

# Technology diffusion and growth

Erzo G.J. Luttmer<sup>a,b,1</sup>

<sup>a</sup> *University of Minnesota, MN, United States*

<sup>b</sup> *Federal Reserve Bank of Minneapolis, MN, United States*

Received 8 March 2010; final version received 17 December 2010; accepted 20 December 2010

Available online 18 February 2011

---

## Abstract

Suppose firms are subject to decreasing returns and permanent idiosyncratic productivity shocks. Suppose also firms can only stay in business by continuously paying a fixed cost. New firms can enter. Firms with a history of relatively good productivity shocks tend to survive and others are forced to exit. This paper identifies assumptions about entry that guarantee a stationary firm size distribution and lead to balanced growth. The range of technology diffusion mechanisms that can be considered is greatly expanded relative to Luttmer (2007) [21]. If entrants can make only small improvements over the technologies used by the least productive incumbents, then the firm size distribution approximates Zipf's law and entry and exit rates are high, as in the data.

© 2011 Elsevier Inc. All rights reserved.

*JEL classification:* L1; O3

*Keywords:* Productivity; Imitation; Selection; Diffusion

---

## 1. Introduction

This paper describes a competitive economy in which aggregate productivity growth is endogenous and driven by firm-level experimentation, selection, and imitation by new entrants. The structure of the economy combines elements of Luttmer [21,23], who in turn builds on an

---

*E-mail address:* [luttmer@umn.edu](mailto:luttmer@umn.edu).

<sup>1</sup> I thank Daron Acemoglu, Robert E. Lucas, Jr., and three anonymous referees for helpful comments. 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.

extensive literature on size distributions.<sup>2</sup> The assumption of perfectly competitive markets is a simplifying assumption that shuts down idiosyncratic firm product demand as a source of heterogeneity. The focus is instead on highlighting the types of assumptions about entry that can ensure the existence of a stationary firm size distribution and a balanced growth path when the process of individual firm productivity growth is highly non-stationary.

There are two ingredients. First, there is a mechanism by which new entrants can benefit from the successes of incumbent firms.<sup>3</sup> The assumption here is that entering firms can improve on the technology used by firms at the very low end of the incumbent productivity distribution. Because of this limited ability to imitate, new firms are small, as they are in the data. They can become really productive and large only through stochastic post-entry improvements in productivity. Second, as in Luttmer [23], the supply of new firms created by entrepreneurs is not perfectly elastic with respect to the market value of new firms. This ensures that the growth rate of the economy and the number of firms are jointly determined by an entry condition and a labor market clearing condition. This forces average firm demand for labor to be finite at the equilibrium growth rate. This is typically not guaranteed if a zero-profit entry condition by itself is sufficient to compute the equilibrium growth rate of the economy. In such an economy, a balanced growth path may fail to exist even if a stationary firm size distribution can be constructed.

The employment size distribution of US firms is quite stable over time. Parametric approximations of this distribution (in most studies, a Pareto or Fréchet distribution, or something very similar) are very close to implying an infinite mean. With the technology diffusion and entry assumptions just described, this happens naturally whenever the initial technology available to entrants is sufficiently unproductive or entry is sufficiently costly. Without the equilibrating forces implied by these assumptions, otherwise arbitrary restrictions on firm and aggregate growth would be needed to account for this phenomenon.

This paper emphasizes the combination of trial and error and selection as an important driving force of aggregate growth. The focus is on the role of selection at the level of populations of firms or organizations. Alchian [2] argued for the importance of trial and error, imitation, and selection in understanding the behavior of producers. Nelson and Winter [24] describe models of growth based on selection. The classic paper on selection at the industry level is Jovanovic [17]. His firms learn about a fixed productivity parameter, and those who learn they are productive remain in the industry while those who learn they are not will exit. Selection is a transitory phenomenon. Here firms are subject to new shocks all the time, resulting in perpetual selection and growth.

Although technological progress at the firm level is taken to be exogenous, the growth rate of the aggregate economy is endogenous and depends on entry cost and spillover parameters. These parameters affect the speed with which the selection process improves aggregate output. Atkeson and Burstein [5] describe a closely related economy in which firms can expend resources to increase their own rate of technological progress. Acemoglu and Cao [1] consider a quality-ladder economy in which both incumbents and entrants can innovate. In these economies, reducing the cost of entry can slow down incumbent innovation, an effect that is absent in the current paper.

<sup>2</sup> The US firm size distribution is well approximated by Zipf's law: the number of firms with  $n$  or more employees behaves like  $1/n$ . Classic models of the firm size distribution are Simon and Bonini [26], Steindl [27], Lucas [19]. Gabaix [12] gives an interpretation of Zipf's law for cities. See Gabaix [13] for a survey on power laws in economics and finance, and Luttmer [22] for a survey of models of growth and firm heterogeneity.

<sup>3</sup> As in much of the work on endogenous growth, the mechanism involves an externality. See Boldrin and Levine [6] for an economy that grows without externalities.

Consumers and incumbent firms are introduced in Sections 2 and 3. The key technical material about stationary size distributions is contained in Section 3.1. Section 4 shows that an economy in which entry productivity is exogenous may not have a balanced growth path. Section 5 does the same for an economy with endogenous technological progress and entry decisions that are perfectly elastic with respect to firm value. Section 6 provides a simple remedy. Section 7 presents a calibration and Section 8 concludes.

## 2. Consumers

The population of consumers is  $H_t = He^{\eta t}$  and everyone has one unit of labor. Population growth is non-negative. There is one consumption good at all times. Dynastic preferences over per-capita consumption flows  $c_t$  are

$$\int_0^{\infty} e^{-\rho t} \ln(c_t) dt.$$

The subjective discount rate  $\rho$  may depend on  $\eta$ . The use of logarithmic utility here is mostly for simplicity. More general homothetic preferences can be considered.

Everyone is a price taker and subject to a standard dynastic present-value budget constraint. Throughout, the focus will be on balanced growth. Along a balanced growth path, per-capita consumption and wages are

$$[c_t, w_t] = [c, w]e^{\kappa t},$$

and the resulting interest rate is  $r = \rho + \kappa$ . It is assumed that  $\rho > \eta$  so that the present-value of aggregate consumption is finite.

## 3. Incumbent firms

A firm is a technology for producing consumption goods using labor that is subject to decreasing returns to scale. Firms are the same except for a productivity index, denoted by  $z$ . This productivity index changes continuously over time, as a result of firm-specific random shocks.

### 3.1. Type- $z$ firms

A type- $z$  firm at time  $t$  can produce  $zl^\beta$  units of output with  $l$  units of labor, where  $\beta \in (0, 1)$ . Type- $z$  firms at time  $t$  solve

$$v_t[z] = \max_l \{zl^\beta - w_t l\}.$$

This behaves like a profit function. Since  $z$  multiplies a production function exhibiting decreasing returns,  $v_t[z]$  is a convex function of  $z$ . Let  $l_t[z]$  and  $y_t[z]$  be the optimal levels of employment and output. Then

$$\begin{bmatrix} l_t[z] \\ v_t[z]/w_t \\ y_t[z]/w_t \end{bmatrix} = \frac{1}{\beta} \begin{bmatrix} \beta \\ 1 - \beta \\ 1 \end{bmatrix} \left( \frac{\beta z}{w_t} \right)^{1/(1-\beta)}.$$

Output and labor inputs are also convex functions of  $z$ . Fixing the number of firms and the aggregate labor supply, a mean-preserving spread of productivity will raise aggregate output and wages in this economy.

To survive, a firm must incur a flow cost of  $\lambda_F$  units of labor. Interruption of this cost causes its productivity to permanently collapse to zero. Technology is like the volatile memory of a computer. It will be convenient to define  $s_t[z]$  via

$$e^{s_t[z]} = \frac{v_t[z]}{\lambda_F w_t} = \frac{1}{\lambda_F} \frac{1-\beta}{\beta} \left( \frac{\beta z}{w_t} \right)^{1/(1-\beta)}.$$

With this definition, firm employment and profits can be written as

$$\begin{bmatrix} w_t l_t[z] + w_t \lambda_F \\ v_t[z] - w_t \lambda_F \end{bmatrix} = w_t \lambda_F \begin{bmatrix} \left( \frac{\beta}{1-\beta} \right) e^{s_t[z]} + 1 \\ e^{s_t[z]} - 1 \end{bmatrix}. \quad (1)$$

Thus  $s_t[z] = 0$  corresponds to zero profits. Clearly, employment and profitability are perfectly correlated for this technology, and thus  $s_t[z]$  can be viewed as a measure of both. The fact that employment, profits, and output are convex functions of  $z$  makes them convex functions of  $s_t[z]$  as well.

### 3.2. Productivity dynamics

New firms can enter with an initial productivity given by  $Z_t = Z e^{\kappa t}$ . The cost of entry and the determination of  $Z$  and  $\kappa$  are described in later sections. Following entry, the productivity of a particular time- $t$  entrant evolves with age according to

$$Z_{t,a} = Z_t \exp(\theta a + \sigma_Z W_a),$$

where  $W_a$  is a standard Brownian motion that is independent across firms. As noted,  $Z_{t,a}$  drops permanently to zero if the firm stops paying the fixed cost. The parameters  $\theta$  and  $\sigma_Z > 0$  are taken as exogenous throughout.

Since productivity and wages grow at the same rate  $\kappa$ ,  $s_t[Z_t]$  is constant over time. Define  $S = s_t[Z_t]$ . Then

$$e^S = \frac{1}{\lambda_F} \frac{1-\beta}{\beta} \left( \frac{\beta Z}{w} \right)^{1/(1-\beta)}. \quad (2)$$

Entrant employment and profitability are inversely related to the level of wages in the economy. Since  $w_t l_t[z] = \beta y_t[z]$ , a stationary firm employment distribution combined with a number of firms that grows at the same rate  $\eta$  as aggregate employment results in per-capita output that grows at the same rate  $\kappa$  as wages.

The fact that wages trend with entry productivity also implies that  $s_t[Z_{t-a,a}]$  only depends on  $a$  and not on  $t$ . Thus the state of a firm of age  $a$  is  $s_a = s_t[Z_{t-a,a}]$ , and this evolves with firm age according to  $s_a = S + [(\theta - \kappa)a + \sigma_Z W_a]/(1 - \beta)$ . Along a balanced growth path, there is no aggregate state to keep track of. It will be convenient to write

$$\begin{bmatrix} \mu \\ \sigma \end{bmatrix} = \frac{1}{1-\beta} \begin{bmatrix} \theta - \kappa \\ \sigma_Z \end{bmatrix}$$

so that

$$s_a = S + \mu a + \sigma W_a,$$

as long as the firm does not stop paying the fixed cost.

### 3.3. Exit

Apart from static production decisions, the only choice the firm faces is whether or not to continue. Because of (1), the value of a firm of size  $s$  can be written as  $w_t \lambda_F V(s)$ . The value function  $V(s)$  is given by

$$V(s) = \sup_{\tau} E_0 \left[ \int_0^{\tau} e^{-\rho a} (e^{s_a} - 1) da \right],$$

where  $s_0 = s$ , and  $\tau$  is a stopping time that depends on the observed history of  $s_a$ . The value of the firm is finite if and only if  $\rho > \mu + \sigma^2/2$ , which says that the present-value of  $\{E_0[e^{s_a}]\}_{a \geq 0}$  discounted at the rate  $\rho$  is finite. The solution to the stopping problem is to exit when  $s$  reaches an exit barrier  $B$ , defined by

$$e^B = \frac{\xi}{1 + \xi} \left( 1 - \frac{\mu + \sigma^2/2}{\rho} \right), \quad \xi = \frac{\mu}{\sigma^2} + \sqrt{\left( \frac{\mu}{\sigma^2} \right)^2 + \frac{\rho}{\sigma^2/2}}. \quad (3)$$

The resulting value function is

$$V(s) = \frac{1}{\rho} \frac{\xi}{1 + \xi} \left( e^{s-B} - 1 - \frac{1 - e^{-\xi(s-B)}}{\xi} \right) \quad (4)$$

for all  $s \geq B$  and  $V(s) = 0$  otherwise (Dixit and Pindyck [9], Luttmer [21]). Observe that this value function, measured in units of labor, only depends on  $\rho/\sigma^2$  and  $\mu/\sigma^2$ .

**Remark.** The exponent  $\xi$  and the value function  $V(\cdot)$  are increasing in  $\mu$ . The exit barrier  $B$  is decreasing in  $\mu$ .

Rapid firm growth lowers the exit barrier  $B$  and raises the value of a firm. The level of productivity  $X_t$  at which exit at time  $t$  takes place can be written as

$$X_t = Z_t e^{-(1-\beta)(S-B)}. \quad (5)$$

The exit level of productivity trends up with entry productivity and wages. Firms that cannot keep up are driven out of business.

### 3.4. The size distribution

There is a continuum of firms measured by  $N_t$  at time  $t$ . Assuming that  $S > B$ , firms can enter at  $s_0 = S$ , evolve according to  $ds_a = \mu da + \sigma dW_a$ , and then exit if  $s_a$  reaches  $B$ . Without entry, there is a continuous flow of firms that exit at  $B$ , and the number of firms declines. Sufficiently large entry rates result in a rising number of firms. The following gives conditions under which a stationary size distribution exists and describes the relation between the entry rate  $\varepsilon$  and the resulting growth rate  $\omega$  of the number of firms. Along the balanced growth paths to be constructed later, the number of firms has to grow at the same rate  $\eta$  as aggregate employment and the entry rate will be positive.<sup>4</sup>

<sup>4</sup> Examples in Luttmer [21,22] exhibit balanced growth with no entry and a non-stationary size distribution that spreads out forever. But there are no fixed costs in those examples.

### 3.4.1. Constant entry rates

Suppose there is a non-negative flow  $\varepsilon N_t$  of new firms entering at time  $t$ . Conjecture that there is a stationary size distribution with a continuous density  $f$  so that the number of firms grows at some constant rate  $\omega$ . Away from any boundaries or entry points, the density  $e^{\omega t} f(s)$  must satisfy the Kolmogorov forward equation  $\partial[e^{\omega t} f(s)]/\partial t = -\partial[\mu e^{\omega t} f(s)]/\partial s + \frac{1}{2}\partial^2[\sigma^2 e^{\omega t} f(s)]/\partial s^2$  (see for example Cox and Miller [7] or Feller [11]). This gives

$$\omega f(s) = -\mu Df(s) + \frac{1}{2}\sigma^2 D^2 f(s), \quad s \in (B, S) \cup (S, \infty). \quad (6)$$

The conjectured continuity means that at the entry point  $S$ ,

$$\lim_{s \uparrow S} f(s) = \lim_{s \downarrow S} f(s). \quad (7)$$

Exit at  $B$  and the need for  $f$  to be integrable imply

$$f(B) = f(\infty) = 0. \quad (8)$$

The flow of firms exiting at  $B$  per unit of time is given by  $\frac{1}{2}\sigma^2 Df(B)$ . This can be shown by integrating the Kolmogorov forward equation for a particular age cohort and aggregating over all age cohorts. More heuristically, over a small enough interval of time  $\Delta$ , the diffusion term in  $ds_a = \mu da + \sigma dW_a$  dominates, and half of the approximate measure  $f(B + \sigma\sqrt{\Delta}) \times \sigma\sqrt{\Delta}$  of firms near the boundary  $B$  will exit. For the number of firms to grow at the rate  $\omega$ , the entry rate must be

$$\varepsilon = \omega + \frac{1}{2}\sigma^2 Df(B). \quad (9)$$

Finally,  $f$  must integrate to 1.

Integrating (6) using (7) and (8) gives

$$\omega = \frac{1}{2}\sigma^2 [D_- f(S) - Df(B) - D_+ f(S)]$$

where  $D_- f(S)$  and  $D_+ f(S)$  are the left and right derivatives of  $f$  at  $S$ , respectively. The entry condition (9) is therefore equivalent to

$$\varepsilon = \frac{1}{2}\sigma^2 [D_- f(S) - D_+ f(S)]. \quad (10)$$

Thus, if the entry rate is positive,  $f$  will have a kink at  $S$ .

The differential equation (6) is linear and homogeneous with constant coefficients. On  $(B, S)$  and  $(S, \infty)$ , it has solutions that are linear combinations of  $e^{-\alpha s}$  and  $e^{\alpha_* s}$ , where

$$\alpha = -\frac{\mu}{\sigma^2} + \sqrt{\left(\frac{\mu}{\sigma^2}\right)^2 + \frac{\omega}{\sigma^2/2}}, \quad \alpha_* = \frac{\mu}{\sigma^2} + \sqrt{\left(\frac{\mu}{\sigma^2}\right)^2 + \frac{\omega}{\sigma^2/2}}. \quad (11)$$

Note that  $\alpha\alpha_* = \omega/(\sigma^2/2)$ . If  $\alpha$  and  $\alpha_*$  are complex, then any real-valued linear combination of  $e^{-\alpha s}$  and  $e^{\alpha_* s}$  will change signs indefinitely on  $(S, \infty)$ . If  $\alpha$  is real and non-positive then  $\alpha_*$  is real and non-negative. In that case, no linear combination of  $e^{-\alpha s}$  and  $e^{\alpha_* s}$  converges to zero as  $s$  becomes large. Stationary densities can be constructed only if  $\alpha$  is positive, or

$$\omega \geq -\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2 \quad \text{if } \mu < 0, \quad \omega > 0 \quad \text{if } \mu \geq 0. \quad (12)$$

This defines a lower bound on how fast a stationary population of firms can decline. There can be no stationary distribution with a declining population of firms if  $\mu$  is non-negative.

**3.4.1.1. A zero entry rate.** Without entry, the population of firms can only decline, and thus  $\mu$  must be negative for a stationary distribution to exist. At  $\varepsilon = 0$ , (10) implies that  $f$  is differentiable at  $S$ . Solving (6) with the boundary conditions  $f(B) = 0$  and  $f(\infty) = 0$  gives

$$f(s) = \alpha(-\alpha_*)e^{-\alpha(s-B)} \left( \frac{e^{(\alpha+\alpha_*)(s-B)} - 1}{\alpha + \alpha_*} \right). \quad (13)$$

This is a well-defined density if and only if  $\alpha > -\alpha_* > 0$ . Given that  $\mu$  is negative, this will be the case for any  $\omega$  in  $[-(\mu/\sigma)^2/2, 0)$ . Thus there is a multiplicity of growth rates  $\omega$  and associated stationary densities  $f$ . At the boundary  $\omega = -(\mu/\sigma)^2/2$  the density (13) becomes  $f(s) = \alpha^2(s-B)e^{-\alpha(s-B)}$ , with  $\alpha = -\mu/\sigma^2$ . An argument given in Luttmer [21] indicates that this is the limiting distribution when the initial size distribution has a compact support.<sup>5</sup>

**3.4.1.2. A positive entry rate.** With positive entry, (10) implies that  $f$  is not differentiable at  $S$ . Thus (6) defines two second-order differential equations, one on  $(B, S)$  and one on  $(S, \infty)$ . Conditions for the endpoints are given in (7) and (8), except that the level of  $f(S)$  is unrestricted. But  $f$  integrates to 1. Ignoring (9) and (10), this implies

$$f(s) = \frac{\alpha\alpha_*e^{-\alpha(s-B)}}{e^{\alpha_*(S-B)} - 1} \min \left\{ \frac{e^{(\alpha+\alpha_*)(s-B)} - 1}{\alpha + \alpha_*}, \frac{e^{(\alpha+\alpha_*)(S-B)} - 1}{\alpha + \alpha_*} \right\} \quad (14)$$

for all  $s \geq B$ . This is a well-defined density as long as  $\alpha > 0$ , which is equivalent to (12). Conditional on  $s \geq S$ , the resulting distribution for  $e^s$  is Pareto, and  $\alpha$  is its tail index.

The  $\omega$  and  $f$  that correspond to a particular entry rate  $\varepsilon > 0$  can be determined by imposing (9) or (10). From (14),

$$\frac{1}{2}\sigma^2 Df(B) = \frac{\frac{1}{2}\sigma^2\alpha\alpha_*}{e^{\alpha_*(S-B)} - 1}, \quad (15)$$

which is positive because  $\alpha$  is positive. Imposing (9) and simplifying gives

$$\varepsilon = \frac{\omega}{1 - e^{-(S-B)\left(\frac{\mu}{\sigma^2} + \sqrt{\left(\frac{\mu}{\sigma^2}\right)^2 + \frac{\omega}{\sigma^2/2}}\right)}}. \quad (16)$$

Only real solutions for  $\omega$  can be interpreted as growth rates. Fig. 1 shows (16) for various  $\mu$ . For  $\mu < 0$  the  $\varepsilon = 0$  entry rate that corresponds to (13) is also indicated.

The right-hand side of (16) is strictly increasing in  $\omega$  and approaches  $\omega$  from above for large  $\omega$ . Exit at  $B$  becomes negligible if  $\omega$  is large. If  $\mu$  is non-negative, then the right-hand side of (16) starts at 0 when  $\omega = 0$ . Thus a positive entry rate implies positive growth rate  $\omega$ , and it is unique. If  $\mu$  is negative, then the right-hand side of (16) attains a positive minimum when  $\omega < 0$  reaches the lower bound (12). The entry rate has to be large enough. If it is, there is a unique  $\omega$ , and this growth rate turns positive when  $\varepsilon$  rises above  $(S-B)/(-\mu)$ . Conversely, given a growth rate  $\omega$  and an associated stationary density  $f$ , one can simply compute  $\varepsilon$  from (9). The conditions for existence of a stationary density  $f$  given  $\omega$  can be summarized as follows.

**Proposition 1.** Assume  $S > B$ . If  $\omega$  satisfies (12) then there is a stationary density  $f$  for which the number of firms grows at the rate  $\omega$ . The associated entry rate  $\varepsilon$  is given by (9). If  $\omega < 0$

<sup>5</sup> If firms replicate themselves at a rate  $\varepsilon$  and the population of firms grows at the rate  $\eta$  as a result, then the size density will satisfy (6) with  $\omega = \eta - \varepsilon$ . This is an alternative interpretation for (13), and the stability argument gives  $\eta = \varepsilon - (\mu/\sigma)^2/2$ .

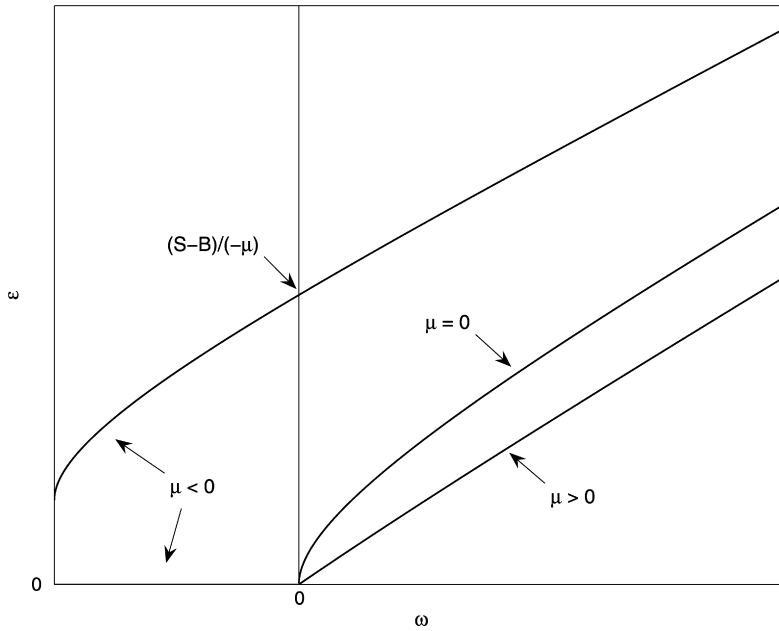


Fig. 1. Entry and growth of the number of firms.

then there are two stationary densities, one of which implies  $\varepsilon = 0$ . If  $\omega \geq 0$  then the stationary density is unique.

It is possible to take a limit in (14) as  $S \downarrow B$ , holding  $\omega$  fixed. This yields  $f(s) = \alpha e^{-\alpha(s-B)}$  for any  $s > B$ . This means that the limiting distribution of  $e^s$  on  $[B, \infty)$  is Pareto. The expression for the exit rate (15) implies that  $\frac{1}{2}\sigma^2 Df(B) \rightarrow \infty$  as  $S \downarrow B$ , and hence the required amount of entry  $\varepsilon$  is infinite as well. The above limiting density is also the density of a regulated Brownian motion on  $[B, \infty)$ . The infinite entry rate matches the fact that the regulator process needed to keep this Brownian motion above  $B$  is not a differentiable function of time (see Harrison [15]).

### 3.5. The mean of $e^s$

Variable firm employment is proportional to  $e^s$ . Given a positive measure of firms,  $e^s$  has to have a finite mean, or else aggregate employment would be infinite. Since  $f(s)$  behaves like  $e^{-\alpha s}$  for large  $s$ ,  $e^s$  has a finite mean if and only if  $\alpha > 1$ . If there is positive entry at  $S$ ,

$$\int_B^\infty e^{s-B} f(s) ds = \frac{\alpha \alpha_*}{(\alpha - 1)(\alpha_* + 1)} \frac{e^{(S-B)(\alpha_*+1)} - 1}{e^{\alpha_*(S-B)} - 1}, \quad (17)$$

and this diverges as  $\alpha \downarrow 1$ . If  $\omega = 0$  this simplifies to  $(\alpha/(\alpha - 1)) \times (e^{S-B} - 1)/(S - B)$ , which is decreasing in  $\alpha$ , and therefore increasing in  $\mu$ . Calculations reported in Appendix A show that this holds for  $\omega \neq 0$  as well.

**Lemma 1.** Holding fixed  $B$  and  $S$ , the mean of  $e^s$  is increasing in  $\mu$ .



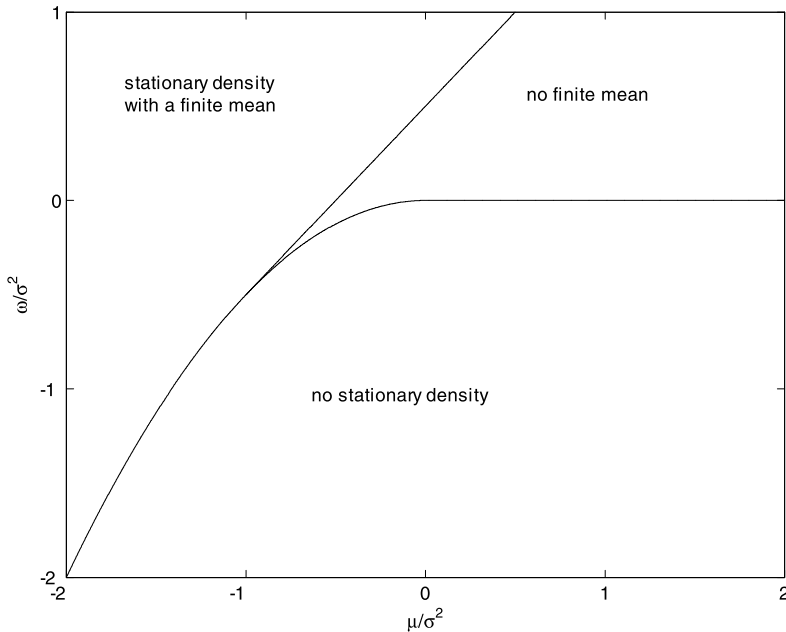


Fig. 2. Conditions for stationarity and a finite mean.

The restriction  $\alpha > 1$  holds automatically if the condition for stationarity (12) is met and  $\mu/\sigma^2 < -1$ . For  $\mu/\sigma^2 \geq -1$ , the restriction  $\alpha > 1$  is equivalent to

$$\omega > \mu + \frac{1}{2}\sigma^2. \quad (18)$$

The conditions for stationarity and a finite mean are shown in Fig. 2. For  $\mu/\sigma^2 > -1$ , the requirement that the mean be finite is strictly stronger than the requirement (12) for stationarity. If  $\omega \geq 0$ , then Fig. 2 implies that (18) is necessary and sufficient for the existence of a stationary distribution with a finite mean. Note that

$$\lim_{\Delta \downarrow 0} E_a[e^{s_a + \Delta - s_a}] = \mu + \frac{1}{2}\sigma^2$$

for any  $s_a > B$ . Thus (18) says that the number of firms must grow faster than variable employment at incumbent firms that are not about to exit. Clearly, for given  $\mu$  and  $\sigma^2$ , it is a condition that can be met by having the number of firms grow fast enough.

#### 4. Entry and technology adoption

Suppose now that entry productivity  $Z_t = Ze^{\kappa t}$  is completely exogenous. Since  $\theta$  is also exogenous, this implies that  $\mu$  is an exogenous parameter.

Although anyone can access  $Z_t$ , to set up new firms does take labor and time. Applying  $\lambda_E$  units of labor continuously generates a new firm following an exponentially distributed waiting time with mean 1. Suppose the project of setting up a firm is initiated at time  $t$  and leads to success at time  $t + \tau$ . The cost of labor is  $w_{t+a}\lambda_E$  per unit of time,  $a \in [0, \tau]$ , and the value of the firm will be  $w_{t+\tau}\lambda_F V(S)$  when it arrives at time  $t + \tau$ . The interest rate is  $r = \rho + \kappa$  and

wages grow at the rate  $\kappa$ . The present-value of the project is thus  $(\lambda_F V(S) - \lambda_E)/(1 + \rho)$ . Since anyone can start a project, these profits cannot be positive.

Without entry, the number of firms has to decline. If the distribution of firms is stationary, then aggregate firm employment declines as well. But we are assuming that the population grows at a non-negative rate  $\eta$  and everyone has one unit of labor. This means there can be no balanced growth path without entry. Thus entry is positive, and the profits from a project to start a firm must be zero. That is,

$$\lambda_E = \lambda_F V(S). \quad (19)$$

Since  $V(s)$  is increasing and convex, with an asymptote that behaves like  $e^s$ , a unique solution for  $S$  is guaranteed. This solution varies with entry and fixed costs only to the extent that  $\lambda_E/\lambda_F$  is affected.

If the number of firms  $N_t$  grows at some rate  $\omega$  that satisfies (18), then so does the amount of labor  $\lambda_E \varepsilon N_t$  required to set up new firms, as well as the amount of labor employed by all firms. The only way this can be part of a balanced growth path is if  $\omega = \eta$ . The resulting entry rate  $\varepsilon$  is then determined by (9). Recall from (1) that a firm with productivity  $z$  employs  $\lambda_F + l_t[z]$  units of labor. Labor market clearing therefore requires

$$\frac{H}{N} = \lambda_E \left( \eta + \frac{1}{2} \sigma^2 Df(B) \right) + \lambda_F \left( 1 + \frac{\beta}{1 - \beta} \int_B^\infty e^s f(s) ds \right). \quad (20)$$

A balanced growth path can now be constructed as follows. The density  $f$  is defined by (11) and (14), evaluated at  $\omega = \eta$ , and given  $B$  and  $S$ . The exit boundary  $B$  is determined by (3) and the entry size  $S$  is determined by the zero-profit condition (19). The labor market clearing condition (20) then determines the number of firms. The level of wages follows from  $S$  and its definition (2).

This construction works only as long as (18) holds for  $\omega = \eta$ . Without this condition, the right-hand side of (20) would not be finite, and this is inconsistent with a positive measure of firms.

## 5. Endogenizing growth

There can be no balanced growth path if (18) is violated at  $\omega = \eta$ . Condition (18) depends on parameters that have so far all been taken as exogenous. In this section, the growth rate  $\kappa$  of the entry productivity of new firms is endogenous. This makes  $\mu = (\theta - \kappa)/(1 - \beta)$  endogenous and sets the stage for finding equilibrium mechanisms that ensure the existence of a balanced growth path.<sup>6</sup>

### 5.1. A spillover at the bottom

Continue to assume that the drift of incumbent productivity is given exogenously by  $\theta$ . As before, suppose that entrepreneurs can expend a flow of  $\lambda_E$  units of labor to create new firms at a unit rate. Rather than assuming entrant productivity grows exogenously, suppose that an entrepreneur who creates a new firm can copy the technology of firms that are about to exit, and

<sup>6</sup> The parameter  $\theta$  is made endogenous in Atkeson and Burstein [5], where firms can raise  $\theta$  at a cost. See Luttmer [22] for a discussion.

improve the productivity of this technology by a factor  $\Gamma^{1-\beta}$ , for some  $\Gamma > 1$ . The resulting entry productivity is

$$Z_t = \Gamma^{1-\beta} X_t.$$

As in Arrow [4], this introduces a knowledge spillover. In contrast to the quality-ladder model of Grossman and Helpman [14], new firms here do not enter at the top of a productivity ladder, ahead of everyone else. Instead, assuming  $\Gamma$  is relatively close to 1, entrants only get to skip the bottom few rungs, occupied by firms that are about to exit. From there they have to climb the productivity ladder, by trial and error.<sup>7</sup>

Recall from (5) that  $X_t/Z_t = e^{-(1-\beta)(S-B)}$ , and hence

$$S = B + \ln(\Gamma). \quad (21)$$

The zero-profit condition for entrepreneurs is still  $\lambda_E = \lambda_F V(S)$ . Combining this with the solution (4) for  $V(\cdot)$  and (21) yields

$$\lambda_E = \frac{\lambda_F}{\rho} \frac{\xi}{1+\xi} \left( \Gamma - 1 - \frac{1 - \Gamma^{-\xi}}{\xi} \right). \quad (22)$$

This is a function of  $\xi$ , which in turn is a function only of  $\kappa$ , via  $\mu = (\theta - \kappa)/(1 - \beta)$ . Thus the requirement that entrepreneurs make zero profits implies an equilibrium condition that depends only on the balanced growth rate  $\kappa$ .

The right-hand side of (22) approaches 0 from above as  $\xi \downarrow 0$ . It is increasing in  $\xi$ , with a large- $\xi$  asymptote equal to  $(\Gamma - 1)\lambda_F/\rho$ . The size of the productivity improvement entrants can make puts an upper bound on the value of a new firm. Thus (22) will have a solution for  $\xi$  as long as

$$\frac{\lambda_E}{\lambda_F} < \frac{\Gamma - 1}{\rho},$$

and this solution is unique. Entry costs cannot be too high, or the improvements available to entrants over the firms that are just exiting cannot be too small. In particular, taking a limit and letting  $\Gamma \downarrow 1$ , so that  $S \downarrow B$ , is guaranteed to rule out a solution to (22) unless entry costs also go to zero.

Although  $\mu$  is now an equilibrium variable, nothing about the equilibrium condition (22) guarantees that (18) holds at  $\omega = \eta$ . Solving the definition (3) of  $\xi$  for  $\mu$  and combining the solution with the restriction (18) gives

$$\frac{1}{\xi} \left( \frac{1}{2} \sigma^2 \xi (1 + \xi) - \rho \right) < \eta. \quad (23)$$

Fig. 3 shows the left-hand side of (23) as a function of  $\xi$ , together with  $\rho/\lambda_F$  times the right-hand side of the zero-profit condition (22), for  $\Gamma = 2$ ,  $\sigma = .2$  and  $\rho = .02$ . If  $\eta = .01$  then  $\xi$  cannot be much above .78, and hence  $\rho\lambda_E/\lambda_F$  cannot be much above .2. Slightly higher values of  $\rho\lambda_E/\lambda_F$

<sup>7</sup> For alternative models of spillovers with productivity distributions, see Kortum [18], Eaton and Kortum [10], Luttmer [21], Lucas [20], and Alvarez, Buera and Lucas [3].

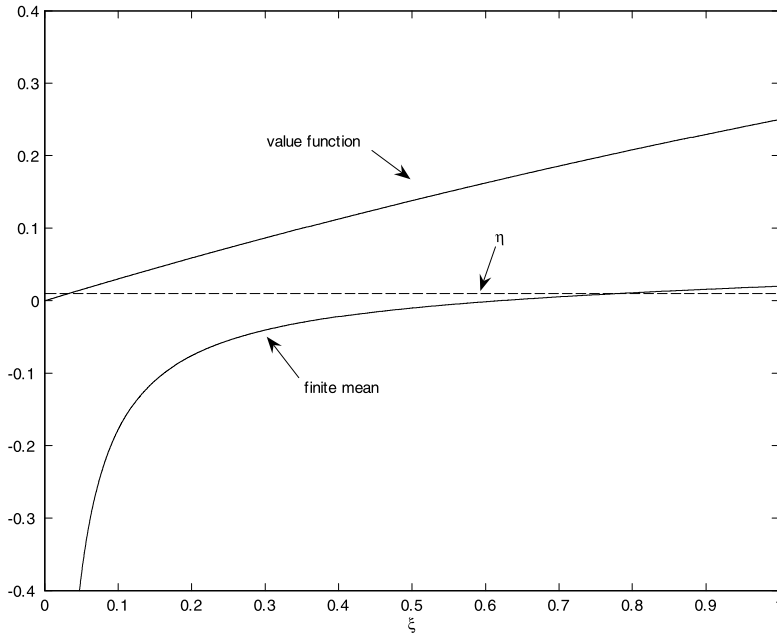


Fig. 3. Existence of a balanced growth path.

lead to non-existence of a balanced growth path, even though  $\lambda_E/\lambda_F$  is still much smaller than  $(\Gamma - 1)/\rho$  and (22) has a unique solution.

### 5.2. The random imitation alternative

One possible mechanism for ensuring that (18) holds in equilibrium is given in Luttmer [21]. There, entrants at time  $t$  do not start with a common  $Z_t = \Gamma^{1-\beta} X_t$ , but with idiosyncratic  $\delta^{1-\beta} z$ , where the  $z$  are random draws from the population of incumbent producers. The fact that entrepreneurs draw at random from the time- $t$  population means that their incentives to enter are driven by the population average of  $V(s_t[z])$ . Since  $V(s_t[z])$  behaves like  $e^{s_t[z]}$ , this population average will explode precisely when  $\mu + \sigma^2/2$  approaches  $\eta$  from below. Imposing the zero-profit condition therefore forces the equilibrium growth rate to be such that the mean of  $e^s$  is finite, which corresponds to condition (18). Given this, the right-hand side of (20) is finite, and one can use this condition to solve for the number of firms  $N$ .<sup>8</sup>

This captures the intuition that the presence of large and profitable firms should induce entry and reduce the growth rate of incumbent firms. Note also that even though potential entrants are drawing incumbents randomly, actual entry is selective if  $\delta < 1$ : all random draws at time  $t$  satisfy  $z > X_t$ , but only those potential entrants for whom  $\delta^{1-\beta} z > X_t$  actually enter. A difficulty with this mechanism is that  $\delta^{1-\beta} z$  will still be very large for some fraction of the entrants, even if the parameter  $\delta$  is set far below 1 so that the ability of entrants to imitate is very poor. In the data, almost without exception, new firms are very small.

<sup>8</sup> If  $\delta = 0$ , the size distributions consistent with balanced growth are the ones defined in (13). There is entry, not at a point  $S$ , but throughout  $(B, \infty)$ , with an intensity that is proportional to the incumbent size density.

## 6. Less-than-perfectly elastic entry

In the economies described so far, the zero-profit condition for entrepreneurs is  $\lambda_E = \lambda_F V(S)$  and  $V(\cdot)$  depends on  $\mu$ . With exogenous technological progress,  $\mu$  is exogenous and the zero-profit condition determines  $S$ . In the economy with a spillover at the bottom,  $S = B + \ln(\Gamma)$  and the zero-profit condition pins down  $\mu$ , and hence the growth rate  $\kappa = \theta - (1 - \beta)\mu$  of the economy. But it does so without reference to the labor market clearing condition (20). A balanced growth path may then not exist because there is no guarantee that  $\mu + \sigma^2/2 < \eta$ .

In a setting of exogenous aggregate growth but endogenous firm growth, Luttmer [23] avoids this problem by using a Roy model of occupational choice to make the entry of new firms less than perfectly elastic with respect to the market value of new firms. As a result, the number of firms and their growth rate are jointly determined by an entry condition and a labor market clearing condition. This forces average firm labor demand to be finite. This device can also be used here, with endogenous aggregate growth and exogenous growth of incumbent firms.

### 6.1. Workers and entrepreneurs

Suppose that agents in the economy vary in their skills in supplying labor and creating new firms. Specifically, suppose a type- $(x, y)$  agent can supply  $x$  units of labor or create new firms at a Poisson rate  $y$ . The distribution of talent in the population is described by some distribution function  $T(x, y)$ . Comparative advantage determines occupational choice. Wages are  $w_t$  per unit of labor and, along a balanced growth path, the value of a new firm is  $w_t \lambda_F V(S)$ . In units of labor, this equals

$$q = \lambda_F V(S).$$

A type- $(x, y)$  agent will choose to be a worker if  $qy < x$  and an entrepreneur if  $qy > x$ . The per-capita supplies of labor and new firms are therefore

$$\begin{pmatrix} L(q) \\ E(q) \end{pmatrix} = \int \begin{pmatrix} x 1\{qy < x\} \\ y 1\{qy \geq x\} \end{pmatrix} dT(x, y).$$

Suppose that  $T$  is continuous with a finite mean and full support in  $\mathbb{R}_{++}^2$ . This implies that the supplies of labor and new firms are positive and vary continuously with  $q$ . Clearly,  $L(q)$  is a decreasing function and  $E(q)$  an increasing function. As  $q$  goes to zero,  $E(q)$  goes to zero while  $L(q)$  converges to the mean endowment of labor. On the other hand, if  $q$  becomes large, then the per-capita labor supply goes to zero while the supply of new firms grows to the mean of ability to create new firms. Hence  $L(q)/E(q) \rightarrow \infty$  as  $q \rightarrow 0$  and  $L(q)/E(q) \rightarrow 0$  as  $q \rightarrow \infty$ . Thus the relative supplies of labor and new firms vary inversely with  $q$  and range throughout  $(0, \infty)$ . The ratio  $L(q)/E(q)$  will be proportional to a negative power of  $q$  if labor and entrepreneurial skills are given by two independent Fréchet distributions.

### 6.2. Balanced growth

Along a balanced growth path, the de-trended supply of new firms is  $E(q)H = \varepsilon N$ , and this supply must sustain a number of firms that grows at the rate  $\eta$ . Accounting for exit using (9), this means that

$$\frac{N}{H} = \frac{E(q)}{\eta + \frac{1}{2}\sigma^2 Df(B)}. \quad (24)$$

Labor is now used only by incumbent firms, and hence the labor market clearing condition simplifies from (20) to

$$\frac{N}{H} = \frac{L(q)}{\lambda_F(1 + \frac{\beta}{1-\beta} \int_B^\infty e^s f(s) ds)}. \quad (25)$$

New firms enter with a productivity that implies  $S = B + \ln(\Gamma)$ , as in (21). Combining this with  $q = \lambda_F V(S)$  yields

$$q = \frac{\lambda_F}{\rho} \frac{\xi}{1 + \xi} \left( \Gamma - 1 - \frac{1 - \Gamma^{-\xi}}{\xi} \right). \quad (26)$$

The parameter  $\xi$  and the exit barrier  $B$  are functions only of  $\mu$ , defined in (3). Hence  $q$  is a function only of  $\mu$ . Because of  $S = B + \ln(\Gamma)$ , the stationary density  $f$ , defined in (14), also depends only on  $\mu$ . Using (3), (14) and  $S = B + \ln(\Gamma)$ , the equilibrium conditions (24)–(26) therefore jointly determine the value of a new firm  $q$ , the number of firms  $N$ , and the employment growth rate  $\mu + \sigma^2/2$  of incumbent firms. From there the rest of the balanced growth path follows. If there is a solution to (24)–(26), then (25) forces  $\mu$  to be such that  $\mu + \sigma^2/2 < \eta$ .

One can view (24) and (25) as, respectively, the balanced growth supply and demand of firms. The supply depends on the entrepreneurial supply of new firms and the rate at which new firms need to be created to ensure the number of firms grows at the rate  $\eta$ . The demand for firms depends on the supply of labor and the size of the average firm. Both supply and demand depend directly on the price of new firms, through the occupational choices made by agents, and also indirectly: via  $\mu$ , the price of a new firm depends on the growth rate of the economy, and this growth rate determines the exit rate and average firm demand for labor.

### 6.2.1. Existence and uniqueness

A balanced growth path exists and is unique if the demand and supply of firms intersect uniquely. The following lemma collects some key derivatives that can be used to show that this is indeed the case.

**Lemma 2.** *Given the definitions (3), (14), and (26), of  $B$ ,  $f$  and  $q$ ,*

$$\frac{\partial}{\partial \mu} e^B < 0, \quad \frac{\partial q}{\partial \mu} > 0, \quad \frac{\partial}{\partial \mu} Df(B) < 0, \quad \frac{\partial}{\partial \mu} \int_B^\infty e^s f(s) ds > 0,$$

when  $\omega = \eta$  and the entry condition  $S = B + \ln(\Gamma)$  holds.

More rapid firm growth implies a lower exit barrier  $B$  and a higher value  $q$  for new firms that enter with  $S = B + \ln(\Gamma)$ . The effect on  $B$  follows immediately from (3), and the effect on  $q$  from (3) and (26). Evaluating the entry rate (9) at  $\omega = \eta$  and using (15) and  $S = B + \ln(\Gamma)$  gives  $\varepsilon = \eta + \frac{1}{2}\sigma^2 Df(B) = \eta/(1 - \Gamma^{-\alpha*})$ , and this is easily seen to be decreasing in  $\mu$  when  $\Gamma > 1$ . Rapidly growing firms are less likely to exit and so it takes less entry to maintain a growing population of firms. As shown in Lemma 1, holding fixed  $B$  and  $S$ , the mean of  $e^s$  is increasing in  $\mu$ . But here  $B$  declines and  $S - B = \ln(\Gamma)$ , which tends to lower the mean of  $e^s$ . It is shown in Appendix B that the effect described in Lemma 1 dominates.

The supply of firms (24) is increasing in  $\mu$ , since  $E(q)$  is an increasing function,  $q$  is increasing in  $\mu$ , and the entry rate  $\varepsilon$  is decreasing in  $\mu$ . Furthermore, letting  $\mu \rightarrow -\infty$  will cause  $q$

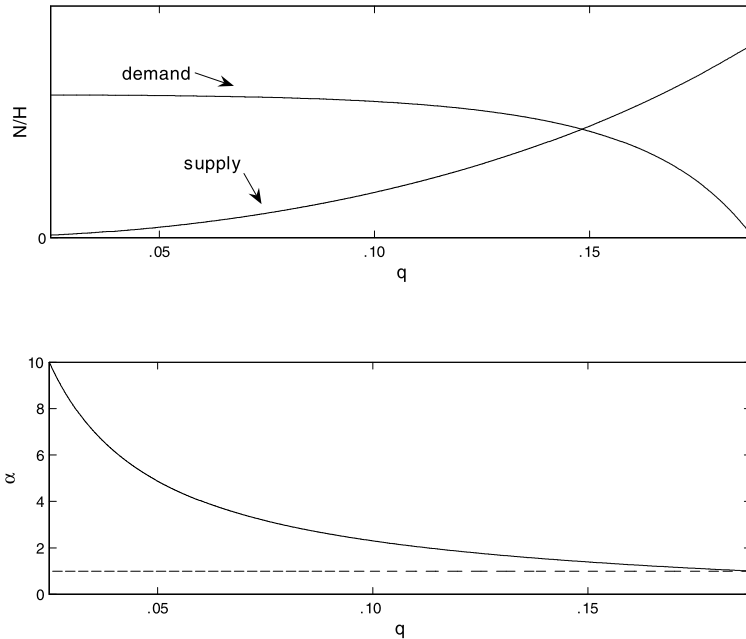


Fig. 4. The supply and demand of firms.

to approach 0 and  $\varepsilon = \eta/(1 - \Gamma^{-\alpha_*})$  to approach  $\infty$ . Thus the supply of firms can be made arbitrarily close to 0 by taking  $\mu$  sufficiently small.

The demand for firms (25) is decreasing in  $\mu$ , since  $L(q)$  is a decreasing function,  $q$  is increasing in  $\mu$ , and the mean firm size is increasing in  $\mu$ . In addition, letting  $\mu + \frac{1}{2}\sigma^2$  approach the upper bound  $\eta < \rho$  will cause the mean firm size to diverge, while  $q$  remains well defined (the value of a firm only diverges when  $\mu + \frac{1}{2}\sigma^2$  reaches  $\rho$ ) and  $L(q)$  remains bounded. Thus the demand for firms will shrink to zero for  $\mu + \frac{1}{2}\sigma^2$  close to  $\eta$ .

**Proposition 2.** *There is a unique equilibrium growth rate  $\kappa = \theta - (1 - \beta)\mu$ .*

The demand and supply of firms are shown in Fig. 4 as a function of  $q$ , which is a monotone function of  $\mu$  implied by (3) and (26). The second panel of Fig. 4 shows the relation between the tail index  $\alpha$  and the price  $q$  of new firms.

#### 6.2.2. Incumbent and entrant innovation

The incumbent productivity growth rate  $\theta$  only appears in the expression  $\kappa = \theta - (1 - \beta)\mu$ , and not in the equilibrium conditions represented in Fig. 4. A change in the growth rate of incumbent productivity therefore translates one-for-one into a change in the growth rate of aggregate productivity. Such a change leaves no mark on the cross-sectional distribution of productivity, the resulting size distribution of firms, entry and exit rates, or the market value in units of labor of new and existing firms.

A change in the distribution of skills in this economy does give rise to shifts in the demand and supply curves in Fig. 4. This typically has growth effects, not just level effects, and it will affect the shape of the size distribution of firms. An increase in the supply of entrepreneurial

effort lowers the equilibrium price  $q$ . This implies an increase in  $\alpha$  and thus a decrease in  $\mu$ , leading to an increase in the growth rate  $\kappa$  of aggregate productivity. More entrepreneurial effort at a given  $q$  implies a size distribution with a thinner tail and a higher growth rate of aggregate productivity. The tail index of the size distribution will be close to the  $\alpha = 1$  asymptote if the supply of entrepreneurial effort is sufficiently low. This is one possible interpretation of the fact that  $\alpha$  is close to 1 in the data. By Lemma 2, it is an interpretation that is associated with relatively low entry rates.

A change in the ability of entrepreneurs to improve the technology of exiting firms also affects the growth rate of aggregate productivity and the size distribution of firms. To examine the effects of an increase in the step size  $\Gamma$  available to entrepreneurs, eliminate  $q$  from (24) and (25) using (26), to obtain a supply of firms that is increasing in  $\mu$  and a demand for firms that is decreasing in  $\mu$ , by Lemma 2. How these supply and demand curves shift is determined by the fact that, holding fixed  $\mu$  and imposing  $S = B + \ln(\Gamma)$ ,

$$\frac{\partial q}{\partial \Gamma} > 0, \quad \frac{\partial}{\partial \Gamma} Df(B) < 0, \quad \frac{\partial}{\partial \Gamma} \int_B^{\infty} e^s f(s) ds > 0. \quad (27)$$

The slope of  $q$  with respect to  $\Gamma$  is immediate from (26). The entry rate  $\eta + \frac{1}{2}\sigma^2 Df(B) = \eta/(1 - \Gamma^{-\alpha_*})$  is clearly decreasing in  $\Gamma$ . A higher  $\Gamma$  means that new firms enter farther away from the exit barrier and thus survive longer. It then requires less entry to maintain a growing population of firms. The fact that the mean firm size is increasing in  $\Gamma$  follows from (17) evaluated at  $S - B = \ln(\Gamma)$ . With (27) in hand one can use (24)–(26) to argue that an increase in  $\Gamma$  causes the supply of firms to rise and the demand for firms to fall, for given  $\mu$ . It follows that  $\mu$  must decrease. An increase in  $\Gamma$  makes new firms more valuable, causing agents to substitute away from supplying labor (demanding firms) towards supplying entrepreneurial effort (supplying firms), and the effects on the supply and demand for firms are magnified by the fact that less entry is required and firms will be larger. Equilibrium can only be restored with a decline in  $\mu$ , and hence a rise in the growth rate  $\kappa = \theta - (1 - \beta)\mu$  of aggregate productivity. The resulting size distribution will have a thinner right tail. In this economy, endowing entrants with the ability to make larger productivity improvements over exiting firms causes aggregate growth to accelerate and reduces the number of relatively large firms in the population.

The observation that  $\alpha$  is increasing in  $\Gamma$  suggests an alternative interpretation of the fact that  $\alpha > 1$  is close to 1 in the data: although there may be an abundant supply of entrepreneurial effort, entrants can make only small productivity improvements over exiting firms.

**Proposition 3.** *The tail index  $\alpha$  approaches 1 from above as  $\Gamma$  shrinks to 1. The associated entry and exit rates grow without bound.*

To prove this, again consider the supply and demand for firms as a function of  $\mu$ . Observe that the supply (24) at any given  $\mu$  converges to zero as  $\Gamma \downarrow 1$ , for two reasons. First, by (26),  $q$  goes to zero as  $\Gamma \downarrow 1$  and holding fixed  $\mu$ . This means that the supply of entrepreneurial effort goes to zero. Second, the required entry rate  $\eta + \frac{1}{2}\sigma^2 Df(B) = \eta/(1 - \Gamma^{-\alpha_*})$  explodes as  $\Gamma$  approaches 1 from above. To examine the limiting demand for firms, note from (17) that the mean of  $e^{s-B}$  converges to  $\alpha/(\alpha - 1)$  as  $\Gamma \downarrow 1$ , for any given  $\mu + \sigma^2/2 < \eta$ . Since  $q$  goes to zero, the supply of labor converges to its maximum. In the limit, the demand for firms (25) becomes a function of  $\mu$  only via the effect of  $\mu$  on the average firm size. Because the mean of  $e^{s-B}$  diverges as  $\alpha \downarrow 1$ , this limiting demand curve still has the property that  $N/H \downarrow 0$  as



$\mu + \sigma^2/2 \uparrow \eta$ . Given that the supply converges to zero as  $\Gamma \downarrow 1$ , it follows that the intersection of the supply and demand curves must occur at a  $\mu + \sigma^2/2$  that approaches  $\eta$  from below. Since  $\alpha_*$  remains positive in the limit, entry and exit rates grow without bound.

### 6.2.3. Comparison with perfectly elastic entry

In terms of  $\mu$ , the structure of the equilibrium conditions (24)–(26) is

$$E(q) = G_1(\mu)N/H, \quad L(q) = G_2(\mu)N/H, \quad q = G_3(\mu),$$

where  $G_1(\mu)$  is the required entry rate,  $G_2(\mu)$  is the average firm demand for labor, and  $G_3(\mu)$  is the value of a new firm. By condition (18),  $G_2(\mu)$  is finite if and only if  $\mu + \sigma^2/2 < \eta$ . With less-than-perfectly-elastic entry,  $E(q)$  and  $L(q)$  vary continuously and Proposition 2 shows that a unique solution is guaranteed. With perfectly elastic entry, the pair  $[E(q), L(q)]$  equals  $[0, 1]$  if  $\lambda_E > q$  and  $[1/\lambda_E, 0]$  if  $\lambda_E < q$ . This means that  $\lambda_E = q = G_3(\mu)$  is the only candidate equilibrium. But  $G_2(\mu)$  may not be finite at the implied  $\mu$ , and in that case there will be no balanced growth path. The random imitation process in Luttmer [21] implies that  $G_2(\mu)$  is finite whenever  $G_3(\mu)$  is, ensuring the existence of a balanced growth path.

## 7. A calibration

Growth in the economy described above is the result of random technological improvements made by incumbents, and the replacement of the least successful firms by new entrants who can do better. What does the empirical evidence on firm heterogeneity say about the importance of these two components? Some very simple back-of-the-envelope calculations can be used to identify the parameters of this economy and answer this question.

US Census data compiled by the Small Business Administration show that the employment size distribution of employer firms is stable over time and close to Pareto for firms with more than about 20 employees. Based on 2002 data, Luttmer [21] reports a point estimate of 1.06 for the tail index  $\alpha$ . There is some uncertainty about this estimate. A value of 1.1 could work too, but larger values of  $\alpha$  are hard to reconcile with the observed number of large firms. The total number of firms grows at roughly the US population growth rate of around 1%, although there are some fluctuations. Entry of new firms is at a stable rate of about 11% per annum. Employment among firms with 500 or more employees grows at an annual rate of around .36% over the period 1988–2006, although this number is not precisely estimated (serially uncorrelated measurement or approximation error implies a standard error of about .38%). Consistent with the class of models considered here, not many of these firms exit over the course of a year.<sup>9</sup>

Solving the formula (11) for the tail index  $\alpha$  for  $\mu$  gives

$$\mu + \frac{1}{2}\sigma^2 = \frac{\eta}{\alpha} - \frac{1}{2}\sigma^2(\alpha - 1). \quad (28)$$

Over small periods of time, the left-hand side is the growth rate of employment among incumbent firms. A tail index close to 1 implies that this growth rate will be close to  $\eta$ , but the discrepancy will be larger the larger the variance of productivity shocks. If the employment growth rate of .36% per annum at firms with more than 500 employees is used to infer  $\mu + \sigma^2/2$ , then (28) yields  $\sigma = .44$ . Davis et al. [8] report estimates of the annual standard deviation of firm

<sup>9</sup> The number is about 2.5%, higher than one might expect if firms only exit at  $B$ . No doubt some of these exits are due to takeovers and mergers.

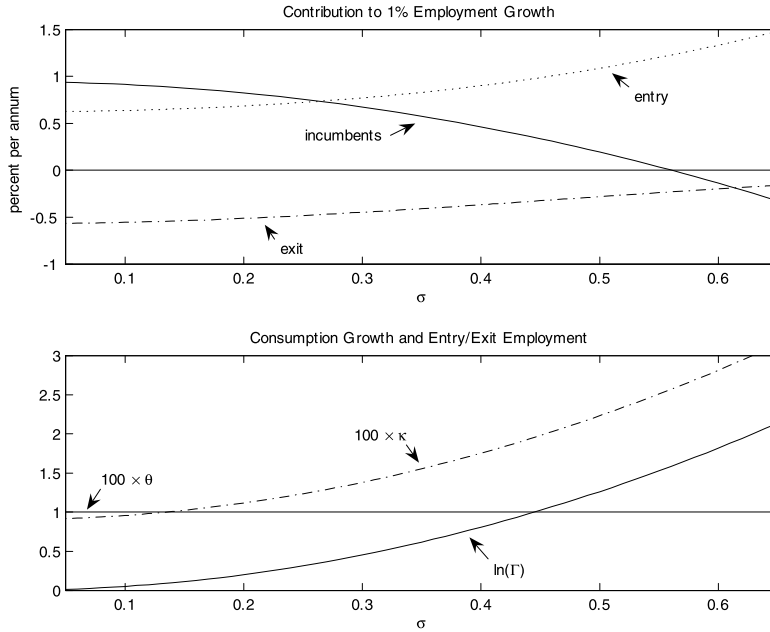


Fig. 5. Growth decomposition and implied spillover.

employment growth that range widely, roughly from 15% to 65%, depending on whether a firm is publicly traded or not, and on the sample period. Their most recent estimate for the whole economy is about 37%.

Recall that the entry rate is  $\varepsilon = \eta / (1 - \Gamma^{-\alpha^*})$  and that (11) implies  $\alpha\alpha^* = \eta / (\sigma^2/2)$ . Solving these two conditions for  $\Gamma$  gives

$$\ln(\Gamma) = \frac{1}{2}\sigma^2 \times \frac{\alpha}{\eta} \ln\left(\frac{\varepsilon}{\varepsilon - \eta}\right). \quad (29)$$

Thus  $\ln(\Gamma)$  is proportional to  $\sigma^2/2$ , with a coefficient that can be inferred from the tail index  $\alpha \approx 1.06$ , the population growth rate  $\eta \approx .01$ , and the entry rate  $\varepsilon \approx .11$ . The resulting coefficient is 10.1, and then  $\sigma \approx .44$  implies  $\ln(\Gamma) \approx .98$ . Hence, variable employment at entering firms is about 2.7 times variable employment at exiting firms. The implicit difference in productivity will be much smaller if  $\beta$  is close to 1 (there is no physical capital in this economy). At  $\beta = .9$ , entrants have a productivity advantage over exiting firms of about 10%, and the standard deviation  $\sigma_Z$  of incumbent productivity growth is about 4.4% per annum.

The first and second panels of Fig. 5 show the incumbent employment growth rate (28) and the spillover parameter (29) as functions of  $\sigma$ . Included in the first panel is a further decomposition of the growth rate  $\eta$  of aggregate employment,

$$\eta = \underbrace{\frac{\varepsilon e^S - (\varepsilon - \eta)e^B}{\int_B^\infty e^s f(s) ds}}_{\text{entry-exit}} + \underbrace{\mu + \frac{1}{2}\sigma^2}_{\text{incumbents}}.$$

This can be verified mechanically using (11) and (17). The first term is the entry rate times the ratio of entry employment over average employment, which accounts for variable employment

created by entry. The second term represents variable employment lost as a result of exiting firms, and the third term is the growth rate of employment at incumbent firms. In SBA data, the gross employment flows from entry and exit hover around 3% of aggregate employment, which significantly exceeds the gross flows shown in Fig. 5. In the data, the ratio of entry and exit employment over aggregate employment is higher than implied by  $\sigma$  in the range reported here. One likely reason for this is that  $\int_B^\infty e^{s-S} f(s) ds$  and  $\int_B^\infty e^{s-B} f(s) ds$  are very sensitive to small changes in the tail index  $\alpha$  near its asymptote  $\alpha = 1$ .

The second panel of Fig. 5 also shows the implied growth rate of per-capita consumption  $\kappa = \theta + (1 - \beta)(\frac{1}{2}\sigma^2\alpha - \eta/\alpha)$  at  $\beta = .9$ . The Cobb–Douglas technology used by all firms implies that a firm with productivity  $z$  produces  $y_t[z] = w_t l_t[z]/\beta$  units of consumption. The contributions of incumbent firms to aggregate consumption growth can therefore be calculated by simply adding the  $\kappa$  shown in the second panel to the  $\mu + \sigma^2/2$  shown in the first panel. The result is  $\kappa + \mu + \sigma^2/2 = \theta + \beta\eta/\alpha + (1 - \alpha\beta)\sigma^2/2$ . Thus the implied contribution of incumbent firms to aggregate consumption growth is linear and increasing in  $\sigma^2$ , with a small slope when  $\beta$  is close to 1.

At  $\sigma = .44$ , an incumbent employment growth rate of .36% implies  $\mu \approx -.0932$ . Thus, according to these estimates, the positive incumbent employment growth rate is the combination of a strong negative drift in the logarithm of incumbent employment, and a lot of re-allocation of employment driven by productivity shocks. This means that  $\kappa = \theta + (1 - \beta) \times .0932$ , and hence randomness and re-allocation add almost 1% to the logarithmic drift  $\theta$  of incumbent productivity growth if  $\beta = .9$ , or about half the 2% growth rate of US per-capita consumption. This is consistent with the important role for re-allocation found in Restuccia and Rogerson [25] and Hsieh and Klenow [16].

## 8. Concluding remarks

This paper shows that in an economy with random productivity growth at the firm level, it takes only a very weak spillover to produce balanced growth and a stationary firm size distribution. It is critical that there are always some agents for whom starting a new firm beats being a worker. Entrants need to be able learn from incumbents, but it is precisely when they can make only relatively small improvements over the technology used by the least productive incumbents that the striking Zipf-like features of the firm size distribution arise. Moreover, this is also how the model can account for the high entry and exit rates observed in the data.

Firm employment size in this paper is log-linear in firm productivity and entrants are small because of their low productivity. Luttmer [23] presents an alternative theory of the firm size distribution in which firms grow by stochastically replicating units of organization capital. Entrants are small because they have not yet had the chance to replicate their organization capital. Persistent but potentially very small productivity differences can give rise to persistent differences in the average rate at which organization capital is accumulated. Large firms tend to be those that have had a productivity advantage for a long time.

An advantage of this alternative theory is that stochastic replication provides a natural reason for firm growth rate variances to decline with size, as they do in the data. But aggregate productivity is exogenous in this alternative theory. Incorporating elements of both accounts of the firm size distribution into tractable and quantitatively plausible models of technology diffusion and endogenous growth remains an important task for further research.

## Appendix A. Proof of Lemma 1

Write (17) as

$$\int_B^\infty e^s f(s) ds = \frac{\alpha e^B}{\alpha - 1} \frac{e^{(1+\alpha_*)(S-B)} - 1}{\frac{e^{\alpha_*(S-B)} - 1}{\alpha_*(S-B)}} = \frac{\alpha e^B}{\alpha - 1} \times \frac{e^{S-B+z} - 1}{\frac{e^z - 1}{z}} \quad (30)$$

where  $z = \alpha_*(S - B)$ . The definitions of  $\alpha$  and  $\alpha_*$  imply that  $\partial\alpha/\partial\mu < 0$  and  $\partial\alpha_*/\partial\mu > 0$ . Thus, for fixed  $B$ , the first factor on the right-hand side is increasing in  $\mu$ . We need to show that the second factor is increasing in  $z$ . Define  $g(y) = (e^y - 1)/y$  for any  $y > 0$ . We need to show that  $g(S - B + z)/g(z)$  is increasing in  $z$  when  $S > B$ . That is, we need  $\partial \ln(g(S - B + z))/\partial z > \partial \ln(g(z))/\partial z$  when  $S > B$ . This is true if  $Dg(z)/g(z)$  is increasing in  $z$ . Now,

$$\frac{d}{dz} \frac{Dg(z)}{g(z)} = \frac{1}{(e^z - 1)^2} \left( \left( \frac{e^z - 1}{z} \right)^2 - e^z \right).$$

To see that this is positive, note that  $e^{z/2} - e^{-z/2} - z$  is an increasing function for  $z > 0$ , and equal to 0 at  $z = 0$ . Hence  $e^{z/2} > e^{-z/2} + z$  for  $z > 0$ . It follows that  $(e^z - 1)/z > e^{z/2}$  for all  $z > 0$ , and this proves the desired result.

## Appendix B. Proof of Lemma 2

Only the derivative of the mean of  $e^s$  with respect to  $\mu$  remains to be shown. By Lemma 1, the second factor on the right-hand side of (30) is increasing in  $\mu$  when  $S - B = \ln(\Gamma)$  is fixed. Write  $a = \eta/(\sigma^2/2)$ ,  $b = \rho/(\sigma^2/2)$  and recall that  $b > a > 0$ . Also write  $x = \mu/\sigma^2$  and note that  $\alpha > 1$  corresponds to  $a > 1 + 2x$ . Using the definitions of  $\alpha$  and  $B$  one can verify that

$$\frac{\alpha e^B}{\alpha - 1} = \frac{\frac{-x + \sqrt{x^2 + a}}{-(1+x) + \sqrt{x^2 + a}}}{\frac{-x + \sqrt{x^2 + b}}{-(1+x) + \sqrt{x^2 + b}}}.$$

Differentiating with respect to  $x$  gives

$$\frac{\partial}{\partial x} \frac{\alpha e^B}{\alpha - 1} = \frac{-x + \sqrt{x^2 + a}}{-x + \sqrt{x^2 + b}} \frac{b - (1+x)\sqrt{x^2 + b} - (a - (1+x)\sqrt{x^2 + a})}{\sqrt{x^2 + a}\sqrt{x^2 + b}(- (1+x) + \sqrt{x^2 + a})^2}.$$

Since  $b > a$ , the question is thus if  $y - (1+x)\sqrt{x^2 + y}$  is increasing in  $y$  when  $y > 1 + 2x$ . The derivative of this function is

$$\frac{\partial}{\partial y} (y - (1+x)\sqrt{x^2 + y}) = 1 - \frac{1+x}{2\sqrt{x^2 + y}}.$$

This is clearly positive if  $1+x \leq 0$ . The derivative is also positive if  $1+x > 0$  and  $3(x^2 + y) + y > 1 + 2x$ . This is true when  $y > 1 + 2x$  and  $y > 0$ , and hence for  $y \in \{a, b\}$ .

## References

- [1] D. Acemoglu, D. Cao, Innovation by Incumbents and Entrants, unpublished manuscript, MIT Department of Economics, 2010.
- [2] A.A. Alchian, Uncertainty, evolution, and economic theory, *J. Polit. Economy* 58 (1950) 211–221.

- [3] F.E. Alvarez, F.J. Buera, R.E. Lucas, Jr., Models of idea flows, NBER working paper 14135, 2008.
- [4] K.J. Arrow, The economic implications of learning by doing, *Rev. Econ. Stud.* 29 (1962) 155–173.
- [5] A. Atkeson, A.T. Burstein, Innovation, firm dynamics, and international trade, *J. Polit. Economy* 118 (2010) 433–484.
- [6] M. Boldrin, D.K. Levine, Factor saving innovation, *J. Econ. Theory* 105 (2002) 18–41.
- [7] D.R. Cox, H.D. Miller, *The Theory of Stochastic Processes*, Chapman and Hall, London, 1965.
- [8] S.J. Davis, J. Haltiwanger, R. Jarmin, J. Miranda, Volatility and dispersion, in: D. Acemoglu, K. Rogoff, M. Woodford (Eds.), *Business Growth Rates: Publicly Traded Versus Privately Held Firms*, NBER Macroeconomics Annual 2006, MIT Press, Cambridge, MA, 2006.
- [9] A.K. Dixit, R.S. Pindyck, *Investment under Uncertainty*, Princeton University Press, Princeton, NJ, 1994.
- [10] J. Eaton, S. Kortum, International technology diffusion: Theory and measurement, *Int. Econ. Rev.* 40 (1999) 537–570.
- [11] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. II, second edition, Wiley, New York, NY, 1971.
- [12] X. Gabaix, Zipf's law for cities: An explanation, *Quart. J. Econ.* 114 (1999) 739–767.
- [13] X. Gabaix, Power laws in economics and finance, *Annual Review of Economics* 1 (2009) 244–294.
- [14] G.M. Grossman, E. Helpman, *Innovation and Growth in the Global Economy*, MIT Press, Cambridge, MA, 1991.
- [15] J.M. Harrison, *Brownian Motion and Stochastic Flow Systems*, Wiley, New York, NY, 1985.
- [16] C.-T. Hsieh, P.J. Klenow, Misallocation and manufacturing TFP in China and India, *Quart. J. Econ.* 124 (2009) 1403–1448.
- [17] B. Jovanovic, Selection and the evolution of industry, *Econometrica* 50 (1982) 649–670.
- [18] S. Kortum, Research, patenting, and technological change, *Econometrica* 65 (1997) 1389–1419.
- [19] R.E. Lucas, Jr., On the size distribution of business firms, *Bell J. Econ.* 9 (1978) 508–523.
- [20] R.E. Lucas, Jr., Ideas, growth, NBER working paper 14133, 2008.
- [21] E.G.J. Luttmer, Selection, growth, and the size distribution of firms, *Quart. J. Econ.* 122 (2007) 1103–1144.
- [22] E.G.J. Luttmer, Models of growth and firm heterogeneity, *Annual Review of Economics* 2 (2010) 547–576.
- [23] E.G.J. Luttmer, On the mechanics of firm growth, *Rev. Econ. Stud.* (2011) doi:10.1093/restud/rdq028.
- [24] R.N. Nelson, S.G. Winter, *An Evolutionary Theory of Economic Change*, The Belknap Press of Harvard University Press, Cambridge, MA, 1982.
- [25] D. Restuccia, R. Rogerson, Policy distortions and aggregate productivity with heterogeneous plants, *Rev. Econ. Dynam.* 11 (2008) 707–720.
- [26] H.A. Simon, C.P. Bonini, The size distribution of business firms, *Amer. Econ. Rev.* 48 (1958) 607–617.
- [27] J. Steindl, Random processes and the growth of firms, in: *A Study of the Pareto Law*, Hafner Publishing Company, New York, NY, 1965.