering firms can imitate incumbent firms, and in which low-productivity incumbents are forced to exit because of fixed costs. Heterogeneity is preserved because individual firm productivities follow geometric Brownian motions. This paper adds imitation by incumbent firms, not just by entering firms, and shows that the noise generated by Brownian shocks to productivity continues to be enough to break the tendency of imitation to eliminate heterogeneity. In fact, the economy has a stationary distribution of firm productivity relative to aggregate productivity with the same Pareto tail that was obtained in Luttmer [2007]. Aggregate productivity grows at an endogenous rate, and the only way the right tail of the firm productivity distribution depends on incumbent imitation is through general equilibrium effects on the equilibrium growth rate of aggregate productivity.

In the model, entrants and incumbents acquire imitation opportunities at Poisson rates. For entrants this takes costly entrepreneurial effort, and for incumbents this happens at a rate that is taken to be exogenous. But entrants and incumbents have exactly the same ability to take advantage of imitation opportunities. When an opportunity to imitate arrives, both can sample a random incumbent firm and copy its productivity at no further cost. Since entrants sample incumbents that have not chosen to exit, adoption is always profitable for entrants. Incumbents only adopt the sampled technology if it implies an improvement in productivity.

The Poisson arrival of imitation opportunities delays the equalizing effect of imitation. But if incumbent productivities are growing at a common deterministic rate in the absence of imitation opportunities, and if the initial distribution of productivity has compact support, then this initial heterogeneity will quickly disappear, as low-productivity firms catch up with the most productive incumbent firms. Alvarez, Buera and Lucas [2007] and Lucas and Moll [2011] avoid this implication by focusing on balanced growth paths that can only arise from initial productivity distributions with such thick right tails that imitation opportunities are never exhausted. This is an assumption about initial conditions that many will find hard to defend. Taken literally, it means that all the growth that has happened since ancient times has been about more and more producers imitating ideas that were present in the population all along. The contribution of this

paper is to show that a little bit of independent randomness in the rates at which individual firms innovate is enough to preserve heterogeneity in the presence of imitation, and produce long-run growth, even if the range of productivities at some initial date is extremely limited.

The economy can be simplified by abstracting from fixed costs and assuming a fixed population of agents with heterogeneous productivities who simply produce for their own consumption instead of for the market. In this simplified economy, the distribution of productivities happens to satisfy a partial differential equation—an example of a reaction-diffusion equation—that was introduced by Fisher [1937] and Kolmogorov, Petrovskii and Piskunov [1937] to study the geographic spread of an advantageous gene. Cavalli-Sforza and Feldman [1981] have used this equation to describe the geographic spread of culture in human populations. As discussed in more detail in Section 2.2.1 below, the interpretation of random improvements and imitation given here is very different from the classic interpretation of this equation. Adapted to an economy with noise and imitation, the key result of Kolmogorov, Petrovskii and Piskunov [1937] shows that an initial distribution of productivity concentrated at a mass point transitions over the long run into a distribution of de-trended productivity that is stationary, with a trend that grows at a specific rate. Later investigations have shown that this long-run trend is approached from below. In a rough sense, this implies a model of historical development in which the economy transitions over time from an economy with limited variation in productivity to one that grows at a steady rate, with a thick-tailed distribution of productivity.

Such a convergence result remains an open problem in the market economy that arises when firms participate in product and labor markets, and make optimal entry and exit decisions. But the balanced growth path for this economy can be described by a very simple system of ordinary differential equations that can be solved numerically. Furthermore, this system of differential equations shows precisely why the right-tail of the de-trended productivity distribution is approximately Pareto, and why incumbent imitation has no direct effect on how thick this right tail will be.

Some Related Work This paper aims to add to a literature on models of growth and heterogeneity that feature search and imitation. In Jovanovic and Rob [1989], the engine of growth is imitation in random meetings. Fogli and Veldkamp [2011] study the spread of ideas when meetings are not random but based on a network structure. Jovanovic and Eeckhout [2002] consider the effect of imitation on firm size distributions using a more

reduced-form copying mechanism. Kortum [1997] has a distribution of productivities that shifts to the right over time as agents search. This paper extends Luttmer [2007] by adding incumbent imitation to an economy with only entrant imitation. This follows and builds on the work of Alvarez, Buera and Lucas [2007], Lucas and Moll [2011], König, Lorenz and Zilibotti [2012], and Perla and Tonetti [2012].¹

Alvarez, Buera and Lucas [2007] and Lucas and Moll [2011] abstract from entry and exit and focus on incumbent imitation. They show that imitation alone gives rise to balanced growth when the initial distribution of productivities has a sufficiently thick right tail, but not otherwise. Lucas and Moll [2011] assume producers have a producer-specific fixed factor that must be divided between the production of consumption goods and the creation of imitation opportunities. They are able to characterize the efficient allocation and show the extent to which search and imitation externalities compromise welfare.

Perla and Tonetti [2012] consider a fixed population of incumbent firms imitating incumbent firms, incorporating a trade-off between production and imitation. In their economy, only actively producing firms can be sampled by searching firms, and only relatively productive firms choose to produce rather than imitate. As a result, imitating firms always sample from a population of firms that are more productive than themselves. Balanced growth would not arise from an initial distribution with bounded support, but does with a thick-tailed initial distribution of productivity.

König, Lorenz and Zilibotti [2012] also consider imitation of incumbent firms by incumbent firms, using a mechanism similar to that in Lucas and Moll [2011] and Perla and Tonetti [2012]. Individual firms decide not between current production and imitation, but between firm-specific innovation and imitation. This is an important and crucial difference because the resulting innovation generates stochastic improvements in the individual productivities of firms that preserve heterogeneity, in spite of the tendency of imitation to make firms similar. As a result, König, Lorenz and Zilibotti [2012] obtain convergence from limited initial heterogeneity to long-term growth with a stable non-degenerate distribution of de-trended productivity. In their economy, innovation moves a firm up a quality ladder at a Poisson rate. This is probably easier to interpret than Brownian shocks to productivity, but it does not allow for the direct application of Kolmogorov, Petrovskii and Piskunov [1937] that is available in simple versions of the economy described here. More importantly, complementing entrant imitation with

¹See Atkeson and Burstein [2011] for a very different way to model knowledge spillovers in an economy with realistic firm heterogeneity and entry and exit.

incumbent imitation in the growth model of Luttmer [2007] results in a highly tractable benchmark model that can accommodate firms operating subject to decreasing returns or as monopolistic competitors. And firms are making forward-looking decisions that generate realistic predictions about firm entry and exit.

The Rest of the Paper Section 2 makes the strong equalizing effect of imitation explicit in a very simple example in which incumbent innovation is deterministic. It then adds Brownian noise, while maintaining the assumption of a fixed population of producers. Section 3 characterizes the stationary distributions that arise from Brownian innovation in an economy with imitating entrants and incumbents, and with exit at an exogenously specified exit barrier. Section 4 then lays out the complete economy with optimal entry and exit decisions and characterizes the balanced growth path. Section 5 concludes.

2. INNOVATION AND IMITATION IN A FIXED POPULATION

Consider an economy with a large population of agents who produce and consume. Everyone has one unit of labor and a certain level of productivity that can evolve over time. The imitation mechanism is that of Alvarez, Buera and Lucas [2007] and Lucas and Moll [2011].²

2.1 Deterministic Innovation

Write x_t for log productivity of a particular producer at time t . Suppose productivity evolves deterministically as a result of innovation by the producer, and stochastically because of randomness in the imitation process. Specifically, suppose

$$dx_t = \theta_I dt + \Delta_t dN_t, \tag{1}$$

where θ_I represents innovation by this producer, N_t is a Poisson process with arrival rate β that counts opportunities to imitate. When an imitation opportunity arrives, the producer randomly selects another producer from the population. The producer with the imitation opportunity can copy the technology of the randomly selected producer, and will do so if this generates an increase in productivity. The resulting increase in productivity is represented by $\Delta_t \geq 0$.

²In comparing results, note well that the state variable used below is log productivity, not productivity. This simplifies Kolmogorov forward equations.

Write $P(t, x)$ for the cumulative distribution of log productivity at time t . The initial distribution $P(0, x)$ is given and then the process of innovation and imitation determines how $P(t, x)$ evolves over time.

2.1.1 Degenerate Initial Conditions

In the simplest scenario, suppose all producers have the same initial productivity x_0 , so that $P(0, x)$ is a non-decreasing step function with a unit step at x_0 . Since all producers improve their own productivity at the same rate θ_I , the population remains homogeneous, and producers cannot improve their productivity by imitating others. As a result, the degenerate distribution simply shifts to the right, so that $P(t, x) = P(0, x - \theta_I t)$.

2.1.2 Stationary Distributions

Suppose instead that the initial productivity distribution $P(0, x)$ is smooth, and conjecture that $P(t, x)$ will be as well for all t . Then $P(t, x)$ must satisfy the partial differential equation

$$D_t P(t, x) = -\theta_I D_x P(t, x) - \beta P(t, x)[1 - P(t, x)]. \quad (2)$$

The negative term $-\theta_I D_x P(t, x)$ reflects the deterministic flow of individual producers to the right. There is also a flow $\beta P(t, x)$ of producers with log productivities in $(-\infty, x]$ that sample from the population of producers. These producers move out of the range $(-\infty, x]$ if they sample another producer with a productivity in excess of x . This happens with probability $1 - P(t, x)$. Note that $P(t, x)$ can only decrease with t . There is no entry into the range of log productivities $(-\infty, x]$ from producers not already in that range.

Conjecture that there is a cumulative distribution function F so that

$$P(t, x) = F(x - \theta_E t),$$

for some θ_E positive. Thus $F(x) = P(0, x)$ is the initial distribution of log productivity. Inserting this into (2) and evaluating the result at $z = x - \theta_E t$ shows that F must satisfy

$$(\theta_E - \theta_I)DF(z) = \beta F(z)[1 - F(z)] \quad (3)$$

for all z . The right-hand side is non-negative and can only be zero when $F(z) \in \{0, 1\}$. It follows that $P(t, x) = F(x - \theta_E t)$ can only be a distribution function that solves (2) if $\theta_E > \theta_I$. To simplify the calculations, write

$$\gamma = \frac{\beta}{\theta_E - \theta_I} > 0. \quad (4)$$

The differential equation for F then becomes

$$DF(z) = \gamma F(z)[1 - F(z)],$$

which is solved by the logistic function

$$F(z) = \frac{1}{1 + \left(\frac{1}{F(0)} - 1\right) e^{-\gamma z}}. \quad (5)$$

Since $\gamma > 0$, this is increasing by construction. Note that the right tail $1 - F(z)$ behaves like $e^{-\gamma z}$ for z large. Thus the level of productivity e^z behaves like a Pareto distribution with right tail index γ .³

Using (4) and (5) the smooth stationary distributions that solve (2) can be written as

$$P(t, x) = \frac{1}{1 + \left(\frac{1}{P(0,0)} - 1\right) \exp\left(-\frac{\beta(x - \theta_E t)}{\theta_E - \theta_I}\right)}. \quad (6)$$

This solution is indexed by two parameters: the initial value $P(0, 0)$ and the trend parameter $\theta_E > \theta_I$. The solution is non-degenerate if and only if $P(0, 0) \in (0, 1)$. Note well that the initial distribution $P(0, x)$ depends on the parameter $\theta_E > \theta_I$. Thus the assumed shape of the initial distribution determines the rate at which productivity grows over time.

Starting from the right initial conditions, the long-run growth rate of this economy is above the rate at which individual producers improve productivity. In Luttmer [2007], this happens because of entry, exit, and imitation by entrants. Here entry and exit plays no role and imitation by incumbents suffices to obtain accelerated growth. In Lucas and Moll [2011], $\theta_I = 0$, but β is not independent of x because high-productivity producers choose to imitate less than low-productivity producers.

2.1.3 Smooth Initial Conditions

It turns out that in this environment, accelerated long-run growth is special, in that it can only arise from fortuitous initial conditions. To make this explicit requires solving (2) for more general initial conditions. The partial differential equation (2) can be reduced to a system of ordinary differential equations by de-trending the productivity state variable. Define

$$Q(t, z) = P(t, z + \theta_I t).$$

³One can also construct stationary distributions on $[b, \infty)$ with a point mass at b . These are not smooth everywhere. Distributions with support (b, ∞) are discussed below.

Evaluating (2) at $x = z + \theta_1 t$ gives

$$D_t Q(t, z) = -\beta Q(t, z)[1 - Q(t, z)]. \quad (7)$$

This is now a system of ordinary differential equations, instead of a partial differential equation. The solution is $Q(t, z) = 1/(1 + [-1 + 1/Q(0, z)]e^{\beta t})$. The original distribution is $P(t, x) = Q(t, x - \theta_1 t)$, and $P(0, x) = Q(0, x)$ in particular. Hence

$$P(t, x) = \frac{1}{1 + \left(\frac{1}{P(0, x - \theta_1 t)} - 1 \right) e^{\beta t}} \quad (8)$$

for any smooth initial conditions $P(0, x)$. To check this against the stationary distributions obtained earlier, note that (17) and (6) imply

$$\begin{aligned} P(t, x) &= \frac{1}{1 + \left(\frac{1}{\frac{1}{1 + \left(\frac{1}{P(0,0)} - 1 \right) e^{-\gamma(x - \theta_1 t)}}} - 1 \right) e^{\beta t}} \\ &= \frac{1}{1 + \left(\frac{1}{P(0,0)} - 1 \right) e^{-\frac{\beta}{\theta_E - \theta_1}(x - \theta_1 t) + \frac{\beta}{\theta_E - \theta_1}(\theta_E - \theta_1)t}} = \frac{1}{1 + \left(\frac{1}{P(0,0)} - 1 \right) e^{-\gamma(x - \theta_E t)}}, \end{aligned}$$

as required. If $P(0, x)$ is the logistic distribution (6), then (8) implies that $P(t, x)$ is a logistic distribution.

But (8) is not a logistic distribution if $P(0, x)$ is not logistic. It is clear from the description of the economy that initial conditions with a compact support lead to $P(t, x)$ that have a compact support as well—nobody can overtake the most productive producer or fall behind the least productive producer. To make explicit what happens to $P(t, x)$ over time, suppose U and $L < U$ are finite and assume $P(0, L) = 0$ and $P(0, U) = 1$. If $P(0, x)$ varies smoothly in between, then (8) implies

$$P(t, x + \theta_1 t) = \frac{1}{1 + \left(\frac{1}{P(0, x)} - 1 \right) e^{\beta t}}.$$

This means that $P(t, x + \theta_1 t) = 0$ for $x < L$ and $P(t, x + \theta_1 t) = 1$ for $x > U$. No producer at time t will be less productive than $L + \theta_1 t$, or more productive than $U + \theta_1 t$. In between, $\beta > 0$ implies

$$\lim_{t \rightarrow \infty} P(t, x + \theta_1 t) = 0, \quad x \in (L, U).$$

That is, the productivity distribution will converge to a distribution concentrated at $U + \theta_1 t$, and not to any logistic limiting distribution. Because the initial distribution has

compact support and everyone innovates at the same rate, one cannot hope to obtain a stationary distribution by taking out a trend that differs from $\theta_I t$. The long-run growth rate is determined by the rate at which individual producers innovate. But imitation does matter. The easier it is to imitate, in the sense that imitation opportunities arise more frequently, the more quickly will any initial productive heterogeneity disappear.

For an example, consider an initial distribution with support $[0, 1]$ given by $P(0, x) = \{1 + \cos(\pi(x - 1))\} / 2$ for $x \in [0, 1]$. Clearly, this is smooth everywhere. At fixed intervals of time, the distributions $P(t, x)$ are shown in Figure 1.

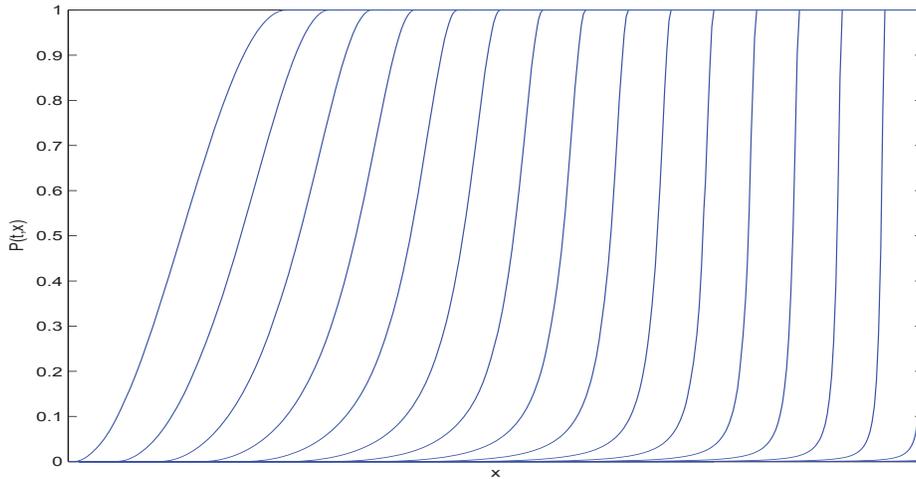


FIGURE 1 The distributions $P(t, x)$ for $P(0, x)$ with compact support.

As expected, the $P(t, x)$ converge to a point mass that shifts to the right at rate θ_I . All initial heterogeneity disappears over time, as low-productivity producers receive more and more draws that allow them to approach the productivity frontier. In this economy, imitation changes the level of aggregate output, as more and more producers use technologies close to the frontier. But the long-run growth rate of the economy is unaffected by imitation and completely determined by the rate at which individual producers innovate.

2.2 Brownian Innovation

Imitation tends to eliminate heterogeneity, even if there is randomness in who gets to imitate whom. An obvious way to preserve heterogeneity is to introduce randomness in the innovation processes of individual producers. Consider the same economy as before, but now add Brownian terms to the individual innovation processes. Thus innovation makes individual productivity go up and down stochastically. One can imagine that

all change is irreversible and producers do not know in advance whether a change is for the better. Or the market conditions encountered by a particular producer could change. Imitation is still an unambiguous source of productivity improvement: producers only adopt a technology when it is more productive than their own, and the implicit assumption is that it can be copied perfectly.

Therefore, instead of (1), suppose that the log productivity of a particular producer evolves according to

$$dx_t = \theta_I dt + \sigma dW_t + \Delta_t dN_t, \quad (9)$$

where W_t is a standard Brownian motion, N_t a Poisson process with arrival rate β , and Δ_t an increment in log productivity that is obtained by copying another randomly selected producer—provided this leads to a productivity improvement. Every producer is subject to different Brownian noise, and this will be a force for increasing productivity dispersion. In a large population, any initial discreteness in the initial productivity distribution is smoothed out instantaneously.

As before, let $P(t, x)$ be the distribution of log productivity at time t . Then

$$D_t P(t, x) = -\theta_I D_x P(t, x) + \frac{1}{2} \sigma^2 D_{xx} P(t, x) - \beta P(t, x) [1 - P(t, x)] \quad (10)$$

for all $t \geq 0$ and $x \in (-\infty, \infty)$. Without the last term on the right-hand side, this is the Kolmogorov forward equation, written in terms of distribution functions on the unbounded domain $(-\infty, \infty)$. The last term on the right-hand side of (10) arises from imitation, exactly as in the deterministic case (2). As before, take out the common trend $\theta_I t$ by considering $Q(t, z) = P(t, z + \theta_I t)$. Then (10) becomes

$$D_t Q(t, z) = \frac{1}{2} \sigma^2 D_{zz} Q(t, z) - \beta Q(t, z) [1 - Q(t, z)], \quad (11)$$

which generalizes (7).

Without imitation, when $\beta = 0$, (10) is solved by the distribution function of a normal random variable with mean $x_0 + \theta_I t$ and variance $\sigma^2 t$, which reduces to a point mass at x_0 when $t = 0$. One can then integrate over the initial distribution $P(0, x_0)$ to obtain a solution that satisfies the desired initial condition. Since random walks, with or without trends, are non-stationary, the distributions $P(t, x)$ that solve (10) will not converge when $\beta = 0$.

The properties of (10) are very different when β is positive. Consider solutions to (10) of the form $P(t, x) = F(x - \theta_E t)$. Evaluating (10) at $z = x - \theta_E t$ gives

$$(\theta_E - \theta_I) D F(z) + \frac{1}{2} \sigma^2 D^2 F(z) = \beta F(z) [1 - F(z)], \quad (12)$$

for all $z \in (-\infty, \infty)$. Since F is supposed to be a distribution function, $F(z) \in [0, 1]$, $F(-\infty) = 0$ and $F(\infty) = 1$. Note that $\theta_E \leq \theta_I$ would imply that F is convex. But distribution functions on $(-\infty, \infty)$ cannot be convex, and so (12) cannot be solved by a cumulative distribution function unless $\theta_E > \theta_I$. If there is a stationary distribution, then the growth rate of aggregate productivity must be higher than the average rate at which individual productivities improve over time.

The partial differential equation (11) and the associated stationary distributions implied by (12) have been widely studied. Equation (12) can be solved for a distribution function on $(-\infty, \infty)$ provided that $\theta_E \geq \theta_I + \sigma\sqrt{2\beta}$.⁴ Kolmogorov, Petrovskii and Piskunov [1937] and McKean [1975, 1976] prove that starting from a point mass $P(0, x)$, the distributions $P(t, x + \theta_E t)$ converge to a non-degenerate limiting distribution when θ_E is taken to be at its lower bound

$$\theta_E = \theta_I + \sigma\sqrt{2\beta}. \quad (13)$$

The same is also true for initial distributions with compact support.⁵ Furthermore, McKean [1975] shows that the median of $P(t, x)$ always lags behind the trend $\theta_E t$. In this sense, the θ_E defined in (13) is a maximal attainable long-run growth rate to which an economy converges from below, given compactly supported initial conditions. This yields a precise prediction for the long-run growth rate θ_E , and for the resulting stationary productivity distribution. Clearly, imitation is essential for accelerated growth. We have already seen that growth at a rate above θ_I cannot arise from limited initial heterogeneity when there is no noise. The formula (13) shows that the amount of long-run acceleration that is possible with noisy innovation is linear in the standard deviation of that noise.

2.2.1 Some Traditional Interpretations of (11)

Fisher [1937] proposed (11) to describe the geographic spread of an advantageous gene, and Kolmogorov, Petrovskii and Piskunov [1937] adopted the same interpretation. More broadly, (11) is an example of a reaction-diffusion equation. The second term on the right-hand side of (11) can be a more general function of $Q(t, z)$ in such equations. Reaction-diffusion equations have a number of different interpretations in biology and the physical sciences. In the original interpretation of Fisher [1937], z is a location, $Q(t, z)$ is the population density at a point z , and $-\beta$ is positive (note that $1 - Q(t, z)$ satisfies

⁴Their argument is essentially the one described below for an economy with entry and exit.

⁵See Bramson [1983] for assumptions on $P(0, x)$ that will result in convergence to a stationary distribution that solves (12) for a growth rate θ_E that exceeds the one given in (13).

the same equation with β replacing $-\beta$). The quadratic term $-\beta Q(t, z)[1 - Q(t, z)]$ then measures logistic population growth at the location z , with growth rates shrinking from near $-\beta$ to near zero as the population in a particular location gets close to its “carrying capacity,” normalized to 1. In this interpretation, $\frac{1}{2}\sigma^2 D_{zz}Q(t, z)$ measures random movement of population across locations, usually referred to as diffusion. Cavalli-Sforza and Feldman [1981, Section 1.9] use Fisher’s model to interpret cultural transmission in human populations.

Here, z measures a productivity rather than a location, and $Q(t, z)$ is a distribution function rather than a density. The classic diffusion terminology is potentially confusing in the present context. The diffusion term $\frac{1}{2}\sigma^2 D_{zz}Q(t, z)$ in (11) does not represent technology diffusion, but continuous random and autonomous changes in the productivities of individual producers. Instead, technology diffusion takes place in jumps, when producers imitate randomly selected producers with productivities greater than their own. This is accounted for by the term $\beta Q(t, z)[1 - Q(t, z)]$ in (11) and not by the diffusion term $\frac{1}{2}\sigma^2 D_{xx}Q(t, z)$.

Considering the case $\beta = \sigma^2 = 1$, McKean [1975] shows that, starting from $Q(0, z)$ equal to a point mass at 0, the $Q(t, z)$ that solves (11) is the cumulative distribution function of the first order statistic of a branching Brownian motion. In the present context, suppose initially there is one producer with log productivity zero. Log productivity then follows a standard Brownian motion. At a unit Poisson rate, this producer generates a spinoff.⁶ The initial log productivity of this spinoff producer is that of its parent. Immediately after the spinoff, the log productivities of parent and spinoff proceed as independent standard Brownian motions. Both parent and spinoff will produce further spinoffs at independent Poisson rates, and so on. This generates a stochastically growing population of producers, with a growth rate that settles down over time to the unit Poisson rate at which producers generate spinoffs. In this growing but finite population, $Q(t, z)$ is the distribution at time t of the productivity of the most productive producer. Note well that this is not the cross-sectional distribution of productivities at time t that is of interest here.

⁶See Chatterjee and Rossi-Hansberg [2007] for model of idea flows and the firm size distribution based on spinoffs.

3. ENTRY AND SELECTIVE EXIT

Consider again the economy with Brownian innovation. A simple way to introduce population growth into this economy is to assume that, at a Poisson rate η , every producer generates a new entrant who inherits its productivity, as in the above discussion of McKean [1975]. Equivalently, one can assume there is a flow of entrants equal to η times the measure of incumbent producers, who enter with productivities sampled at random from the incumbent population. All this does is cause the population to grow at the rate η . In this growing population, (10) and all its implications still hold. In particular, productivity will tend to grow at the rate (13) in the long run. One can also assume that some producers exit randomly. Conclusions about the long-run distribution of productivity are not affected.

A slightly different scenario arises when there is a flow $E_t = Ee^{\eta t}$ of new entrants, where E and η are positive parameters specified independently of the size of the initial population of producers. As before, entrants sample an entry productivity from the incumbent population. Both the entry rate and the population growth rate will converge to η in the long run, but these rates will not equal η from the start unless the initial population size happens to be E/η . Nevertheless, one expects the convergence result of Kolmogorov, Petrovskii and Piskunov [1937] and McKean [1975] to apply.

Again take the flow of entrants to be $E_t = Ee^{\eta t}$ so that the population has to grow at the rate η in the long run. Suppose further that producers exit when their productivity crosses $b + \theta_E t$ from above. The constant b simply depends on what t is associated with calendar time, and exit at $b + \theta_E t$ will force the distribution as well as its density to be zero at this exit barrier, just as they would be at $-\infty$ in an economy without exit (see Cox and Miller [1965]). But with $b > -\infty$, the slope of the stationary density at the exit barrier scales with the exit rate, and this now becomes an equilibrium variable. As shown in Luttmer [2007], even given a particular value for θ_E , there can be multiple stationary distributions, indexed not by the growth rate θ_E , but by the exit rate at the barrier $b + \theta_E t$. The question then again arises: to which of these stationary distributions will productivity distributions converge given compactly supported initial conditions?

3.1 Population Dynamics

Let $S(t, z)$ denote the measure of incumbent producers at time t with log productivity $z + \theta_E t$ or higher. Thus $S(t, z)$ integrates a density from the right, and it describes the distribution of log productivity relative to the trend $\theta_E t$ of the exit barrier. The drift

of incumbent relative to exit productivities is thus $\mu = \theta_I - \theta_E$. If $\theta_E > \theta_I$ then the exit barrier catches up with incumbent productivity on average, and μ will be negative. The flow E_t of entering producers samples from the distribution $1 - S(t, z)/S(t, b)$ and entrants can start with the productivity they have sampled. As before, incumbent producers receive imitation opportunities at the Poisson rate β .

Starting from $S(0, z)$, the measure of incumbent producers at time t evolves according to

$$D_t S(t, z) = -\mu D_z S(t, z) + \frac{1}{2} \sigma^2 D_{zz} S(t, z) + (E_t + \beta [S(t, b) - S(t, z)]) \frac{S(t, z)}{S(t, b)}. \quad (14)$$

The left-hand side and the first two terms on the right-hand side are the result of integrating the Kolmogorov forward equation for the density. The last term on the right-hand side measures the flow of incumbent producers who reach $[z, \infty)$ as new entrants or as incumbent producers with an imitation opportunity. The flow of entrants is E_t , and a flow $\beta[S(t, b) - S(t, z)]$ of incumbents receive imitation opportunities that could move them into $[z, \infty)$. But because they can only imitate a randomly selected incumbent producer, they succeed with probability $S(t, z)/S(t, b)$. The exit barrier $z = b$ gives rise to the boundary conditions

$$S(t, z) = S(t, b), \quad D_z S(t, z) = 0, \quad z \leq b. \quad (15)$$

There are no producers below $z = b$ and exit at this barrier means that the density at $z = b$ is also zero. Given an initial condition $S(0, z)$, the equations (14) and (15) determine how the measure of surviving producers $S(t, z)$ evolves over time.

Evaluating (14) at the exit barrier $z = b$ and using the boundary conditions (15) gives

$$D_t S(t, b) = E_t + \frac{1}{2} \sigma^2 D_{zz} S(t, b). \quad (16)$$

The density at the exit barrier is $D_z S(t, b) = 0$, and so the slope $-D_{zz} S(t, b)$ of this density must be positive. The population of survivors changes according to $D_t S(t, b)$, and (16) means that $-\frac{1}{2} \sigma^2 D_{zz} S(t, b)$ is the flow of incumbents exiting at the exit barrier $z = b$.

3.1.1 Stationary Distributions

If the exit rate is constant, then the flow of entrants $E_t = E e^{\eta t}$, with E and η positive, implies a population growth rate that converges to η in the long run. Conjecture that there is a balanced growth path of the form

$$S(t, z) = M_t R(z), \quad M_t = M e^{\eta t},$$

for some population level M and some right cumulative distribution function $R(z)$, both to be determined. Then (14) and (15) can be written as

$$\eta R(z) = -\mu DR(z) + \frac{1}{2}\sigma^2 D^2 R(z) + \left(\frac{E}{M} + \beta [1 - R(z)] \right) R(z),$$

for all $z \geq b$ together with the boundary conditions $R(b) = 1$ and $DR(b) = 0$. A distribution function $1 - R(z)$ that solves this second-order differential equation for some population level M defines a balanced growth path. Evaluating the differential equation at $z = b$ gives the stationary analog to (16),

$$\eta = \frac{E}{M} + \frac{1}{2}\sigma^2 D^2 R(b).$$

This emphasizes the fact that the entry rate E/M is jointly determined with the distribution $R(z)$. The slope of the density at $z = b$ implies an exit rate $\frac{E}{M} - \eta$.

Note that $R(z)$ shows up in linear terms on both sides of the above differential equation, and the coefficient E/M that appears on the right-hand side is determined jointly with $R(z)$. It will be convenient to define

$$\delta = \beta + \frac{E}{M} - \eta. \tag{17}$$

This is the sum of the rate at which incumbent producers imitate, and the exit rate $E/M - \eta$. Since the exit rate must be non-negative, this implies that $\delta \geq \beta$. The differential equation for $R(z)$ can now be written as

$$0 = -\mu DR(z) + \frac{1}{2}\sigma^2 D^2 R(z) + \delta R(z) - \beta R^2(z), \tag{18}$$

for all $z \geq b$, together with the boundary conditions

$$R(b) = 1, \quad DR(b) = 0. \tag{19}$$

Any combination of a distribution function $1 - R(z)$ and an entry parameter $\delta \geq \beta$ that satisfies (18)-(19) defines a balanced growth path. Note that $\delta \geq \beta$ implies that $\delta R(z) - \beta R^2(z)$ is non-negative for any $R(z) \in [0, 1]$. Also, $DR(z)$ has to be non-positive since $1 - R(z)$ is supposed to be a distribution function. If μ were non-negative, then (18) would imply that that $1 - R(z)$ is convex on (b, ∞) . This is not consistent with a stationary distribution with support (b, ∞) , and so we must assume that $\mu < 0$. In other words, stationary distributions on (b, ∞) can only be constructed if $\theta_E > \theta_I$.

To compare (18)-(19) with the fixed population case, note that taking $R(z) = 1 - F(z)$, $\mu = \theta_I - \theta_E$ and $\delta = \beta$ in (18) gives (12). In the fixed population case, θ_E was

jointly determined with the stationary distribution $F(z)$, and stationary distributions can be constructed for any $\theta_E \geq \theta_I + \sigma\sqrt{2\beta} = \theta_I + \sigma\sqrt{2\delta}$. Compactly supported initial conditions imply the particular long-run growth rate $\theta_E = \theta_I + \sigma\sqrt{2\delta}$. Here, $\mu = \theta_I - \theta_E$ is taken as given and $\delta > \beta$ and $R(z)$ are jointly determined. Luttmer [2007] explicitly solves (18)-(19) in the special case $\beta = 0$ and finds that compactly supported initial distributions converge to the stationary distribution associated with $\theta_E = \theta_I + \sigma\sqrt{2\delta}$. The formula $\theta_E = \theta_I + \sigma\sqrt{2\delta}$ therefore describes the limiting growth rate in the two special cases $\delta = \beta$ and $\delta > \beta = 0$, suggesting that the same may be true for $\beta \in (0, \delta)$.

3.1.2 Constructing Stationary Distributions

The following adapts some of the initial steps in Kolmogorov, Petrovskii and Piskunov [1937] and McKean [1975] to account for exit. Fix some exit boundary $b > -\infty$ and exit parameter $\delta > \beta$. Write $f(z) = -DR(z)$ for the density at z . The differential equation (18)-(19) is equivalent to the first-order system

$$DR(z) = -f(z) \tag{20}$$

$$Df(z) = \frac{\mu f(z) + [\delta - \beta R(z)]R(z)}{\sigma^2/2} \tag{21}$$

with the initial conditions $R(b) = 1$ and $f(b) = 0$. This differential equation is autonomous and so the solution only depends on $z - b$. Note that the differential equation becomes linear when $\beta = 0$. This is the case solved in Luttmer [2007].

Starting from any initial conditions, the pair of differential equations (20)-(21) implies trajectories $[R, f]$ that satisfy $dR/df = -(\mu f + (\delta - \beta R)R)/(f\sigma^2/2)$ for all $f > 0$. Only solutions that satisfy $R(z) \in [0, 1]$ and $f(z) \geq 0$ can be interpreted as stationary distributions. Observe that

$$\begin{aligned} DR(z) = 0 &\Rightarrow f(z) = 0, \\ Df(z) = 0 &\Rightarrow f(z) = -\frac{1}{\mu}[\delta - \beta R(z)]R(z). \end{aligned}$$

Since $\mu < 0$, the $Df(z) = 0$ curve is a parabola with a maximum at $\delta/(2\beta)$ and zeros at $R(z) = 0$ and $R(z) = \delta/\beta > 1$. As illustrated in Figure 2 below, this implies the stationary points $[0, 0]$ and $[\delta/\beta, 0]$. Under the alternative assumptions $b = -\infty$ and $\delta = \beta$, these stationary points would correspond to $[R(\infty), f(\infty)] = [0, 0]$ and $[R(-\infty), f(-\infty)] = [1, 0]$. In that case, stationary distributions are associated with unstable trajectories out of the saddle point $[1, 0]$ that converge to $[0, 0]$ (Kolmogorov,

Petrovskii and Piskunov [1937] and McKean [1975].) Here $b > -\infty$ and $\delta > \beta$ implies that the initial condition $[R(b), f(b)] = [1, 0]$ is not a stationary point of the differential equation. At this point, (21) yields $Df(b) > 0$, and hence a positive exit rate at the boundary b .

To examine the differential equation near the stationary point $[0, 0]$, linearize (20)-(21) to obtain

$$D \begin{bmatrix} R(z) \\ f(z) \end{bmatrix} \approx \begin{bmatrix} 0 & -1 \\ \frac{\delta}{\sigma^2/2} & \frac{\mu}{\sigma^2/2} \end{bmatrix} \begin{bmatrix} R(z) \\ f(z) \end{bmatrix}. \quad (22)$$

The imitation parameter β multiplies a quadratic term in (21) and so it drops out from the linearization. The eigenvalues of the matrix on the right-hand side of (22) are

$$\lambda_- = \frac{\mu}{\sigma^2} - \sqrt{\left(\frac{\mu}{\sigma^2}\right)^2 - \frac{\delta}{\sigma^2/2}}, \quad \lambda_+ = \frac{\mu}{\sigma^2} + \sqrt{\left(\frac{\mu}{\sigma^2}\right)^2 - \frac{\delta}{\sigma^2/2}}.$$

Since $\mu < 0$, both eigenvalues are negative if real. Complex eigenvalues result in trajectories that spiral around the stationary point $[0, 0]$. Along such trajectories, $R(z)$ and $f(z)$ become negative, and so parameters that gives rise to complex eigenvalues are not associated with stationary distributions. Since we also have $\delta > \beta$, this means that δ must satisfy

$$\beta < \delta \leq \frac{1}{2} \left(\frac{\mu}{\sigma}\right)^2. \quad (23)$$

The first inequality arises because exit rates are positive, and the second is necessary for the existence of a stationary distribution. Observe the upper bound for δ together with $\mu < 0$ implies that $\sigma\sqrt{2\delta} \leq -\mu = \theta_E - \theta_I$, and thus $\theta_E \geq \theta_I + \sigma\sqrt{2\delta}$. This is the bound we already have in (13) for the special case $\delta = \beta$, and in Luttmer [2007] for the special case $\delta > \beta = 0$.

Figure 2 shows the phase diagram for the case $\delta/(2\beta) < 1 < \delta/\beta$. Included are the tangent to the curve defined by $Df(z) = 0$ at $R(z) = 0$, the threshold $R(z) \leq 1$, and the trajectory that starts at $[R(b), f(b)] = [1, 0]$. For any trajectory $[R, f]$ that converges to the stationary point $[0, 0]$,

$$\lim_{R \rightarrow 0} \frac{df}{dR} = -\frac{1}{\sigma^2/2} \left(\mu + \frac{\delta}{\lim_{R \rightarrow 0} f/R} \right).$$

Thus $s = \lim_{R \rightarrow 0} df/dR = \lim_{R \rightarrow 0} f/R$ must satisfy $s\sigma^2/2 = -\mu - \delta/s$. Solving this yields $s \in \{-\lambda_-, -\lambda_+\}$. That is, the slope of any trajectory converging to the origin is just the negative of one of the two eigenvalues of the linearized system. Near the origin, trajectories that start above the line $[f, -\lambda_- f]$ curve around the origin,

while trajectories that start below this line converge to the origin monotonically, with asymptotic slope $-\lambda_+$. Following the steps of Kolmogorov, Petrovskii and Piskunov [1937] and McKean [1975] one can show that any δ that satisfies (23) is associated with a stationary distribution that solves (20)-(21) starting from $[R(b), f(b)] = [1, 0]$, and that the trajectory that starts with initial values $[R(b), f(b)] = [1, 0]$ converges to $[0, 0]$ with slope $-\lambda_+$. Since $f(z) = -DR(z)$, it therefore follows that

$$\lim_{z \rightarrow \infty} \left(-\frac{DR(z)}{R(z)} \right) = -\lambda_+ = -\frac{\mu}{\sigma^2} - \sqrt{\left(\frac{\mu}{\sigma^2}\right)^2 - \frac{\delta}{\sigma^2/2}}. \quad (24)$$

That is, the stationary distribution behaves like an exponential distribution for large z , with scale parameter $-\lambda_+$. The fact that the distribution of z behaves like an exponential distribution means that the distribution of e^z behaves like a Pareto distribution. When δ is at the upper bound implied by (23), the tail index of this Pareto distribution simplifies to $-\mu/\sigma^2$, as in Luttmer [2007].

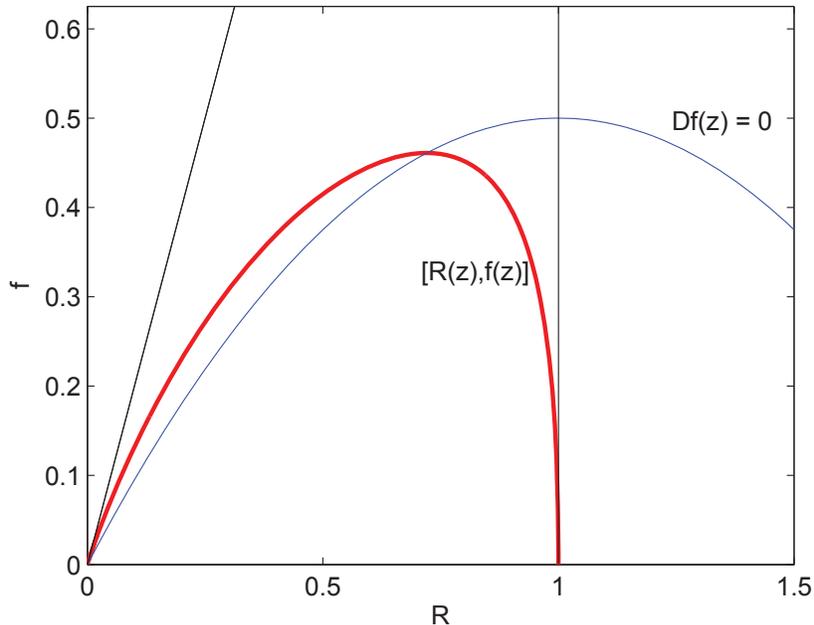


FIGURE 2 Phase Diagram for $[DR(z), Df(z)]$

Remarkably, the tail index (24) does not depend on the imitation Poisson rate β , in contrast to both the case of deterministic innovation studied in Alvarez, Buera and Lucas [2007] and Lucas and Moll [2011], and the case of Brownian innovation without exit presented above. Because the linearized differential equation (22) near the stationary point

$[0, 0]$ does not depend on $\beta < \delta$, the tail index (24) of the stationary distribution does not depend on β either. Given $\mu < 0$ and $\delta > \beta$, the precise rate at which incumbents imitate has no effect on the shape of the right tail of the productivity distribution. The condition $\delta > \beta$ says that there is exit. The economy with no exit and productivities distributed throughout $(-\infty, \infty)$ has $\delta = \beta$, and with this (24) does again depend on β .

4. BALANCED GROWTH

The economy with a fixed population in Section 2 can be interpreted as an economy with agents who produce for their own consumption. The growing exit barrier $b + \theta_E t$ in Section 3 requires a more intricate interpretation. The following takes producers to be firms and embeds the imitation mechanism of Section 3 in an economy very similar to Luttmer [2007]. This yields equilibrium conditions that explain why firms exit and why aggregate productivity grows at the rate θ_E , along a balanced growth path.

4.1 Consumers

Consider an economy with a growing population $H_t = H e^{nt}$ of consumers whose dynastic preferences over per-capita consumption flows C_t/H_t are determined by

$$\int_0^\infty e^{-\rho t} H_t \ln(C_t/H_t) dt,$$

for some positive subjective discount rate ρ . Given complete markets, these preferences imply an interest rate equal to $\rho + D[C_t/H_t]/[C_t/H_t]$. At any point in time, consumers can choose to supply labor or choose to be an entrepreneur. Being an entrepreneur means looking for opportunities to create a new firm. If the opportunity to create a new firm arrives, the entrepreneur can randomly select an incumbent firm and adopt its technology, at no further cost. Abilities vary, and there is a atomless distribution of labor and entrepreneurial skills in the population. Labor skills determine the flow of labor services a consumer can supply, and entrepreneurial skills determine the Poisson rate at which a consumer can start new firms. Comparative advantage determines the services consumers choose to supply. Write q_t for the expected market value at time t of a randomly sampled firm, measured in units labor. Consumers for whom the product of q_t and the Poisson rate at which they are able to create new firms exceeds 1 will choose to be entrepreneurs, and all other consumers will choose to be workers. Given a talent distribution, one can compute the supplies of labor and entrepreneurial services,

$$L_t = H_t L(q_t), \quad E_t = H_t E(q_t). \tag{25}$$

The assumption that the talent distribution is smooth ensures that $L(\cdot)$ and $E(\cdot)$ are smooth. Assuming the talent distribution has a mean, both will be bounded. The function $L(\cdot)$ is decreasing and $E(\cdot)$ is increasing. An increase in the expected value of trying to start up a firm induces more agents to switch from supplying labor to being an entrepreneur. Along a balanced growth path, q_t will be constant at some equilibrium level q , and hence L_t and E_t will grow at the same rate as the population.

4.2 Firms

Firms are production functions that use labor, subject to decreasing returns to scale. Following entry, a fixed cost must be paid continuously to continue the firm.

4.2.1 Static Decisions

A type- k firm is a Cobb-Douglas technology that can be used to produce a flow of $[k/(1-\alpha)]^{1-\alpha}[l/\alpha]^\alpha$ units of consumption using l units of labor, where $\alpha \in (0, 1)$ is the factor share of labor. At any point in time t , there are many incumbent firms with firm-specific fixed factor levels k_t . These change according to

$$d \ln(k_t) = \theta_1 dt + \sigma dW_t + \Delta_t dN_t. \quad (26)$$

As before, W_t is a standard Brownian motion, N_t is a Poisson process with arrival rate β , and Δ_t is an increment obtained by allowing the firm to adopt the technology of another firm, sampled at random from the population of incumbent firms. Firms must continuously expend a flow of $\phi > 0$ units of labor to survive, inducing exit of firms that are not sufficiently productive. Wages at time t are denoted by w_t , in units of consumption per unit of time. A type- k_t firm facing wages w_t at time t chooses labor inputs l by solving

$$\max_l \left\{ \left(\frac{k_t}{1-\alpha} \right)^{1-\alpha} \left(\frac{l}{\alpha} \right)^\alpha - w_t l \right\}.$$

This implies labor inputs equal to

$$\frac{l_t}{\alpha} = \frac{1}{w_t^{1/(1-\alpha)}} \frac{k_t}{1-\alpha}.$$

After fixed costs, the resulting profits are

$$\frac{k_t}{w_t^{\alpha/(1-\alpha)}} - \phi w_t = \phi w_t (e^{z_t} - 1),$$

where the state variable z_t is given by

$$z_t = \ln(k_t) - \ln(\phi) - \frac{\ln(w_t)}{1 - \alpha}. \quad (27)$$

Thus z_t is normalized so that profits are zero when z_t is zero and $e^{z_t} - 1$ measures profits in units of labor. Labor inputs expressed in terms of z_t are $l_t = e^{z_t} \phi \alpha / (1 - \alpha)$. Up to an intercept, z_t measures log variable employment, a measure of firm size.

Let κ be the growth rate of wages along a balanced growth path. Then (26) and (27) imply that $dz_t = \mu dt + \sigma dW_t + \Delta_t dN_t$, where

$$\mu = \theta_I - \theta_E, \quad \theta_E = \frac{\kappa}{1 - \alpha}. \quad (28)$$

The drift θ_I is a parameter, and θ_E and κ are equilibrium variables associated with a balanced growth path. From Section 3 we know that (26)-(28) gives rise to stationary distributions if combined with exit at a lower threshold b for z_t and entry at randomly sampled points above b . As before, let $R(\cdot)$ denote the right tail of such a distribution and write $M_t = M e^{\eta t}$ for the number of firms along a balanced growth path. Stationary distributions must satisfy the differential equation (18)-(19) for some $\delta = \beta + E(q)H/M - \eta$, as in (17). Following the convergence considerations discussed in Section 3, select the particular stationary distribution implied by

$$\delta = \frac{1}{2} \left(\frac{\mu}{\sigma} \right)^2. \quad (29)$$

With this, there is a unique stationary distribution $R(\cdot)$, given the exit threshold b and the drift $\mu = \theta_I - \theta_E$.

4.2.2 Exit and Entry Decisions

The construction so far gives rise to profits and labor income that grow at a per-capita rate κ along a balanced growth path. It follows that per-capita consumption will also grow at the rate κ , and hence the interest rate must be $\rho + \kappa$. Since profits in units of consumption are $\phi w_t (e^{z_t} - 1)$ and w_t grows at the rate κ , it follows that the value of a firm in state z is $\phi w_t V(z)$, where

$$V(z) = \max_{\tau \geq 0} E_0 \left[\int_0^\tau e^{-\rho t} (e^{z_t} - 1) dt \right],$$

with z_t determined by (26)-(27) and the initial condition $z_0 = z$. The stopping time τ must depend on the information available to the firm—the aggregate state of the

economy and its own history. The Bellman equation for this value function is

$$\rho V(z) = e^z - 1 + \mu DV(z) + \frac{1}{2}\sigma^2 D^2V(z) + \beta \int_z^\infty [V(y) - V(z)] d[1 - R(y)], \quad (30)$$

where the optimal barrier b must satisfy

$$V(b) = 0, \quad DV(b) = 0. \quad (31)$$

The last term on the right-hand side of (30) describes the capital gains that arise from the arrival, at the rate β , of opportunities to imitate by randomly selecting another firm and copy its technology if it is more productive than z . The condition $V(b) = 0$ arises from the fact that exit leaves the firm with nothing, and $DV(b) = 0$ is a familiar smooth-pasting condition that is required for b to be optimal. Given μ and $R(\cdot)$, the Bellman equation (30)-(31) determines the value function V and the optimal threshold b below which this value function will be equal to zero. Note well that the distribution $R(\cdot)$ depends itself on a hypothesized exit barrier. In equilibrium, this exit barrier must be the same as the exit barrier b that follows from (30)-(31). Given μ , this results in a fixed point condition for b .

Potential entrants draw a state z from the distribution $R(z)$. The result will be a new firm with value $\phi V(z)$ in units of labor. The expected value of entry is thus

$$q = \phi \int V(y) d[1 - R(y)] \quad (32)$$

in units of labor. The supply of entrepreneurial services (25) combined with the definition $\delta = \beta + E(q)H/M - \eta$ imply that the per-capita number of firms can be written as

$$\frac{M}{H} = \frac{E(q)}{\eta + \delta - \beta}. \quad (33)$$

The amount of variable labor used by a firm in state z is $e^z \phi \alpha / (1 - \alpha)$, and the fixed labor requirement is ϕ . Aggregating over all firms gives the labor market clearing condition

$$L(q) = \frac{\phi M}{H} \int \left(1 + \frac{\alpha e^y}{1 - \alpha} \right) d[1 - R(y)]. \quad (34)$$

One can think of (33) as the steady-state supply of firms, and (34) as the demand for firms derived from the supply of labor $L(q)H$. The market clearing price is q , but q also has to satisfy the forward-looking condition (32). This implies an equilibrium condition that drives the determination of the drift parameter μ , via the dependence of $V(\cdot)$ and $R(\cdot)$ on this equilibrium parameter.

A balanced growth path is now defined by an exit barrier b , a drift parameter μ with an associated entry parameter δ determined by (29), the implied stationary distribution $1 - R(z)$ determined by (18)-(19), a value function $V(z)$ that satisfies (30)-(31), and a value of entry q defined as (32) that also satisfies the market clearing conditions (33)-(34). The level of consumption and wages at the initial date is determined by matching the stationary distribution $1 - R(z)$ with the time-0 distribution of fixed factors, using (27).

4.3 Constructing Balanced Growth Paths

Holding fixed $\mu = \theta_I - \theta_E$, the conditions for a balanced growth path can be reduced to a system of first-order differential equations. After solving these equations, one can back out the fixed cost parameter ϕ that is consistent with equilibrium.

The Bellman equation (30) depends on the expected gains from imitation. Given an imitation opportunity, these gains are given by

$$G(z) = \int_z^\infty [V(y) - V(z)] d[1 - R(y)]$$

for a firm with a current state z . The expected capital gains $\beta G(z)$ appear on the right-hand side of the Bellman equation (30), and the value of entry (32) is given by $q = \beta G(b)$. Note that $DG(z) = -DV(z)R(z)$. Since $V(z)$ is an increasing function, the integral that defines $G(z)$ converges if and only if $G(z) \rightarrow 0$ for large z . This is a boundary condition for $DG(z) = -DV(z)R(z)$ that can be used to solve for $G(z)$ given $V(z)$ and $R(z)$.

The construction is now as follows. Fix some $\mu < 0$ and let $\delta = (\mu/\sigma)^2/2$, as in (29). The stationary distribution is determined by

$$\begin{aligned} DR(z) &= -f(z), \\ Df(z) &= \frac{\mu f(z) + \delta R(z) - \beta R^2(z)}{\sigma^2/2}, \\ 1 &= R(b), \quad 0 = f(b). \end{aligned} \tag{35}$$

This determines $R(\cdot)$ and $f(\cdot)$ up to a shift b . The value function $V(\cdot)$ and the associated gains from imitation $G(\cdot)$ must satisfy

$$\begin{aligned} DV(z) &= U(z), \\ DU(z) &= \frac{\rho V(z) - [e^z - 1 + \beta G(z) + \mu U(z)]}{\sigma^2/2}, \\ DG(z) &= -U(z)R(z). \end{aligned} \tag{36}$$

This is a system of three first-order differential equations for $V(\cdot)$, $U(\cdot)$ and $G(\cdot)$. For fixed b , the value-matching and smooth-pasting conditions (31) provide two boundary conditions, and the requirement that $G(z) \rightarrow 0$ as $z \rightarrow \infty$ provides a third. But b remains to be determined. What is missing is the transversality condition for the infinite-horizon problem of a firm. Rather than impose this transversality condition directly, one can exploit the fact that for z very large, $V(z)$ must behave like the value of a firm with no imitation opportunities (because $G(z) \rightarrow 0$ as z becomes large) and no option to exit (because this option is very unlikely to be exercised as z becomes large). The value of such a hypothetical firm is simply the expected present value of $e^{\mu t + \sigma W_t} - 1$ discounted at the rate ρ . With this, the boundary conditions for (36) can be summarized as

$$\begin{aligned} 0 &= V(b) = DV(b), \\ 0 &= \lim_{z \rightarrow \infty} G(z) = \lim_{z \rightarrow \infty} \left\{ V(z) - \left(\frac{e^z}{\rho - (\mu + \frac{1}{2}\sigma^2)} - \frac{1}{\rho} \right) \right\}. \end{aligned} \quad (37)$$

This provides the four boundary conditions needed to determine $V(\cdot)$, $U(\cdot)$ and $G(\cdot)$, as well as the exit barrier b , from (36). To implement this, one can guess $[b, G(b)]$ and solve (36) starting from the initial values $V(b) = DV(b) = 0$ and $G(b)$. One can then search for the $[b, G(b)]$ for which the two large- z boundary conditions hold approximately. Starting values for this search can be obtained from the $[b, G(b)]$ that can be computed analytically for the special case of $\beta = 0$, no incumbent imitation, that appears in Luttmer [2007].

Given $\mu = \theta_I - \theta_E < 0$ and $\delta = (\mu/\sigma)^2/2$, this construction delivers a stationary distribution, a value function, and an implied value $q = \phi G(b)$ of entry. The market clearing conditions (33)-(34) then require that

$$\frac{L(\phi G(b))/\phi}{E(\phi G(b))} = \frac{1}{\eta + \delta - \beta} \int \left(1 + \frac{\alpha e^y}{1 - \alpha} \right) d[1 - R(y)]. \quad (38)$$

The right-hand side can be interpreted as the average demand for labor, measured in units of the fixed amount of labor needed to operate a firm, implied by a unit flow of entrants. The left-hand side is the supply of labor in the same units divided by the equilibrium flow of entrants—observe that this is a decreasing function of the amount of labor ϕ it takes to continue a firm. Imposing (38) gives an equilibrium condition for $\mu = \theta_I - \theta_E < 0$. More easily, one can conjecture a $\mu < 0$ and then use this condition to solve for the implicit fixed cost ϕ that makes μ an equilibrium.

4.4 A Numerical Example

Suppose θ_E is the equilibrium growth rate for an economy with $\beta = 0$, no incumbent imitation. From Luttmer [2007], this implies $\mu < 0$. Note that (29) implies $\delta = (\mu/\sigma)^2/2 > 0$. Fix θ_E , take some $\beta \in (0, \delta)$, and again solve for the value function $V(z)$, the gains from imitation $G(z)$, and the distribution $1 - R(z)$. Taking the value functions of other firms as given, the ability to imitate increases the value of a particular firm. Furthermore, although the exit threshold will be lower for $\beta > 0$ than for $\beta = 0$, firm productivities grow faster because of imitation, and this shifts much of the productivity distribution to the right. Both effects make for higher firm values. In particular, the value of entry $q = \phi G(b)$ will be higher when $\beta > 0$. This encourages more consumers to become entrepreneurs and reduces labor supply. At the same time, the average firm will be larger, and so the demand for labor will be higher. In addition, firms move away from the exit barrier more quickly, and this implies a lower exit rate $\delta - \beta$. A lower exit rate translates into a higher steady state number of firms, further increasing the demand for labor. There will be an excess demand for labor, implying that the original growth rate θ_E associated with $\beta = 0$ is no longer an equilibrium.

It is easy to see that the new balanced growth rate θ_E must be higher. This lowers the rate $\mu = \theta_I - \theta_E$ at which incumbent productivity grows by itself, relative to entry productivity. This makes incumbent firms smaller—in particular, the tail index $-\mu/\sigma^2$ increases, implying a thinner right tail. The value of incumbent firms will be lower, because of the weaker growth prospects of individual firms, and because the population that can be imitated is less productive. Such a decline in the value of incumbent firms is precisely what is needed to help clear the labor market. More consumers will choose to be workers rather than entrepreneurs, and firms are less productive and demand less labor. Both effects work to reduced the excess demand for labor implied by the growth rate of an economy without imitation.

Quantitatively, everything will hinge on the supply functions $E(\cdot)$ and $L(\cdot)$ of entrepreneurial services and labor services. If these supplies are relative elastic, then it will take only a small change in the value of entry to clear the labor market, and hence only a small change in the equilibrium growth rate of the economy. If these supplies are completely inelastic, then there is an excess demand for labor at the original growth rate simply because the average firm is larger with imitation than without imitation, and not because anyone switches from supplying labor services to supplying entrepreneurial services. In that case, all the adjustment in the labor market clearing condition (38) has to come from an increase in the equilibrium exit rate $\delta - \beta = \frac{1}{2} \left(\frac{\mu}{\sigma}\right)^2 - \beta$ and an increase

in the tail index $-\mu/\sigma^2$ that will result from a higher value of $-\mu = \theta_E - \theta_I$.

To illustrate this in a numerical example, it is easiest computationally to start with an equilibrium with incumbent imitation and then determine the equilibrium for an alternative economy with $\beta = 0$. Figure 3 shows the value functions $V(z)$ and the gains from imitation $G(z)$ associated with $\rho = 0.05$, $\mu = -0.02$ and $\sigma = 0.10$, for an economy with $\beta > 0$ and an economy with $\beta = 0$. Note that these parameters imply a tail index of $-\mu/\sigma^2 = 2$.⁷ In the economy with imitation, the imitation parameter is set equal to $\beta = \delta/2$, where $\delta = (\mu/\sigma)^2/2$. This implies $\beta = 0.01$. Both sets of functions are determined by solving (35)-(36). As expected, the value function for $\beta > 0$ dominates the one for $\beta = 0$, and both converge to the present value of $e^z - 1$ when z is large.

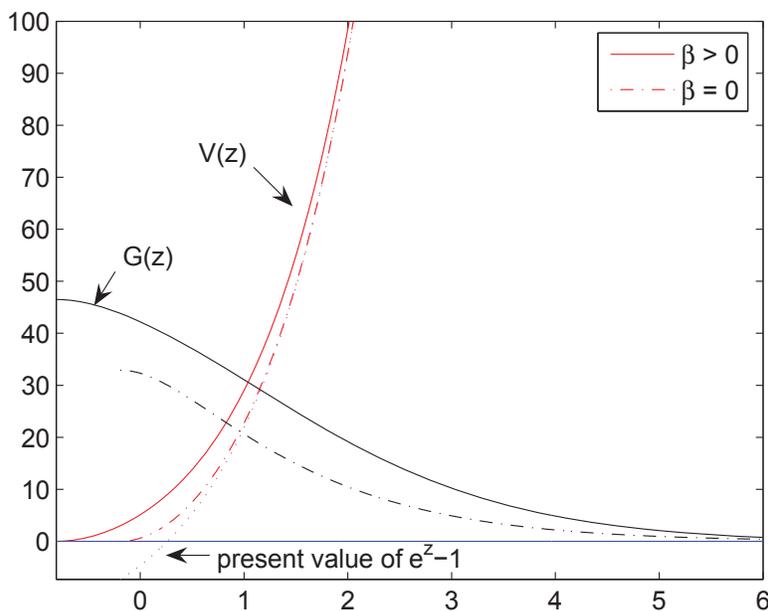


FIGURE 3 Value Functions and Gains from Imitation

Now suppose the μ used in Figure 3 is the equilibrium value of $\mu = \theta_I - \theta_E$ for an economy with imitation. Suppose the supplies of entrepreneurial and labor services are inelastic. With some abuse of notation, write the left-hand side of the equilibrium condition (38) as $L/(\phi E)$. Suppose the population growth rate is $\eta = 0.01$ and the factor share of variable

⁷In the data, this number is slightly above 1. The closer $-\mu/\sigma^2$ is to 1, the more important the far right tail of the productivity distribution is for calculating aggregates and present values. This makes numerical procedures more fragile and requires more rigorous testing of these procedures than performed for this draft. Results for $-\mu/\sigma^2$ close to 1 will be reported in a future draft.

labor equals $\alpha = 0.7$. Then one can calculate the right-hand side of (38) and thus infer $L/(\phi E)$. This gives a number of close to 1,000. Since $\eta + \delta - \beta = 0.02$, this implies that $(\eta + \delta - \beta)L/(\phi E) \approx 20$. That is, fixed costs amount to roughly 5% of firm employment in the economy with incumbent imitation. Now hold fixed $L/(\phi E)$ and compute the equilibrium for $\beta = 0$ by solving (38) for μ , using the density $f(z) = \delta^2(z - b)e^{-\delta(z-b)}$ with $\delta = (\mu/\sigma)^2/2$ and the exit threshold b determined by the optimal stopping problem for μ and $\beta = 0$, obtained in Luttmer [2007]. This gives $\mu = -0.0164$, implying a tail index equal to 1.64. The change from the $\beta > 0$ equilibrium $\mu = -0.02$ to the $\beta = 0$ equilibrium $\mu = -0.0164$ implies a reduction in the growth rate $\theta_E = \theta_I - \mu$ of the economy equal to 0.36% per annum. As expected, incumbent imitation speeds up growth in this economy.

The resulting productivity distributions are shown in Figure 4.

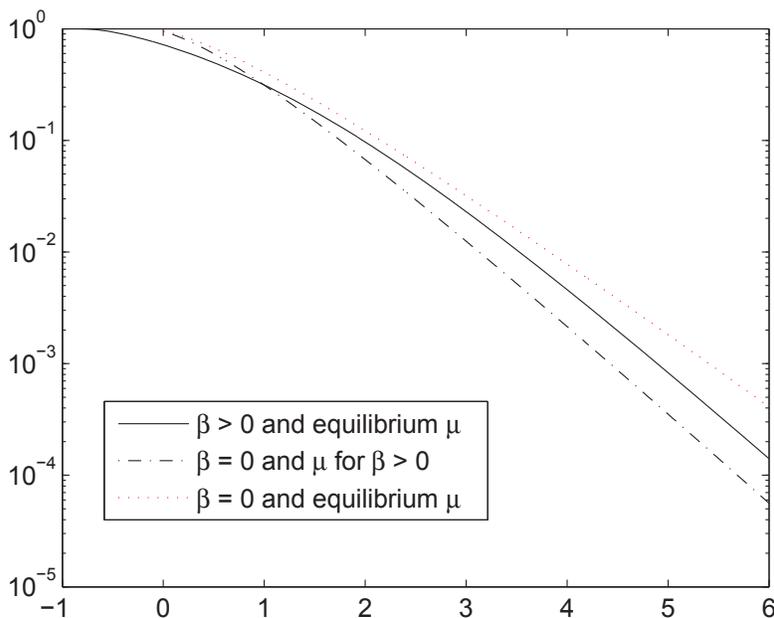


FIGURE 4 Right Cumulative Distribution Functions

The solid curve is the equilibrium distribution for $\beta > 0$. Holding fixed μ and solving the differential equation (35)-(36) again for $\beta = 0$, analytically using Luttmer [2007], gives a productivity distribution with the same tail index, but with a higher exit threshold b . In equilibrium, the economy with $\beta = 0$ and the same $L/(\phi E)$ will then have a lower growth rate θ_E and hence a $\mu = \theta_I - \theta_E < 0$ that is closer to zero. The result is a productivity distribution with a thicker right tail, as shown in Figure 4.

It goes without saying that the improvement in long-run growth rates generated by incumbent imitation ignores the fact that there may be negative effects on the incentives of firms to innovate. Here, θ_1 is an exogenous parameter. Luttmer [2010] suggests a model of managerial effort choice in which this parameter becomes endogenous.

5. CONCLUDING REMARKS

The results in this paper suggests that an economy starting out with a narrow range of initial productivities will, after some period of transition, settle down on a balanced growth path along which aggregate productivity grows faster than the average rate at which individual producers are able to innovate. Both noisy innovation and imitation are essential. Without noise, imitation will eventually make all producers look alike. Without imitation, aggregate growth is limited by the rate at which individual firms can innovate. An interesting quantitative question that arises is: how long does an economy with limited initial heterogeneity linger before it locks on to a path of stable growth?

Although incumbent firms can imitate, their imitation is not frictionless. It is delayed by their undirected search, and by the finite Poisson rate at which imitation opportunities arrive. This delay allows firm productivities to be widely dispersed. In turn this dispersion gives rise to a thick-tailed distribution of firm employment, as observed in the data. Luttmer [2011] gives an alternative interpretation of the employment size distribution of firms that invokes heterogeneous histories of capital accumulation instead of heterogeneous histories of productivity growth. In such an economy, rampant imitation can lead to very similar productivities, without necessarily implying a counterfactual lack of heterogeneity in firm size. Non-rival ideas may be easy to copy, but accumulating the rival capital needed to implement these ideas takes time.

REFERENCES

- [1] Alvarez, F., F. Buera, and R.E. Lucas, Jr., “Models of Idea Flows,” *NBER working paper no. 14135* (2008).
- [2] Atkeson, A. and A.T. Burstein, “Aggregate Implications of Innovation Policy,” UCLA working paper (2011).
- [3] Bramson, M., “Convergence of Solutions of the Kolmogorov Equation to Travelling Waves,” *Memoirs of the American Mathematical Society*, vol. 44, no. 285 (1983).

- [4] Cavalli-Sforza, L.L. and M.W. Feldman, *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton University Press (1981).
- [5] Chatterjee, S. and E. Rossi-Hansberg, "Spin-Offs and the Market for Ideas," *NBER working paper 13198* (2007).
- [6] Cox, D.R. and H.D. Miller, *The Theory of Stochastic Processes*, Chapman & Hall/CRC (1965).
- [7] Eeckhout, J. and B. Jovanovic, "Knowledge Spillovers and Inequality," *American Economic Review*, vol. 92, no. 5 (2002), 1290-1307.
- [8] Fogli, A. and L. Veldkamp, "Germs, Social Networks and Growth," working paper, New York University and University of Minnesota (2011).
- [9] Fisher, R.A., "The Wave of Advance of Advantageous Genes," *Annals of Eugenics*, vol. 7 (1937), 355-369.
- [10] Jovanovic, B., and R. Rob, "The Growth and Diffusion of Knowledge," *Review of Economic Studies*, vol. 56 no. 4 (1989), 569-82.
- [11] Kolmogorov, A.N., I.G. Petrovskii and N.S. Piskunov "A Study of the Diffusion Equation with Increase in the Amount of Substance, and its Application to a Biological Problem," in *Selected Works of A. N. Kolmogorov I*, V. M. Tikhomirov, Editor, Kluwer Academic Publishers 1991 (1937).
- [12] Kortum, S., "Research Patenting and Technological Change," *Econometrica*, vol. 65, no. 6 (1997), 1389-1419.
- [13] König, M., J. Lorenz, and F. Zilibotti, "Innovation vs Imitation and the Evolution of Productivity Distributions," CEPR Discussion Papers 8843 (2012).
- [14] Lucas, R.E., Jr. and B. Moll, "Knowledge Growth and the Allocation of Time," *NBER working paper no. 17495* (2011).
- [15] Luttmer, E.G.J., "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, Vol. 122, No. 3 (2007), 1103-1144.
- [16] Luttmer, E.G.J., "Models of Growth and Firm Heterogeneity," *Annual Review of Economics*, vol. 2 (2010), 547-576.

- [17] Luttmer, E.G.J., “On the Mechanics of Firm Growth,” *Review of Economic Studies*, vol. 78 (2011), 1042–1068.
- [18] McKean, H.P., “Applications of Brownian motion to the Equation of Kolmogorov-Petrovskii-Piskunov,” *Communications in Pure and Applied Mathematics*, vol. 28 (1975), 323–331; “Erratum,” vol. 29 (1976), 553-554.
- [19] Perla, J. and C. Tonetti, “Equilibrium Imitation and Growth,” working paper, New York University (2012).