

Federal Reserve Bank of Minneapolis
Research Department

Four Models of Knowledge Diffusion and Growth*

Erzo G.J. Luttmer

Working Paper 724

May 2015

ABSTRACT _____

This paper describes how long-run growth emerges in four closely related models that combine individual discovery with some form of social learning. In a large economy, there is a continuum of long-run growth rates and associated stationary distributions when it is possible to learn from individuals in the right tail of the productivity distribution. What happens in the long run depends on initial conditions. Two distinct literatures, one on reaction-diffusion equations, and another on quasi-stationary distributions suggest a unique long-run outcome when the initial productivity distribution has bounded support.

Keywords: Growth; Knowledge diffusion

JEL classification: O33

*Luttmer: University of Minnesota. Comments 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

This paper brings together the key ingredients of four related models of knowledge diffusion and long-run growth that were developed in Luttmer [2007, 2012a, 2012b, 2015]. The reader is referred to the original papers for fully specified economies, and for more context, motivation, and important references to the literature.

Suppose there are many agents with different productivities who can learn from each other. Such social learning tends to shift the distribution of productivities to the right. In three of the four models considered here (not Luttmer [2012a]), and in related models elsewhere (Lucas and Moll [2014] is a prominent example), this results in a continuum of long-run equilibrium growth rates and a continuum of stationary detrended productivity distributions. In particular, it can happen that any non-negative growth rate is a possible long-run equilibrium even though nobody is making any new discoveries. As will be illustrated with some simple examples, this occurs when the initial distribution of productivities has a sufficiently fat tail, so that the opportunities to learn from others are never exhausted. A natural way to avoid such a scenario is to consider only balanced growth paths that are accessible from initial productivity distributions with bounded support. As will be laid out in detail, this selects the balanced growth path with the lowest possible growth rate in Luttmer [2007, 2012b, 2015]. There is long-run growth in these economies because agents not only learn from each other, but also continuously make small independent discoveries. New discoveries that can spread in a population are essential for growth to continue in the long run.

Section 2 describes some simple examples of social learning that can be analyzed using only the most elementary tools. They are the easiest way to understand the multiplicity of growth rates and stationary distributions that arises more generally in large economies. The differential equations that govern these examples are degenerate special cases of equations that appear later. Section 3 presents a basic example, inspired by the discussion of scale effects in Staley [2011], of what happens when one combines social learning with individual discovery in a finite economy. A closely related large economy is introduced in Section 4, following Staley [2011] and Luttmer [2012b, 2015]. Section 5 considers the basic elements of an economy with entry and exit in which only new entrants can learn from incumbent producers, as in Luttmer [2007]. The same multiplicity issues arise and Section 5 recaps the solution proposed in Luttmer [2007].

Section 6 also considers entry and exit but limits the ability of entrants to learn from incumbents in the right tail of the productivity distribution. This produces a unique stationary distribution of productivities, as in Luttmer [2012a].

2. LEARNING FROM OTHERS

Useful knowledge does not seem to spread instantaneously. When some are more knowledgeable than others, it may take time to find out who has the most useful knowledge. Or it may be clear who knows more, but learning from someone who is more knowledgeable takes time. If such learning also requires the cooperation of the more knowledgeable individual, and if this individual faces a capacity constraint, then everyone cannot catch up at the same time. The following describes how useful knowledge spreads with these two alternative sources of delay. For more on some of these and related examples, see Alvarez, Buera and Lucas [2008].

2.1 Search Delays

Consider a large population of producers with heterogeneous productivities. Write $P(t, y)$ for the cross-sectional distribution of productivity states at time t . Suppose any producer can randomly select someone else in the population at an average rate β and instantaneously adopt the technology of the randomly selected producer if it is more productive. Producers in $(-\infty, y]$ therefore exit this interval by sampling producers in (y, ∞) at an average rate $\beta[1 - P(t, y)]$. Since $P(t, y)$ is the number of such producers, this implies¹

$$D_t P(t, y) = -\beta P(t, y)[1 - P(t, y)]. \quad (1)$$

This is a system of logistic differential equations in t , one for each y . The unique solution is the logistic function

$$P(t, y) = \frac{1}{1 + \left(\frac{1}{P(0, y)} - 1\right) e^{\beta t}}. \quad (2)$$

The initial distribution $P(0, y)$ matters a lot. For example, if $P(0, u) = 1$ for some $u \in (-\infty, \infty)$ then $P(t, u) = 1$ for all $t > 0$. On the other hand, if $P(0, y) < 1$ for all y , then $P(t, y)$ will forever decline. That is, the distribution of productivities will forever move to the right.

2.1.1 Many Stationary Distributions

The economy has a large population of interacting producers, and the cross-sectional distribution of their productivities is updated over time. After some form of de-trending, these cross-sectional distributions may become time-invariant. With a certain abuse of terminology, refer to those time-invariant cross-sectional distributions as stationary

¹Throughout, $Df(x)$ is the derivative of $f(x)$ and $D_y g(y, z) = \partial g(y, z)/\partial y$.

distributions. Any point mass at a single point is a stationary distribution without any de-trending: $P(t, y) = P(0, y)$ for all y and all $t > 0$. But with some form of de-trending, one can construct many more stationary distributions.

One class of stationary distributions is obtained by conjecturing that $P(t, y) = F(y - \kappa t)$ for some strictly positive κ . With this conjecture, (1) implies

$$\kappa DF(y) = \beta F(y)[1 - F(y)]. \quad (3)$$

This is another logistic differential equation. Without even solving it, note that

$$\lim_{y \rightarrow \infty} \frac{DF(y)}{1 - F(y)} = \frac{\beta}{\kappa}.$$

This means the right tail $1 - F(y)$ will behave like the exponential $e^{-(\beta/\kappa)y}$, as confirmed by the explicit solution

$$F(y) = \frac{1}{1 + \left(\frac{1}{F(0)} - 1\right) e^{-(\beta/\kappa)y}}, \quad y \in (-\infty, \infty). \quad (4)$$

So the initial distribution $P(0, y) = F(y)$ will result in a linear trend κt .

A second class of stationary distributions is obtained by conjecturing that $P(t, y) = F(ye^{-\kappa t})$ for some positive κ . This yields

$$\kappa y DF(y) = \beta F(y)[1 - F(y)] \quad (5)$$

and hence

$$\lim_{y \rightarrow \infty} \frac{y DF(y)}{1 - F(y)} = \frac{\beta}{\kappa}.$$

That is, $F(y)$ will be a power law: its right tail $1 - F(y)$ will behave like the power function $y^{-(\beta/\kappa)}$. The explicit solution is

$$F(y) = \frac{1}{1 + \left(\frac{1}{F(1)} - 1\right) y^{-\beta/\kappa}}, \quad y \in (0, \infty). \quad (6)$$

Of course, the two stationary distributions (4) and (6) are linked via an exponential change of variables. But the point is that choosing $P(0, y)$ to be of the form (4) or (6) determines whether growth will be linear or exponential. More generally, the right tail properties of $P(0, y)$ will determine the type of growth that emerges in the long-run.

2.1.2 Relation to Extreme Value Theory

The reader will have noticed the similarity between (4) and (6) on the one hand, and, respectively, the Gumbel $F(y) = \exp(-e^{-y})$ and Fréchet $F(y) = \exp(-y^{-\alpha})$ distributions of extreme value theory on the other. This is not a coincidence. Suppose that when producers i and j meet at time t , they both adopt $\max\{y_{i,t}, y_{j,t}\}$. The $y_{i,t}$ and $y_{j,t}$ are themselves the result of taking maxima in previous meetings, and so on. Such a process of random meetings and taking maxima is recursive. Over time, it amounts to taking the maximum of an exponentially growing number of random draws from the initial distribution $P(0, y)$.

To make this explicit, observe that the maximum of N_t independent draws $\{y_n\}_{n=1}^{N_t}$ from some distribution F will have a distribution

$$\Pr[\max\{y_1, \dots, y_{N_t}\} \leq y | N_t] = [F(y)]^{N_t}.$$

Suppose now that N_t follows a pure birth process with birth rate βN_t . Then N_t follows a geometric distribution,

$$\Pr[N_t = n] = e^{-\beta t}(1 - e^{-\beta t})^{n-1}, \quad n \in \mathbb{N}.$$

The resulting marginal distribution of the maximum of the number of draws obtained by time t is

$$\begin{aligned} \Pr[\max\{y_1, \dots, y_{N_t}\} \leq y] &= \sum_{n=1}^{\infty} e^{-\beta t}(1 - e^{-\beta t})^{n-1} [F(y)]^n \\ &= \frac{e^{-\beta t} F(y)}{1 - (1 - e^{-\beta t}) F(y)} = \frac{1}{1 + \left(\frac{1}{F(y)} - 1\right) e^{\beta t}}. \end{aligned}$$

This is just the solution (2) of (1), with $P(0, y) = F(y)$. Thus, if $F(y)$ is given by (4), then $\max\{y_1, \dots, y_{N_t}\} - \kappa t$ will have the same distribution. Alternatively, if $F(y)$ is given by (6) then $\max\{y_1, \dots, y_{N_t}\} e^{-\kappa t}$ will also have that distribution. As in Kortum [1997], exponential growth results from an exponentially growing number of draws from a distribution that follows a power law.

As in the case of extreme-value theory, one can also consider an initial distribution $P(0, z)$ with bounded support and scale the ever decreasing gap between the sample maximum and the upper bound of the support of $P(0, z)$ to obtain a time-invariant distribution similar to the Weibull distribution.

2.2 Learning Delays

Delays in learning from others may also happen because not every “student” can be learning from the same “teacher” at the same time, and learning itself takes time. Suppose again we have a large population of producers with different productivities, and assume that they can form student-teacher pairs. In any pair, the student learns what the teacher knows after a random delay that is exponentially distributed with mean $1/\beta$. Everyone below the mean of the productivity distribution is a student, and everyone above the median is a teacher. The precise assignment of which student is matched with which teacher does not matter for keeping track of the cross-sectional distribution of productivities. See Luttmer [2015] for a market mechanism that implements this.

Let x_t be the median of the cross-sectional distribution $P(t, \cdot)$ at time t . For $x < x_t$, students in $(-\infty, x]$ exit this interval at the rate β . For $y > x_t$, teachers in (y, ∞) succeed in lifting students out of $(-\infty, x_t] \subset (-\infty, y)$ at the same rate β . As illustrated in Figure 1, this yields $D_t P(t, x) = -\beta P(t, x)$ for all x below the median and $D_t P(t, y) = -\beta[1 - P(t, y)]$ for all y above the median of $P(t, \cdot)$.

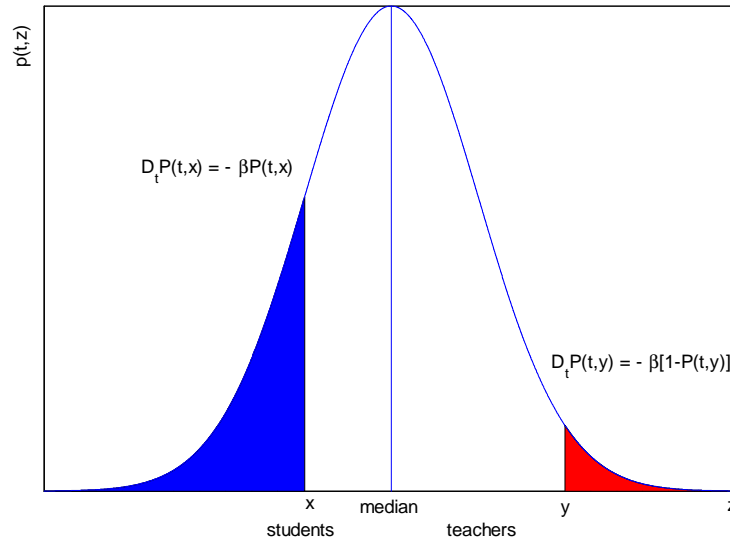


FIGURE 1 Social Learning Delays

Another way to write this piecewise linear differential equation is

$$D_t P(t, y) = -\beta \min \{P(t, y), 1 - P(t, y)\} \quad (7)$$

Note the similarity between (1) and (7): the right-hand side of the differential equation is $-\beta$ times a hump-shaped function of $P(t, y) \in [0, 1]$ that is equal to zero at the

endpoints $P(t, y) = 0$ and $P(t, y) = 1$. In the case of search delays, the hump-shaped function is a quadratic function. Here it is a tent.

Observe that the median x_t can only move to the right over time. For all $s \in [0, t]$ and any $y > x_t$, the right tail $1 - P(s, y)$ will grow at the rate β . It follows that the median x_t is determined by

$$\frac{1}{2} = P(t, x_t) = e^{\beta t} [1 - P(0, x_t)]. \quad (8)$$

This immediately shows the critical role of the right tail of $P(0, \cdot)$ in determining the pace at which the median x_t moves to the right. As in the case of (1), the entire solution to (7) can be computed explicitly. One can verify that it is given by

$$P(t, y) = \begin{cases} e^{-\beta t} P(0, y), & y \in (-\infty, x_0), \\ \frac{1}{2} \frac{1/2}{e^{\beta t} [1 - P(0, y)]}, & y \in (x_0, x_t), \\ 1 - e^{\beta t} [1 - P(0, y)], & y \in (x_t, \infty). \end{cases} \quad (9)$$

Again, the initial distribution is critical. To construct stationary distributions, it is easiest to use the density $p(t, y) = D_y P(t, y)$ and note that it satisfies

$$D_t p(t, y) = \begin{cases} -\beta p(t, y), & y \in (-\infty, x_t), \\ +\beta p(t, y), & y \in (x_t, \infty). \end{cases} \quad (10)$$

Conjecture that $p(t, y) = f(y - \kappa t)$ for some positive κ and some density f with median 0. Writing $z = y - \kappa t$ gives

$$-\kappa D f(z) = \begin{cases} -\beta f(z), & z \in (-\infty, 0), \\ +\beta f(z), & z \in (0, \infty). \end{cases}$$

This is a piecewise linear differential equation. The continuous solution is

$$f(z) = f(0) \times \begin{cases} e^{(\beta/\kappa)z}, & z \in (-\infty, 0), \\ e^{-(\beta/\kappa)z}, & z \in (0, \infty). \end{cases}$$

Requiring f to integrate to 1 yields $f(0) = (\beta/\kappa)/2$. This defines a class of stationary distributions, indexed by the trend parameter κ . Figure 2 shows the time lapse $\{p(n\Delta, z)\}_{n=0}^N$ for some positive Δ and $p(0, z)$ equal to, respectively, a Gaussian distribution and a version of the above stationary distribution. Trivially, growth would eventually stop completely if the initial distribution had been taken to have bounded support. Figure 2 shows that even Gaussian initial conditions result in stagnation, while the double exponential initial conditions result in constant growth.

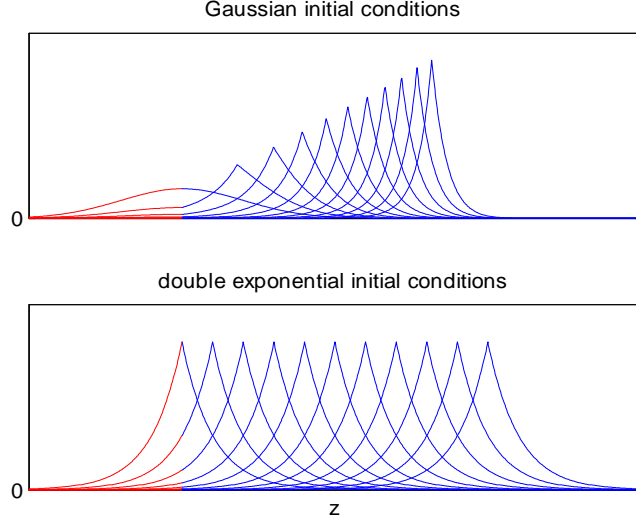


FIGURE 2 The Effect of Initial Conditions

As in the case of search delays, there is also a class of stationary distributions of the form $P(t, y) = F(ye^{-\kappa t})$, with densities $p(t, y) = f(ye^{-\kappa t})e^{-\kappa t}$. Write $z = ye^{-\kappa t}$. Then the differential equation (10) becomes

$$-\kappa z Df(z) - \kappa f(z) = \begin{cases} -\beta f(z), & z \in (0, 1), \\ +\beta f(z), & z \in (1, \infty). \end{cases}$$

The continuous solution is

$$f(z) = f(1) \times \begin{cases} z^{-1+\beta/\kappa}, & z \in (0, 1), \\ z^{-(1+\beta/\kappa)}, & z \in (1, \infty). \end{cases}$$

This defines a class of stationary distributions that are associated with exponential growth. To emphasize, whether the trend is linear or exponential depends entirely on the assumed initial conditions.

3. MEETING DELAYS IN A FINITE ECONOMY WITH INDIVIDUAL DISCOVERIES

Whether delay occurs because producers have to search for others who are more productive, or because one-on-one learning takes time, it is clear that there will be no long-run growth without individual discoveries if $P(0, z)$ has bounded support. In a finite population, the cross-sectional distribution of productivities necessarily has bounded support.

To introduce individual discoveries, consider an economy with only two producers, indexed by $j \in \{1, 2\}$, who meet infrequently, following random delays that are exponentially distributed with mean $1/\alpha$. When these producers meet, both have the opportunity to adopt the technology of the other. If each producer finds the other randomly at the rate β , then the meeting rate will be $\alpha = 2\beta$.

Write τ_n for the n^{th} meeting time, and suppose that in between meeting times τ_n and $\tau_{n+1} > \tau_n$, the productivities of the two producers wander around, independently and rather aimlessly, as a result of increments in two independent Brownian motions $W_{1,t}$ and $W_{2,t}$. More precisely, in between meetings the productivity state $y_{j,t}$ of producer $j \in \{1, 2\}$ evolves according to

$$dy_{j,t} = \sigma dW_{j,t}, \quad (11)$$

for some $\sigma > 0$. Define $y_t = \max\{y_{1,t}, y_{2,t}\}$. When they meet, the two producers compare notes and both will end up with this productivity state. The update of this state between meetings n and $n + 1$ is therefore given by

$$y_{\tau_{n+1}} - y_{\tau_n} = \sigma \max\{W_{1,\tau_{n+1}} - W_{1,\tau_n}, W_{2,\tau_{n+1}} - W_{2,\tau_n}\}. \quad (12)$$

Because of the Brownian increments, this process will not get stuck. Long-run growth will emerge. Selection is an essential ingredient: only the best of two recent histories survives when two parties meet.

To calculate the growth rate, write $\Phi(x)$ for the standard Gaussian distribution function and $\phi(x) = (2\pi)^{-1/2}e^{-x^2/2}$ for its density. The distribution of the maximum of two independent standard Gaussians is $\Phi^2(x)$, which has a density $2\phi(x)\Phi(x)$. The mean of the maximum is therefore

$$\int_{-\infty}^{\infty} 2x\phi(x)\Phi(x)dx = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} xe^{-x^2/2}\Phi(x)dx = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{e^{-x^2}dx}{\sqrt{\pi}} = \frac{1}{\sqrt{\pi}}.$$

Integration-by-parts explains the second equality. Since $W_{1,t}$ and $W_{2,t}$ are independent Gaussian with mean zero, variance t , and independent increments,

$$\mathbb{E} \left[\max\{\sigma(W_{1,\tau_{n+1}} - W_{1,\tau_n}), \sigma(W_{2,\tau_{n+1}} - W_{2,\tau_n})\} \mid W_{1,\tau_n}, W_{2,\tau_n} \right] = \sigma\sqrt{(\tau_{n+1} - \tau_n)/\pi}.$$

The mean growth rate between two meeting dates is therefore equal to

$$\mathbb{E} \left[\frac{y_{\tau_{n+1}} - y_{\tau_n}}{\tau_{n+1} - \tau_n} \mid y_{\tau_n} \right] = \int_0^{\infty} \frac{\sigma\sqrt{\tau/\pi}}{\tau} \times \alpha e^{-\alpha\tau} d\tau = \sigma\sqrt{\alpha}. \quad (13)$$

To see the second equality, change variables to $t = \alpha\tau$ and then to $s = \sqrt{2t}$, so that

$$\int_0^{\infty} \frac{\alpha e^{-\alpha\tau} d\tau}{\sqrt{\pi\tau}} = \sqrt{\alpha} \int_0^{\infty} \frac{e^{-t} dt}{\sqrt{\pi t}} = \sqrt{\alpha} \times 2 \int_0^{\infty} \frac{e^{-s^2/2} ds}{\sqrt{2\pi}} = \sqrt{\alpha}.$$

Growth can also be measured as $E[y_{\tau_{n+1}} - y_{\tau_n} | y_{\tau_n}] / E[\tau_{n+1} - \tau_n | y_{\tau_n}] = \frac{1}{2}\sigma\sqrt{\alpha}$ (a similar calculation.) Jensen’s inequality explains the difference between this and (13).

The mean growth rate (13) goes to zero as σ^2 goes to zero: growth vanishes as the variance of the individual discovery processes goes to zero. Also, the productivity state $y_t = \max\{y_{1,t}, y_{2,t}\}$ is expected to have a linear trend. The exponential trend associated with (5)-(6) does not re-appear in this economy. The individual productivity states follow (11), and the “size” of these individual discoveries matters. Of course, one can interpret $y_{j,t}$ as the log productivity of producer j .

With the meeting rate $\alpha = 2\beta$ that arises when the two producers initiate meetings independently at the rate β , the expected growth rate (13) matches what happens in a large economy in which producers make individual discoveries and can learn from others at the rate β . See equations (20)-(21) and (37) below. In various guises, a formula like (13) appears in Luttmer [2007, 2012b, 2015] and Staley [2011]. For an economy in which only the searching producer learns, Staley [2011] gives an argument suggesting that the growth rate will be increasing in the number of producers, but with an asymptote given by (13). This scale effect is intuitive: the value of additional sources of information declines if it takes time to process each and every one of them. It echoes the emphasis in Weitzman [1998] on our limited ability to process potentially fruitful ideas.

4. INDIVIDUAL DISCOVERIES AND SOCIAL LEARNING

This section describes what happens in a large population when individuals make their own discoveries and can learn from others after some delay. Both search delays and learning delays are considered. The differential equations (1) and (7) have to be modified in an important way.

4.1 Brownian Discoveries and Two Types of Delay

Consider a large population of producers, with idiosyncratic productivity states y_t that evolve according to

$$dy_t = \mu dt + \sigma dW_t \tag{14}$$

in between times when producers learn from other producers. The Brownian motions are independent across producers.

Suppose first that producers randomly sample other producers at the rate β . A producer who samples a more productive producer can instantaneously adopt the more productive technology. The state y_t of such a producer jumps, and then continues to

follow (14). Take some $\kappa > \mu$ and define

$$z_t = y_t - \kappa t.$$

Let $F(t, z)$ be the distribution of these de-trended productivity states in the time- t population, and write $f(t, z)$ for its density. This form of social learning adds a term to the usual Kolmogorov forward equation for (14),

$$D_t F(t, z) = -(\mu - \kappa)D_z F(t, z) + \frac{1}{2}\sigma^2 D_{zz} F(t, z) - \beta F(t, z)[1 - F(t, z)]. \quad (15)$$

Intuitively, if $\sigma = 0$ and $\beta = 0$ then $z_t = z_0 + (\mu - \kappa)t$, and differentiating $F(t, z) = F(0, z - (\mu - \kappa)t)$ with respect to t gives the first term on the right-hand side of (15). And if $\mu - \kappa = 0$ and $\beta = 0$, then $D_t F(t, z) = \frac{1}{2}\sigma^2 D_z f(t, z)$ captures the fact that random up and down movement causes the population as a whole to move downhill. The last term on the right-hand side is familiar from (1). Equation (15) is used in Staley [2011] and Luttmer [2012b].

Alternatively, suppose every producer below the median is assigned to a producer above the median and learns to adopt the better after an exponentially distributed delay with mean $1/\beta$. Assignments are updated continuously, as circumstances change. The analog of (7) then becomes

$$D_t F(t, z) = -(\mu - \kappa)D_z F(t, z) + \frac{1}{2}\sigma^2 D_{zz} F(t, z) - \beta \min\{F(t, z), 1 - F(t, z)\}. \quad (16)$$

The partial differential equations (15) and (16) describe the evolution of $F(t, z)$, starting from a given initial value $F(0, z)$. These differential equations are examples of what are called reaction-diffusion equations. In the standard interpretation, the second-order term $\frac{1}{2}\sigma^2 D_{zz} F(t, z)$ describes diffusion (in a physical sense, across space), and the last term in both equations is the “reaction term” that gives these equations their name.² A useful textbook treatment can be found in Volpert, Volpert and Volpert [1994]. Luttmer [2015] describes an economy in which (16) arises as a special case.

4.2 Stationary Distributions

The common shape of the third (reaction) term in (15) and (16) implies similarities for their solutions. In particular, both equations will have stationary solutions $F(t, z) =$

²Importantly, the source of knowledge diffusion here is the “reaction term,” and not the “diffusion term” $\frac{1}{2}\sigma^2 D_{zz} F(t, z)$. See Cavalli-Sforza and Feldman [1981, p. 42] for a classical interpretation of (15) in which this diffusion term does describe knowledge diffusion, but of a single innovation, across space.

$F(z)$ for any κ that satisfies the lower bound

$$\kappa \geq \mu + \sigma^2 \sqrt{\frac{\beta}{\sigma^2/2}}. \quad (17)$$

Observe that this immediately implies that $\kappa > \mu$. Social learning raises the growth rate above what individual producers can do on their own.

To conform with convention, define $R(t, z) = 1 - F(t, z)$ and notice that the reaction-diffusion equations (15) and (16) are of the form

$$D_t R(t, z) = -(\mu - \kappa)D_z R(t, z) + \frac{1}{2}\sigma^2 D_{zz} R(t, z) + \beta Q(R(t, z)),$$

where $Q(R) = (1 - R)R$ in the case of search delays, and $Q(R) = \min\{1 - R, R\}$ in the case of learning delays. Conjecture a stationary solution with density $f(z)$. That is, the right tail is $R(z) = 1 - F(z)$ and $DR(z) = -f(z)$. The partial differential equation for $R(t, z)$ becomes an ordinary differential equation for $R(z)$,

$$0 = -(\mu - \kappa)DR(z) + \frac{1}{2}\sigma^2 D^2 R(z) + \beta Q(R(z)).$$

The boundary conditions are $R(\infty) = 0$ and $R(-\infty) = 1$. Note that this equation is autonomous. If $R(z)$ is a solution, then so is $R(a + z)$ for any constant a . Recall that $z_t = y_t - \kappa t$, and so this is just a change in the origin with respect to which time is measured.

4.2.1 Phase Diagrams

As usual, a second-order differential equation for $R(z)$ can be written as a first-order differential equation equation for the function $[R(z), f(z)]$,

$$DR(z) = -f(z), \quad Df(z) = \frac{-(\kappa - \mu)f(z) + \beta Q(R(z))}{\sigma^2/2}.$$

This two-dimensional system can be examined in a phase diagram, shown in Figure 3 for both versions of $Q(R)$. Observe that the solid trajectories satisfy the boundary condition $R(-\infty) = 1$, in contrast to the dotted trajectories. Both phase diagrams are drawn under the assumption (17), and this accounts for the fact that the trajectories shown converge to the origin. To see why this happens, note that $(1 - R)R \approx R$ and $\min\{1 - R, R\} = R$ near $R = 0$. So the linearized differential equation for $R(z)$ is simply

$$0 \approx -(\mu - \kappa)DR(z) + \frac{1}{2}\sigma^2 D^2 R(z) + \beta R(z). \quad (18)$$

The solutions to this linear second-order differential equation with constant coefficients are linear combinations of $e^{-\zeta_{\pm}z}$. The coefficients ζ_{\pm} must solve a quadratic equation that is solved by

$$\zeta_{\pm} = \frac{\kappa - \mu}{\sigma^2} \pm \sqrt{\left(\frac{\kappa - \mu}{\sigma^2}\right)^2 - \frac{\beta}{\sigma^2/2}}. \quad (19)$$

If β is too large, then these roots will be complex. In the phase diagrams shown in Figure 3, this will lead to trajectories that spiral around the origin. The resulting solution would not be everywhere positive, and so no stationary distribution can be constructed in this case. If β is positive but small enough to ensure that the ζ_{\pm} are real, then $\kappa - \mu < 0$ would imply that both ζ_{\pm} are negative. This would result in solutions for $R(z)$ that explode as $z \rightarrow \infty$ rather than converge to zero. Taken together, these observations give rise to the bound given in (17).

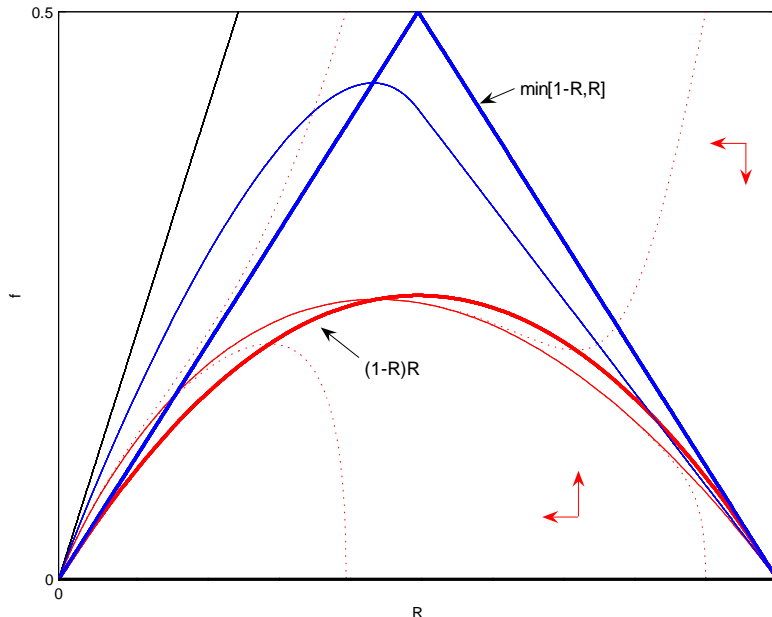


FIGURE 3 Phase Diagrams for $Q(R) = (1 - R)R$ and $Q(R) = \min\{1 - R, R\}$.

The Tail Index The solution for large z will be well approximated by a linear combination of the functions $e^{-\zeta_{\pm}z}$. Since $\zeta_+ \geq \zeta_- > 0$, this implies that

$$-\frac{DR(z)}{R(z)} \approx \zeta_-$$

for large z . That is, ζ_- measures the exponential rate of decay of the right tail of the stationary distribution. If z represents log productivity, then this says that productivity itself, e^z , follows a power law (a Pareto-like distribution) with tail index ζ_- .

4.2.2 Initial Conditions with Bounded Support

As long as κ satisfies the bound (17), and subject to the caveat that one can shift the location of $R(z)$, the phase diagram shown in Figure 3 pins down a unique solution for the distribution $F(z) = 1 - R(z)$. But there is a continuum of κ that satisfy the bound (17). A sufficiently strong linear trend can make the de-trended productivity states $z_t = y_t - \kappa t$ stationary, an even stronger result than the stationarity of expected changes in the productivity state obtained for the finite economy described in Section 3. But that two-person economy gave a definite prediction for the average growth rate: as computed in (13) for the case $\mu = 0$, it equals $\kappa = \mu + \sigma\sqrt{2\beta}$. Here instead, as in the social-learning-only examples introduced in Section 2, we still have many possible growth rates and associated stationary distributions.

It turns out that only the solution associated with κ at the lower bound in (17) can arise when the support of $F(0, z)$ is bounded. This result is due to Kolmogorov, Petrovskii and Piskunov [1937], who were studying (15), as was Fisher [1937], as a model of the spread of an advantageous gene across space (the Cavalli-Sforza and Feldman [1981] application mentioned above is directly based on this.) From hereon, refer to the bound (17) as the KPP bound. An alternative proof of this result was given by McKean [1976]. The subsequent related literature is extensive. Bramson [1984] obtained important results for more general initial conditions $F(0, z)$. Even if $F(0, z)$ has an unbounded support, the long-run growth rate will still be the κ implied by the KPP bound provided the right tail of $F(0, z)$ is sufficiently thin. And an initial distribution with a sufficiently thick tail will result in a long-run growth rate κ that exceeds the KPP bound.

When the KPP bound (17) holds with equality, the two roots ζ_+ and ζ_- coincide and are equal to

$$\zeta = \sqrt{\frac{\beta}{\sigma^2/2}}. \quad (20)$$

The growth rate κ can then be written as

$$\kappa = \mu + \sigma^2\zeta, \quad (21)$$

which matches (13). This implies that the growth rate κ is increasing in both the learning rate β and the variance σ^2 . The tail index ζ is also increasing in β , but decreasing in σ^2 . A low tail index corresponds to a stationary distribution with a thick tail, with very unequal productivities. So a high β is associated with fast growth and low inequality, while a high σ^2 is associated with fast growth and high inequality.

Continuity Obviously, as made explicit in Section 2, an initial distribution $F(0, z)$ that has bounded support produces no long-run growth without individual discoveries, when $\sigma^2 = 0$. When $\sigma^2 > 0$, an initial distribution with bounded support selects a stationary distribution and long-run growth rate characterized by (20)-(21). These formulas imply that $\kappa \downarrow \mu$ and $\zeta \uparrow \infty$ as $\sigma^2 \downarrow 0$. As σ^2 becomes small, aggregate growth approaches what individual producers can do on their own, on average, and the right tail of the productivity distribution becomes very thin. These limits show that the growth rate and right tail implications of the economy considered here are continuous in σ^2 . Almost no variance in individual discoveries leads to almost no growth.

4.2.3 An Explicit Solution—The Case of Learning Delays

Some of this can be made more explicit in the case of learning delays. Let $f(t, z)$ be the density of $F(t, z)$ and differentiate (16) with respect to z to obtain

$$D_t f(t, z) = -(\mu - \kappa)D_z f(t, z) + \frac{1}{2}\sigma^2 D_{zz} f(t, z) + \begin{cases} -\beta f(t, z), & z \in (-\infty, x_t), \\ +\beta f(t, z), & z \in (x_t, \infty), \end{cases}$$

where x_t is the median,

$$\frac{1}{2} = \int_{-\infty}^{x_t} f(t, z) dz,$$

and the initial value $f(0, z)$ is given. Conjecture now that there is a density $f(z)$ so that $f(t, z) = f(z)$, and take its median to be zero. Then $f(z)$ will have to satisfy

$$0 = -(\mu - \kappa)Df(z) + \frac{1}{2}\sigma^2 D^2 f(z) + \begin{cases} -\beta f(z), & z \in (-\infty, 0), \\ +\beta f(z), & z \in (0, \infty). \end{cases}$$

The density will have to be positive and vanish at $\pm\infty$. This is now a second-order ordinary differential equation that is linear on the two intervals $(-\infty, 0)$ and $(0, \infty)$. The two equations have exponential solutions on their respective domains. On $(0, \infty)$, we now have an exact version of (18), with solutions that are linear combinations of $e^{-\zeta_{\pm} z}$, with the ζ_{\pm} given in (19). On $(-\infty, 0)$, the solutions are linear combinations of $e^{-\xi_{\pm} z}$, where

$$\xi_{\pm} = \frac{\kappa - \mu}{\sigma^2} \pm \sqrt{\left(\frac{\kappa - \mu}{\sigma^2}\right)^2 + \frac{\beta}{\sigma^2/2}}.$$

Note that $\xi_- < 0 < \xi_+$, and so $e^{-\xi_- z} \rightarrow 0$ and $e^{-\xi_+ z} \rightarrow \infty$ as $z \rightarrow -\infty$. This means that $e^{-\xi_+ z}$ cannot be part of the solution. Hence $f(z) = f(0)e^{-\xi_- z}$ for all $z < 0$. The solution on $(0, \infty)$ is of the form $f(z) = A_- e^{-\zeta_- z} + A_+ e^{-\zeta_+ z}$, for some A_- and A_+ that

remain to be determined. Forcing the density to be continuous yields $f(0) = A_- + A_+$. Requiring the density to be differentiable at 0 yields $f(0)\xi_- = A_-\zeta_- + A_+\zeta_+$. In other words,

$$\begin{bmatrix} 1 \\ \xi_- \end{bmatrix} f(0) = \begin{bmatrix} 1 & 1 \\ \zeta_- & \zeta_+ \end{bmatrix} \begin{bmatrix} A_- \\ A_+ \end{bmatrix}.$$

When the KPP bound (17) holds as a strict inequality, ζ_- and ζ_+ will be distinct, and this equation can be solved for a unique $[A_-, A_+]$. The resulting density is given by $f(z) = f(0)e^{-\xi_-z}$ on $(-\infty, 0]$ and by

$$f(z) = f(0)e^{-\xi_-z} \times \frac{(\zeta_+ - \xi_-)e^{-(\zeta_- - \xi_-)z} - (\zeta_- - \xi_-)e^{-(\zeta_+ - \xi_-)z}}{\zeta_+ - \zeta_-}$$

on $[0, \infty)$. Note that this is positive because $\zeta_+ > \zeta_- > \xi_-$. The scale parameter $f(0)$ follows from forcing $f(z)$ to be a probability density on $(-\infty, \infty)$. As κ approaches the KPP bound (17) from above, both ζ_{\pm} converge to the ζ given in (20). Although the coefficients $[A_-, A_+]$ do not converge, the density does converge, to

$$f(z) = f(0)(1 + (\zeta - \xi_-)z)e^{-\zeta z}.$$

Restricted to $(0, \infty)$, $f(z)$ is a mixture of an exponential density and the Gamma density obtained in Lemma 2 of Luttmer [2007]. See also page 10 of Technical Appendix II of Luttmer [2015] (see www.luttmer.org).

5. SELECTION IN THE AGGREGATE

Suppose producers exit when they are not sufficiently productive, and that only new entrants can learn from other producers by randomly sampling the population of incumbent producers. This is the basic scenario proposed in Luttmer [2007]. One could call this the putty-clay model of knowledge diffusion: new entrants can adopt what others have learnt, whereas incumbents have to go with what they can discover on their own.³

Consider an exit barrier b_t , the same for all producers, that evolves according to

$$b_t = b_0 + \kappa t. \tag{22}$$

When the state y_t of some producer hits b_t , this producer exits and is replaced by a new producer. This new producer enters with a productivity state drawn at random from

³The economy described in Luttmer [2012b] is a putty-putty model with social learning by both entrants and incumbents.

the incumbent population of producers, a form of social learning. The drift μ and the diffusion coefficient σ , as well as the trend κ of the exit barrier, are exogenously given parameters. The only assumptions are

$$\kappa > \mu, \quad \sigma > 0. \quad (23)$$

The initial exit threshold b_0 and the initial distribution of productivity states at time $t = 0$ are also given. The support of this initial distribution is contained in (b_0, ∞) . The population of producers is constant by construction. We want to know what happens over time to the cross-sectional distribution of y_t . It is convenient to eliminate the trend of the exit barrier, as follows. Define

$$z_t = y_t - b_t, \quad \theta = \mu - \kappa.$$

Then $dy_t = \mu dt + \sigma dW_t$ yields

$$dz_t = \theta dt + \sigma dW_t. \quad (24)$$

Assumption (23) implies $\theta < 0$ and the exit threshold for z_t is 0 by construction.

5.1 The Cohort Distribution

Recall that ϕ and Φ are the standard normal density and distribution functions. Consider a cohort of producers who all start out in some state $x > 0$. The initial measure of this cohort is 1. It is well known (for example, see Harrison [1985], or the application in Luttmer [2007]) that the density at $t > 0$ of survivors in this cohort is $\psi(t, y|x)$, where

$$\psi(t, y|x) = \frac{1}{\sigma\sqrt{t}} \left[\phi\left(\frac{y-x-\theta t}{\sigma\sqrt{t}}\right) - \exp\left(-\frac{\theta x}{\sigma^2/2}\right) \phi\left(\frac{y+x-\theta t}{\sigma\sqrt{t}}\right) \right]. \quad (25)$$

The measure of survivors in $[0, z]$ at time t is thus

$$\begin{aligned} \int_0^z \psi(t, y|x) dy &= \Phi\left(\frac{z-x-\theta t}{\sigma\sqrt{t}}\right) - \exp\left(-\frac{\theta x}{\sigma^2/2}\right) \Phi\left(\frac{z+x-\theta t}{\sigma\sqrt{t}}\right) \\ &\quad - \left[\Phi\left(\frac{-x-\theta t}{\sigma\sqrt{t}}\right) - \exp\left(-\frac{\theta x}{\sigma^2/2}\right) \Phi\left(\frac{x-\theta t}{\sigma\sqrt{t}}\right) \right]. \end{aligned}$$

Taking the $z \rightarrow \infty$ limit gives the measure of all survivors at time t ,

$$S(t|x) = \int_0^\infty \psi(t, y|x) dy = \Phi\left(\frac{x+\theta t}{\sigma\sqrt{t}}\right) - \exp\left(-\frac{\theta x}{\sigma^2/2}\right) \Phi\left(\frac{-x+\theta t}{\sigma\sqrt{t}}\right).$$

The assumption $\theta < 0$ implies that $S(t|x) \rightarrow 0$ as $t \rightarrow \infty$. The probability distribution of the state z among survivors at time t is given by

$$\begin{aligned} P(t, z|x) &= \frac{1}{S(t|x)} \int_0^z \psi(t, y|x) dy \\ &= \frac{\Phi\left(\frac{z-x-\theta t}{\sigma\sqrt{t}}\right) - \Phi\left(\frac{-x-\theta t}{\sigma\sqrt{t}}\right) - e^{-\frac{\theta x}{\sigma^2/2}} \left[\Phi\left(\frac{z+x-\theta t}{\sigma\sqrt{t}}\right) - \Phi\left(\frac{x-\theta t}{\sigma\sqrt{t}}\right) \right]}{\Phi\left(\frac{x+\theta t}{\sigma\sqrt{t}}\right) - e^{-\frac{\theta x}{\sigma^2/2}} \Phi\left(\frac{-x+\theta t}{\sigma\sqrt{t}}\right)} \end{aligned}$$

Since $\Phi(-u) = 1 - \Phi(u)$, this can also be written as

$$P(t, z|x) = \frac{\Phi\left(\frac{x+\theta t}{\sigma\sqrt{t}}\right) - \Phi\left(\frac{-z+x+\theta t}{\sigma\sqrt{t}}\right) - e^{-\frac{\theta x}{\sigma^2/2}} \left[\Phi\left(\frac{-x+\theta t}{\sigma\sqrt{t}}\right) - \Phi\left(\frac{-z-x+\theta t}{\sigma\sqrt{t}}\right) \right]}{\Phi\left(\frac{x+\theta t}{\sigma\sqrt{t}}\right) - e^{-\frac{\theta x}{\sigma^2/2}} \Phi\left(\frac{-x+\theta t}{\sigma\sqrt{t}}\right)}$$

and thus

$$1 - P(t, z|x) = \frac{\Phi\left(\frac{-z+x+\theta t}{\sigma\sqrt{t}}\right) - e^{-\frac{\theta x}{\sigma^2/2}} \Phi\left(\frac{-z-x+\theta t}{\sigma\sqrt{t}}\right)}{\Phi\left(\frac{x+\theta t}{\sigma\sqrt{t}}\right) - e^{-\frac{\theta x}{\sigma^2/2}} \Phi\left(\frac{-x+\theta t}{\sigma\sqrt{t}}\right)}. \quad (26)$$

Since $\theta < 0$, both the numerator and the denominator on the right-hand side of (26) go to zero as t goes to infinity.

This cohort probability distribution among survivors will also be the actual distribution at time t in a population that starts out concentrated at x and that is maintained by having exiting producers be replaced by new entrants with states drawn at random from the population of survivors. This is precisely the distribution we are trying to describe.

5.2 The Limiting Distribution

We want to examine the limit of $P(t, z|x)$ as t becomes large. This is called the Yaglom limit for the process z_t , after Yaglom [1947]. The following calculations are adapted from the technical appendix of Luttmmer [2007].

Fix some z and x , and write $h = 1/(\sigma\sqrt{t})$ and $\zeta = -\theta/\sigma^2 > 0$. Note that this definition of ζ is consistent with (20)-(21). Then $1 - P(t, z|x) = f(h)/g(h)$, where

$$\begin{aligned} f(h) &= \Phi\left(\left[-z+x-\frac{\zeta}{h^2}\right]h\right) - e^{2\zeta x} \Phi\left(\left[-z-x-\frac{\zeta}{h^2}\right]h\right), \\ g(h) &= \Phi\left(\left[x-\frac{\zeta}{h^2}\right]h\right) - e^{2\zeta x} \Phi\left(\left[-x-\frac{\zeta}{h^2}\right]h\right). \end{aligned}$$

Since $\zeta > 0$, it follows that both $f(h) \rightarrow 0$ and $g(h) \rightarrow 0$ as $h \downarrow 0$. The limit of $f(h)/g(h)$ can be computed using l'Hôpital's rule, as follows. Observe first that

$$Df(h) = \left[-z + x + \frac{\zeta}{h^2}\right] \phi\left(\left[-z + x - \frac{\zeta}{h^2}\right]h\right) - \left[-z - x + \frac{\zeta}{h^2}\right] e^{2\zeta x} \phi\left(\left[-z - x - \frac{\zeta}{h^2}\right]h\right).$$

Note that

$$\begin{aligned} \phi\left(\left[-z + x - \frac{\zeta}{h^2}\right]h\right) &= \phi\left(\left[x - \frac{\zeta}{h^2}\right]h\right) \exp(-\zeta z) \exp(z(x - z/2)h^2), \\ e^{2\zeta x} \phi\left(\left[-z - x - \frac{\zeta}{h^2}\right]h\right) &= \phi\left(\left[x - \frac{\zeta}{h^2}\right]h\right) \exp(-\zeta z) \exp(-z(x + z/2)h^2). \end{aligned}$$

and hence

$$\begin{aligned} Df(h) &= \phi\left(xh - \frac{\zeta}{h}\right) \exp(-\zeta z) \times \\ &\quad \left\{ \left[-z + x + \frac{\zeta}{h^2}\right] \exp(z(x - z/2)h^2) - \left[-z - x + \frac{\zeta}{h^2}\right] \exp(-z(x + z/2)h^2) \right\} \end{aligned}$$

Next, observe that

$$Dg(h) = \left(\left[x + \frac{\zeta}{h^2}\right]\right) \phi\left(\left[x - \frac{\zeta}{h^2}\right]h\right) - e^{2\zeta x} \left[-x + \frac{\zeta}{h^2}\right] \phi\left(\left[-x - \frac{\zeta}{h^2}\right]h\right)$$

and note that

$$-\frac{1}{2} \left(\left(x - \frac{\zeta}{h^2}\right)h\right)^2 = 2\zeta x - \frac{1}{2} \left(\left(-x - \frac{\zeta}{h^2}\right)h\right)^2.$$

Therefore

$$Dg(h) = 2x \phi\left(xh - \frac{\zeta}{h}\right).$$

The desired ratio of derivatives is thus

$$\frac{Df(h)}{Dg(h)} = \frac{e^{-\zeta z}}{2x} \left[\left(z + x - \frac{\zeta}{h^2}\right) \exp(-z(x + z/2)h^2) - \left(z - x - \frac{\zeta}{h^2}\right) \exp(z(x - z/2)h^2) \right].$$

Clearly,

$$\lim_{h \downarrow 0} \frac{1}{2x} [(z + x) \exp(-z(x + z/2)h^2) - (z - x) \exp(z(x - z/2)h^2)] = 1, \quad (27)$$

and

$$\begin{aligned} &\lim_{h \downarrow 0} \frac{1}{2x} \left[\left(-\frac{\zeta}{h^2}\right) \exp(-z(x + z/2)h^2) - \left(-\frac{\zeta}{h^2}\right) \exp(z(x - z/2)h^2) \right] \\ &= \frac{\zeta}{2x} \lim_{h \downarrow 0} \left[\frac{1 - \exp(-z(x + z/2)h^2)}{h^2} + \frac{\exp(z(x - z/2)h^2) - 1}{h^2} \right] \\ &= \frac{\zeta}{2x} [z(x + z/2) + z(x - z/2)] = \zeta z. \end{aligned} \quad (28)$$

The ratio of derivatives therefore converges to

$$\lim_{h \downarrow 0} \frac{Df(h)}{Dg(h)} = (1 + \zeta z)e^{-\zeta z}. \quad (29)$$

By l'Hôpital's rule, $f(h)/g(h)$ converges to the same limit. Observe that this limit does not depend on the initial state x . In fact, the limits (27)-(28) and thus (29) are uniform in x on any compact $X \subset (0, \infty)$. In turn, this implies that the convergence of $f(h)/g(h)$ to $(1 + \zeta z)e^{-\zeta z}$ is uniform in x on any compact $X \subset (0, \infty)$. To summarize,

$$\lim_{t \rightarrow \infty} P(t, z|x) = P(z) = 1 - (1 + \zeta z)e^{-\zeta z} \quad (30)$$

and the convergence is uniform in $x \in X$, for any compact $X \subset (0, \infty)$. The right-hand of (30) is the Gamma distribution with density $\zeta^2 z e^{-\zeta z}$ on $(0, \infty)$.

5.2.1 Initial Conditions with Compact Support

Now suppose the initial conditions are distributed according to a distribution F with compact support $X \subset (0, \infty)$. The protocol for updating this distribution is as before: exiting producers are replaced by new producers whose productivity is drawn at random from the incumbent population. Let $P_F(t, z)$ denote the resulting distribution at age t . Then

$$P_F(t, z) = \frac{\int_X \left[\int_0^z \psi(t, y|x) dy \right] dF(x)}{\int_X \left[\int_0^\infty \psi(t, y|x) dy \right] dF(x)} = \int_X P(t, z|x) w(t|x) dF(x),$$

where the weights $w(t|x)$ are given by

$$w(t|x) = \frac{S(t|x)}{\int_X S(t|x) dF(x)}.$$

Clearly, these weights integrate to 1 against the distribution F . Fix some $z > 0$. Take any $\varepsilon > 0$. The uniform convergence in (30) implies that there is a t_ε such that

$$|P(t, z|x) - P(z)| \leq \varepsilon$$

for all $x \in X$ and all $t \geq t_\varepsilon$. Hence

$$|P_F(t, z) - P(z)| \leq \int_X |P(t, z|x) - P(z)| w(t|x) dF(x) \leq \varepsilon$$

for all $t \geq t_\varepsilon$. We therefore obtain the limit

$$\lim_{t \rightarrow \infty} P_F(t, z) = P(z). \quad (31)$$

That is, as long as the distribution F of the initial state has compact support, the resulting $P_F(t, \cdot)$ converges in distribution to the Gamma distribution with density $\zeta^2 z e^{-\zeta z}$.

5.3 Many Quasi-Stationary Distributions

A distribution Q with the property $P_Q(t, z) = Q(z)$ for all $t \geq 0$ is called a quasi-stationary distribution.⁴ In the environment described here, there are many quasi-stationary distributions. Let $q(z)$ be the probability density of a quasi-stationary distribution. Such a density has to satisfy the Kolmogorov forward equation

$$0 = -(\mu - \kappa)Dq(z) + \frac{1}{2}\sigma^2 D^2q(z) + \varepsilon q(z) \quad (32)$$

for some entry rate $\varepsilon > 0$, with the boundary condition

$$0 = q(0). \quad (33)$$

The last term on the right-hand side of (32) measures the flow of entrants needed to maintain the population, which would otherwise shrink because of exit at 0. Integrating (32) and imposing the boundary condition (33), as well as the restriction that $q(z)$ must be a probability density, gives

$$\varepsilon = \frac{1}{2}\sigma^2 Dq(0).$$

The right-hand side is known to be the flow across the exit boundary 0 (see Cox and Miller [1966]) and so ε is indeed the entry rate required to maintain a constant population.

Observe that (32) is the same equation as the approximate differential equation for the right tail (exact in the case of learning delays) given earlier in (18), except that here ε is not a fixed learning rate but an entry rate that has to be chosen to keep the population constant. And the boundary condition (33) is new. Subject to these qualifications, the implications are the essentially the same. The differential equation (32) has solutions of the form $e^{-\zeta_{\pm}z}$, where

$$\zeta_{\pm} = \frac{\kappa - \mu}{\sigma^2} \pm \sqrt{\left(\frac{\kappa - \mu}{\sigma^2}\right)^2 - \frac{\varepsilon}{\sigma^2/2}}. \quad (34)$$

If these roots are distinct, then the boundary condition (33) implies $q(z) \propto e^{-\zeta_-z} - e^{-\zeta_+z}$. Complex roots can only result in real $q(z)$ that oscillate around zero. We must therefore require that

$$\left(\frac{\kappa - \mu}{\sigma^2}\right)^2 \geq \frac{\varepsilon}{\sigma^2/2}. \quad (35)$$

⁴See Collet, Martínez, and San Martín [2013] for a monograph on the topic.

Suppose now that (35) holds as a strict inequality, so that the roots ζ_{\pm} are real and distinct. Normalizing $q(z)$ to integrate to 1 gives

$$q(z) = \frac{e^{-\zeta_- z} - e^{-\zeta_+ z}}{\frac{1}{\zeta_-} - \frac{1}{\zeta_+}}. \quad (36)$$

Since $\zeta_+ > \zeta_- > 0$, this is positive for all positive z . As a check on the calculations, observe that

$$\frac{1}{2}\sigma^2 Dq(0) = \frac{1}{2}\sigma^2 \times \frac{\zeta_+ - \zeta_-}{\frac{1}{\zeta_-} - \frac{1}{\zeta_+}} = \frac{1}{2}\sigma^2 \times \zeta_+ \zeta_- = \varepsilon,$$

as required. With the entry rate ε at the upper bound (35), the two roots ζ_{\pm} merge into $\zeta = (\kappa - \mu)/\sigma^2$. It is easy to show that the density (36) converges to the Gamma density $\zeta^2 z e^{-\zeta z}$ as $\zeta_{\pm} \rightarrow \zeta$. This Gamma density is just the Yaglom limit derived earlier in (30). That is, the Yaglom limit is a quasi-stationary density, and it is the quasi-stationary density with the highest possible exit rate ε .

To summarize: given a growth rate $\kappa > \mu$, there is a continuum of quasi-stationary distributions with densities defined by (34) and (36), and indexed by exit rates $\varepsilon > 0$ that have to satisfy (35). The quasi-stationary distribution that is the limiting distribution associated with an initial distribution with compact support is the unique quasi-stationary distribution for which the exit rate is maximal. This is the basis for Lemma 2 in Luttmer [2007].

5.4 What this Leaves Out

The Yaglom limit (30) and the quasi-stationary distributions (36) were constructed for an exogenously specified exit threshold $b_t = b_0 + \kappa t$, with $\kappa > \mu$. In a fully specified economy (such as Luttmer [2007]), these exit thresholds are endogenous and depend on beliefs about how interest rates, wages, and the demand for differentiated goods evolve. With the usual rational expectations assumption, beliefs are endogenous and have to be consistent with how the economy will in fact evolve. Given an initial distribution $F(\cdot)$, an equilibrium is a fixed point in a space of infinite histories of cross-sectional distributions $\{P_F(t, \cdot) : t > 0\}$. Characterizing the properties of such a fixed point is a much more difficult problem than the one examined above.

5.4.1 Balanced Growth

Luttmer [2007] focuses on the unique balanced growth path in which the bound (35) holds with equality, as suggested by the Yaglom limit. When (35) holds with equality,

it defines a relation between the entry rate ε and the growth rate κ that can be written as

$$\kappa = \mu + \sigma^2 \times \sqrt{\frac{\varepsilon}{\sigma^2/2}}. \quad (37)$$

The distribution of productivities relative to the exit threshold will then be determined by the Yaglom limit (30). This distribution affects the incentives to enter, since potential entrants sample from this distribution when they try to learn from incumbents. Anyone can try to enter at a cost, and this yields a zero-profit condition that together with (37) determines the entry rate ε and the growth rate κ . A parameter such as the subjective discount rate of consumers will affect the zero-profit condition for entry, and changing this parameter will affect ε and κ . For a concise exposition, see Luttmer [2014].

In an economy without entry and exit that one could construct using the results of Section 3, there will be a unique stationary density given any growth rate κ that satisfies the KPP bound (17). There is no entry rate to consider in such an economy, but also no zero-profit condition for entrants. This will leave κ undetermined unless the KPP bound is required to hold with equality.

6. THE ROLE OF THE RIGHT TAIL

The multiplicity of stationary distributions in these examples is clearly related to what happens in the right tail. This can be illustrated by considering an economy in which all entry is at some fixed distance from the exit boundary, rather than randomly throughout the right tail. Potential entrants can use the technology of exiting producers and improve upon it by some small amount. The following is taken from Luttmer [2012a].

As before, the individual productivity states evolve according to $dy_t = \mu dt + \sigma dW_t$ and $z_t = y_t - b_t$, where $b_t = b_0 + \kappa t$ is an exit barrier and $\kappa > \mu$. Suppose now that exiting producers at b_t are replaced by entrants at $x_t = b_t + \Delta$, for some $\Delta > 0$, an exogenous parameter. A stationary density for the de-trended state z_t will then have to satisfy the Kolmogorov forward equation

$$0 = -(\mu - \kappa)Df(z) + \frac{1}{2}\sigma^2 D^2 f(z), \quad z \in (0, \Delta) \cup (\Delta, \infty) \quad (38)$$

together with the boundary conditions

$$f(0) = 0, \quad f(\Delta_-) = f(\Delta_+). \quad (39)$$

The second part of (39) equates the left and right limits of f at Δ . The scale of $f(z)$ is determined by requiring $f(z)$ to be a probability density. Integrating (38) over $[0, \infty)$

and using the boundary conditions (39) gives

$$\frac{1}{2}\sigma^2 Df(0) = \frac{1}{2}\sigma^2 [Df(\Delta_-) - Df(\Delta_+)].$$

The left-hand side is the flow of exiting producers. The discontinuity in the derivative of the stationary density at Δ measures the inflow of new producers at Δ . As before, write ε for this entry-exit rate.

Recall that $\kappa > \mu$. The differential equation (38) has solutions of the form $e^{-\zeta_{\pm}z}$, where

$$\zeta_+ = \zeta = \frac{\kappa - \mu}{\sigma^2/2}, \quad \zeta_- = 0.$$

In contrast to (19) and (34), the roots ζ_{\pm} are now automatically real, and there is no analog to the KPP bound (17) or entry rate bound (35). The density $f(z)$ will be a linear combination of $e^{-\zeta_{\pm}z}$ on $(0, \Delta)$. The right tail of $f(z)$ must vanish for large z , and so $\zeta_- = 0$ implies that $f(z) \propto e^{-\zeta_+z}$ on (Δ, ∞) . Imposing the boundary conditions (39) and forcing $f(z)$ to integrate to 1 yields

$$f(z) = \frac{1}{\Delta} \times \begin{cases} 1 - e^{-\zeta z}, & z \in (0, \Delta), \\ (1 - e^{-\zeta\Delta})e^{-\zeta(z-\Delta)}, & z \in (\Delta, \infty). \end{cases}$$

The implied entry rate is $\varepsilon = (\sigma^2/2)Df(0) = (1/\Delta)(\sigma^2/2)\zeta = (\kappa - \mu)/\Delta$. The relation between the growth rate $\kappa > \mu$ and the entry rate ε in this economy is not (37) but

$$\kappa = \mu + \Delta \times \varepsilon. \tag{40}$$

Given a growth rate $\kappa > \mu$, here we now have a unique stationary distribution, and the implied entry rate ε is given by (40). This contrasts sharply with the continuum of quasi-stationary distributions and associated entry rates obtained in (34)-(36) for an economy in which entrants can become like randomly drawn incumbents. Luttmer [2012a] describes a fully specified economy in which market clearing conditions and a zero-profit condition for entrants pin down a unique balanced growth path.

This example also highlights the fact that long-run growth can emerge even if nobody can ever learn from the most productive producers in the economy. It is enough for potential entrants to always be able to gain a foothold in the left-tail of a productivity distribution that is continuously shifting to the right. Fortuitous post-entry discoveries by individual producers will put those producers in the right tail. The resulting relation between entry and growth given in (40) is very different from what happens, see (37), when entrants can learn from the most productive incumbents.

7. CONCLUDING REMARKS

The knowledge diffusion models outlined here can be used as building blocks to construct fully specified economies in which productivity growth contributes to consumption growth. The simplest way to use these building blocks is to combine them with a production function that exhibits decreasing returns to scale to labor. As in Luttmer [2007], it is also easy to combine these building blocks with fixed costs, Dixit-Stiglitz preferences, monopolistic competition, and population growth. The number of different goods available to consumers will then grow, and this increase in variety is an independent source of growth. The resulting model avoids the strong scale effect of Romer [1990], in a way that is similar to Young [1998].

REFERENCES

- [1] Alvarez, F.E., F.J. Buera, and R.E. Lucas, Jr., “Models of Idea Flows,” NBER working paper 14135 (2008).
- [2] Bramson, M., *Convergence of Solutions of the Kolmogorov Equation to Travelling Waves*, Memoirs of the American Mathematical Society, vol. 44, no. 285 (1983).
- [3] Cavalli-Sforza, L.L. and M.W. Feldman, *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton University Press (1981).
- [4] Collet, P., S. Martínez, J. San Martín, *Quasi-Stationary Distributions: Markov Chains, Diffusions and Dynamical Systems (Probability and Its Applications)*, Springer-Verlag (2013).
- [5] Cox, D.R. and H.D. Miller, *The Theory of Stochastic Processes*, Chapman and Hall, London (1965).
- [6] Fisher, R.A., “The Wave of Advance of Advantageous Genes,” *Annals of Eugenics*, vol. 7 (1937), 355-369.
- [7] Harrison, J. Michael, *Brownian Motion and Stochastic Flow Systems* (New York, NY: Wiley, 1985).
- [8] 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).

- [9] Kortum, S. “Research, Patenting, and Technological Change,” *Econometrica*, vol. 65, no. 6 (1997), 1389-1419.
- [10] Lucas, R.E., Jr. and B. Moll, “Knowledge Growth and the Allocation of Time,” *Journal of Political Economy*, vol. 122, no. 1 (2014), 1-51.
- [11] Luttmer, E.G.J., “Growth, Selection, and the Size Distribution of Firms,” *Quarterly Journal of Economics*, vol. 122, no. 3 (2007), 1103-1144.
- [12] Luttmer, E.G.J., “Technology Diffusion and Growth,” *Journal of Economic Theory*, vol. 147 (2012a), 602-622.
- [13] Luttmer, E.G.J., “Eventually, Noise and Imitation Implies Balanced Growth,” *Minneapolis Federal Reserve Working Paper 699* (2012b).
- [14] Luttmer, E.G.J., “Lecture Notes on Knowledge Diffusion, Growth, and Income Inequality,” www.luttmer.org (2014).
- [15] Luttmer, E.G.J., “An Assignment Model of Knowledge Diffusion and Income Inequality,” *Federal Reserve Bank of Minneapolis Staff Report 509* (2015).
- [16] 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.
- [17] Romer, P.M., “Endogenous Technical Change,” *Journal of Political Economy*, vol. 98, no. 5 (1990), S71-S102.
- [18] Staley, M., “Growth and the Diffusion of Ideas,” *Journal of Mathematical Economics*, vol. 47 (2011), 470-478.
- [19] Volpert, A.I., V.A. Volpert, and V.A. Volpert, *Traveling Wave Solutions of Parabolic Systems*, Translations of Mathematical Monographs vol. 140, American Mathematical Society (1994).
- [20] Yaglom, A.M. “Certain Limit Theorems of the Theory of Branching Processes,” *Doklady Akademii Nauk SSSR*, vol. 56 (1947), 795-798.
- [21] Young, A., “Growth Without Scale Effects,” *Journal of Political Economy*, vol.106, no.1 (1998), 41-63.
- [22] Weitzman, M.L., “Recombinant Growth,” *Quarterly Journal of Economics*, vol. 113, no. 2, (1998), 331-360.