Federal Reserve Bank of Minneapolis
Research Department

The Size Distribution of Firms in an Economy with Fixed and Entry Costs

Erzo G.J. Luttmer[*]

Working Paper 633

Revised July 2004

ABSTRACT

This paper describes an analytically tractable model of balanced growth that allows for extensive heterogeneity in the technologies used by firms. Firms enter with fixed characteristics that determine their initial technologies and the levels of fixed costs required to stay in business. Each firm produces a different good, and firms are subject to productivity and demand shocks that are independent across firms and over time. Firms exit when revenues are too low relative to fixed costs. Conditional on fixed firm characteristics, the stationary distribution of firm size satisfies a power law for all sizes above the size at which new firms enter. The tail of the size distribution decays very slowly if the growth rate of the initial productivity of potential entrants is not too far above the growth rate of productivity inside incumbent firms. In one interpretation, this difference in growth rates can be related to learning-by-doing inside firms and spillovers of the information generated as a result. As documented in a companion paper, heterogeneity in fixed firm characteristics together with idiosyncratic firm productivity growth can generate entry, exit, and growth rates, conditional on age and size, in line with what is observed in the data.

1. Introduction

This paper is motivated by evidence on the size distribution of firms in the U.S. presented in Axtell (2001). Using 1997 data on firms from the U.S. Census, Axtell (2001) finds that the log of the right tail probabilities of the log-size distribution, with firm size measured by employment, is on a virtual straight line with a slope of -1.059. Figure 1 below shows the data for 2001.[1] This evidence suggests that the distribution of firm size is well approximated by a Pareto distribution with right tail probabilities of the form $1/S^{\theta}$ with $\theta \approx 1.05$, where S is firm size as measured by employment. The remarkable fit of this distribution for the log-size distribution of firms, at least for firms that are not too small, has been documented and interpreted before by many, and perhaps most notably by Simon and Bonini (1958) and Ijiri and Simon (1964). Researchers as far back as Gribrat (1931) have tried to connect the shape of the observed size distribution to data on firm growth, exit and entry. The mechanism elaborated in this paper is most like the one sketched by Gabaix (1999). Axtell (2001) showed that the older empirical findings hold up well in recent Census data based on a large universe of firms. A useful survey of the literature on firm size and Gibrat's law is given in Sutton (1997). Gabaix (1999) contains extensive discussions of the literature on probability models that give rise to power laws, and their application in economics.

This paper constructs an analytically tractable and fully specified general equilibrium model that is consistent with the evidence just cited, and that is amenable to a more detailed comparison with data, going beyond the size information displayed in Figure 1. This adds to a growing literature that is discussed further below. The economy described incorporates several dimensions of firm heterogeneity that can affect firm size, and the paper characterizes the circumstances in which one can expect an approximate power law to emerge, in spite of factors such as scale economies and economic geography that can influence firm size. In contrast to most of the literature, the model does not rely on Cobb-Douglas technologies to generate balanced growth. A companion paper (Luttmer [2004]) describes the extent to which the model is consistent with data on entry, exit and growth of U.S. Manufacturing plants.

---

[1]The data shown in Figure 1 summarize a population of 5,657,774 U.S. firms in 2001. The largest size category, that of 10,000 employees and over, still contains 930 firms. There is a size category of zero employees (in March of 2001) accounting for 703,837 firms that is not shown. The line is fitted by least squares using all the available size categories of 5 employees and higher. The data are originally from the U.S. Census Bureau, and were obtained from the Small Business Administration internet site, and from the Statistics of U.S. Businesses site of the U.S. Census Bureau (the size categories 5,000-9,999 and 10,000 and over).
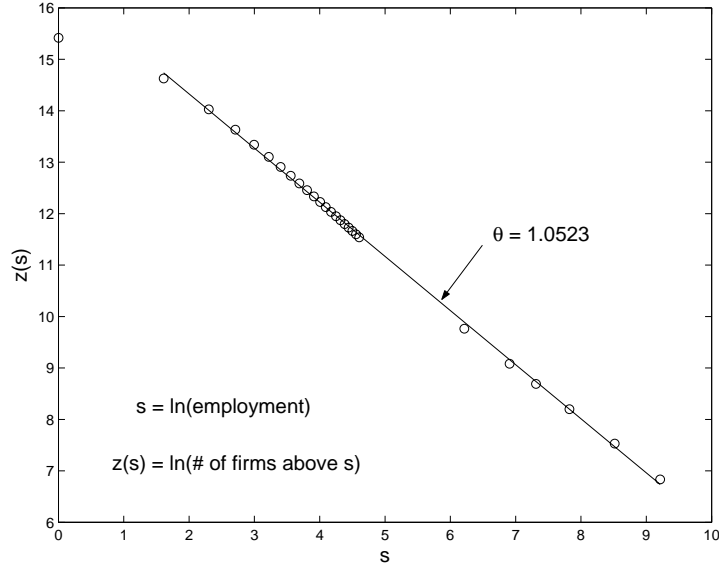
Figure 1: Size Distribution of U.S. Firms in 2001

The economy is one with differentiated final goods and monopolistic competition. A firm is identified with the good of which it is the sole producer. Firms can employ capital and effective labor using a firm-specific technology that exhibits constant returns to scale. Operating a firm requires a fixed cost, and these fixed costs can also vary across firms. Firms experience labor-augmenting improvements in the technology they operate at a rate that is common to all firms. The idiosyncratic total factor productivity of a firm and the weight of its differentiated commodity in the utility function evolve stochastically over time. The stochastic process of these technology and preference shocks is such that firms will eventually close down with probability one. New firms with new technology parameters can enter at a certain cost.

The economy described here has two sources of growth: labor-augmenting technological progress that can be interpreted as accumulation of human capital, and total factor productivity growth in the production of final goods. The aggregate technology for final goods implied by the process of entry and exit of firms is generally not Cobb-Douglas. In a one-sector version of the economy, this would rule out a balanced growth path in which the selection over time of more productive firms leads to total factor productivity growth in the production of final goods. This paper therefore adopts a two-sector specification, as in Uzawa (1961), in which final goods are used only for consumption and a separate technology determines capital accumulation. The production function for new capital depends on labor and existing capital that is not in use for the production of final goods. It exhibits constant returns to scale and must be sufficiently productive.

2

The shares of aggregate output paid to capital, labor and owners of firms are constant along the balanced growth path. The price of capital goods in terms of the composite final consumption good rises at the rate of total factor productivity growth in the final goods producing sector. Wages and firm values rise at the sum of this rate and the rate of labor-augmenting technological progress.

The process of entry and exit leads to a stationary distribution of firm characteristics if the growth rate of total factor productivity inside a typical firm is not too high compared to the rate at which the technology available to new entrants improves over time. This means that incumbents eventually lose out to more recent entrants. In an economy without random productivity shocks, this would lead to an age-size distribution with the property that the most recent entrant is the largest and most productive. The economies described by Grossman and Helpman (1991) and Romer (1990) have this feature. Here instead, random preference and productivity shocks give rise to a selection mechanism, as in Jovanovic (1982), that can make the more efficient incumbents large and profitable relative to average incumbents and new entrants, even though all incumbents will ultimately fall behind and exit. Greenwood and Jovanovic (2001) emphasize this as an important feature of plant-level data.

Conditional on fixed characteristics, the equilibrium distribution of profitability and size is Pareto to the right of the point at which firms enter. This Pareto distribution will have the thick right tail observed in Figure 1 (θ above but close to 1), and survivors of the selection process will be particularly large, if incumbents are not at a great disadvantage relative to new entrants, in terms of the rate at which their technology improves over time. Reducing this disadvantage will cause θ to decrease, and once it reaches 1 the equilibrium described in this paper will cease to exist. Zipf's law —tail probabilities of the form $1/S$— is only a limiting case of the model. As this limit is approached, the number of firms will become arbitrarily small, and aggregate expenditures on fixed and entry costs will converge to zero. In the context of the economy described here, Zipf's law cannot hold if these expenditures are non-negligible.

The relative size of entry and fixed costs determines the range over which the size distribution conditional on fixed characteristics will be Pareto. The lower the entry costs are relative to fixed costs, the lower must be equilibrium profits following entry. The typical entrant will then have a profitability that is close to the level at which it is optimal to exit, and much of the range of the distribution will be Pareto. The overall stationary size distribution will be approximately Pareto if the post-entry selection mechanism dominates the effects of initial heterogeneity in fixed characteristics among entrants.

The key parameter of the model is the difference between the rate at which the technology available to new entrants improves over time and the rate at which the productivity of incumbents is expected to improve. This parameter can be made endogenous by introducing a spillover of technology from surviving incumbents to potential entrants. If this spillover is very strong, then new entrants will constantly drive relatively inefficient incumbents out of business by making use of the technology improvements generated by surviving incumbents. Growth will be fast and the size distribution of firms will have a very thin tail. The data in Figure 1 instead display a very thick tail for which the theoretical size distribution only just has a finite mean. In the context of the model presented here, this means that spillovers from incumbent to entering firms cannot be too strong.

The crucial assumption used in this paper is that the rate of labor augmenting technological progress is the same for all firms. Along the balanced growth path, the rental price of capital in terms of labor decreases at this rate. Not counting fixed costs, any final goods producer with a constant returns technology will therefore choose a ratio of capital over effective labor that is constant over time. Firms with different technologies will choose different ratios, but the model predicts no variation over the life of a firm.

Related Literature   The economy described in this paper can easily be adapted to incorporate decreasing returns of the form introduced by Lucas (1978) to reconcile Gibrat's law with scale economies. This paper follows Atkeson and Kehoe (2002) in requiring firms to incur fixed costs continuously, or be shut down irreversibly. An alternative mechanism for firm exit that may be tractable is along the lines of Hopenhayn (1992). As in Melitz (2003), this paper uses market power instead of decreasing returns. A tractable form of decreasing returns that differs from the span-of-control model of Lucas (1978) is discussed in the concluding remarks. The importance of selection in shaping the distribution of firm characteristics is central in Jovanovic (1982). As in Parente and Prescott (2000) and Lagos (2001), the economy described here has implications for the relation between barriers to entry and the level of total factor productivity.

This paper is perhaps most closely related to the work of Gabaix (1999) on cities, Klette and Kortum (2003) on firms, and by Rossi-Hansberg and Wright (2003, 2004) on cities and firms.

Gabaix (1999) describes how a geometric Brownian motion with a reflecting barrier gives rise to a power law and shows the precise circumstances under which this will lead to Zipf's law. He uses this to construct a model of cities and explain evidence on the

4

size distribution of cities. The entry and exit process described here does not lead to a reflecting barrier, but to a "return process" according to which firms exit below some barrier and enter at a point above this barrier. The two processes are closely related, and the limiting argument used by Gabaix (1999) will be discussed below.

A thought-provoking paper by Eeckhout (2004) suggests strongly that the cross-sectional distribution of city sizes in the US is much better approximated by a log-normal distribution than by the Pareto distribution with $\theta = 1$ that gave rise to Zipf's law. An important difference between the Eeckhout (2004) data and earlier studies is the inclusion of small cities. As a result, the empirical size density of cities eventually comes down for relatively small cities, as does the log-normal density, and in contrast to the Pareto density. By itself this feature of the data does not rule out the model described here as a model for cities. Entry and fixed costs imply an upward sloping size density at low size levels. Nevertheless, the log-normal distribution may still be a better approximation for city size data than the distribution for firm sizes derived in this paper.[2] The populations of cities and firms experience very different levels of entry and exit. Both are very small for cities, and quite substantial for firms. Incumbent cities appear to have more of an edge over entrants than incumbent firms.

The economy described here has many elements in common with Klette and Kortum (2004), who construct a quality ladder model in which firm growth is the result of research and development choices made by firms. The economy grows because firms are able to improve on the quality of existing producers. The spillover described in this paper to endogenize the difference in productivity growth rates between incumbents and entrants has the same effect, provided entrants can start with a technology that is not too far behind the average in the population of incumbents. Entrants in Klette and Kortum (2004) are small relative to the average firm, firm growth satisfies Gibrat's law, and firms are eventually driven out of business with probability one. The resulting size distribution, where size is measured by the number of goods produced by the firm, is logarithmic. This distribution is highly skewed, with a monotonically decreasing density. But a plot as in Figure 1 generates a curve that is concave and that does not asymptote to a straight line for large firm sizes. The right tail of the distribution is too thin. Research and development generates new goods in Kortum and Klette (2004) according to Poisson

---

[2]It would be interesting to see the evolution of the cross-sectional variance over time. Also, given the large data set, the fit of the log-normal distribution could be examined in more detail by plotting the log of the number of cities in the right tail of the empirical distibution against the squared log of city size.

processes, whereas here productivity and taste shocks follow Brownian motions. In part, the different properties of these stochastic processes are at the origin of the different size distributions generated by the two economies.

Rossi-Hansberg and Wright (2004) describe an economy with many sectors and many firms in each sector. Firms operate decreasing return to scale Cobb-Douglas technologies with factor shares that can vary across sectors. Capital is sector specific, and the number and size of firms in a particular sector at a point in time is determined by a free-entry condition.[3] Mean reversion in sectoral technology shocks is used to generate a stationary equilibrium. As in the current paper, Zipf's law only holds as a limiting case. In Rossi-Hansberg (2004), however, moving away from this limit generates concavity in the relation between log firm size and log tail probabilities, at both ends of the distribution. This seems to be at odds with Figure 1.

The analytical techniques used in this paper are standard. Useful references are Dixit and Pindyck (1994) for stopping problems and Karlin and Taylor (1981) and Harrison (1985) for hitting times and stationary distributions of Markov diffusions.

Outline of the Paper Section 2 describes the economy and determines the choices made by individual consumers and firms along a balanced growth path. Section 3 derives the stationary distribution of firm characteristics. This section also contains a discussion of the size distribution of firms and of the precise ways in which selection affects the distribution of firm characteristics. The balanced growth path is determined in Section 4 and the effects of market size and entry and fixed costs are discussed. Section 5 offers two interpretations of the assumption that incumbents cannot grow as fast as new entrants. Section 6 contains concluding remarks and discusses ongoing and possible future research.

2. An Economy with Heterogeneous Firm Technologies

The economy has a final goods sector and a sector in which capital is accumulated. Final goods are produced with capital and labor using firm-specific constant returns to scale technologies. These final goods are perishable and can be used for consumption only.

---

[3]This feature somewhat complicates comparing their model with data. All firms in an industry are identical. Also, it does not matter which of the infinitesimal firms in an industry exit when exit from a particular industry is required. As a result, the model has no determinate implications for the joint age-size distribution of firms.

Capital is produced using capital and labor, subject to constant returns to scale.

2.1 Consumers

Time is continuous and indexed by $t \in [0, \infty)$. There is a continuum $He^{\eta t}$ of identical infinitely-lived consumers alive at time t. The population growth rate η is taken to be non-negative. Each consumer is endowed with one unit of labor at every point in time. Labor supply is inelastic. A typical consumer i has preferences over sequences $\{C_{i,t}\}_{t \geq 0}$ of a composite good given by:

$$\mathrm{E}\left[\int_0^\infty e^{-\rho t} U(C_{i,t}) \mathrm{d}t\right]$$

The subjective discount rate ρ is not too small and the period utility function is:

$$U(C) = \frac{C^{1-\alpha} - 1}{1 - \alpha}$$

for some positive α. A continuum of goods of different types make up the composite good. Preferences are additively separable with weights that define the type of a good. The additive separability implies that all goods of the same type and trading at the same price will be consumed at the same rate. Let $c_{i,t}[u, p]$ be consumer i's consumption at time t of a good of type u that trades at a price p. In equilibrium, there will be a measure G_t of goods that are available at time t, defined on the set of types and prices. The composite good is a version of the one specified in Dixit and Stiglitz (1977). Specifically, for some $\omega \in (0, 1)$:

$$C_{i,t} = \left[\int u^{1-\omega} c_{i,t}^{\omega}[u, p] \mathrm{d}G_t(u, p)\right]^{1/\omega}$$

The number of goods is measured by:

$$N_t = \int \mathrm{d}G_t(u, p)$$

The type u of a good can be viewed as measure of quality. Consumer i chooses $c_{i,t}[u, p]$ to minimize the cost of acquiring $C_{i,t}$. Define the following period-t price index:

$$P_t = \left[\int u p^{-\frac{\omega}{1-\omega}} \mathrm{d}G_t(u, p)\right]^{-\frac{1-\omega}{\omega}}$$

Then cost minimization implies:

$$pc_{i,t}[u, p] = P_t \left(u C_{i,t}\right)^{1-\omega} c_{i,t}^{\omega}[u, p] \tag{1}$$

The price elasticity of demand is $-1/(1-\omega)$, and the implied expenditure share for good (u, p) is $u(p/P_t)^{-\omega/(1-\omega)}$.

Consumers own firms and physical capital, and they can trade in a sequence of complete markets, subject to a present-value borrowing constraint. Since preferences are homothetic and markets are complete, we can assume there is a representative consumer with the same preferences as described here for individual consumers. Aggregate consumption of the composite good will be denoted by C_t and aggregate consumption of a type-(u, p) commodity is $c[u, p]$.

2.2 Capital Accumulation and Balanced Growth

There is a single physical capital good, the stock of which at time t is denoted by K_t. Capital can be used to produce the differentiated final goods, and to reproduce capital. New capital and depreciated capital are perfect substitutes in production. Capital depreciates at a rate δ. The technology for augmenting the capital stock is:

$$\mathrm{d}K_t = [F(K_{I,t}, \Lambda_t L_{I,t}) - \delta K_t]\,\mathrm{d}t \tag{2}$$

$$\mathrm{d}\Lambda_t = \lambda \Lambda_t \mathrm{d}t \tag{3}$$

where $K_{I,t}$ is the amount of capital and $L_{I,t}$ is the amount of labor used in investment. The production function F is strictly increasing, concave, and exhibits constant returns. It is assumed to be sufficiently smooth. The process Λ_t represents labor-augmenting technological progress. This form of technological progress is taken to be exogenous in this paper. Observe that the accumulation of physical and "human capital" Λ_t in (2)-(3) is linear in (K_t, Λ_t) when $K_{I,t}/K_t$ and $L_{I,t}$ are constant. It is possible to make labor-augmenting technological progress endogenous in a way that is consistent with balanced growth, by letting consumers divide their time between supplying labor and accumulating human capital, along the lines of Lucas (1988) and Rebelo (1991).

The price of capital in terms of the composite commodity is denoted by q_t. Capital can be rented by the producers of new capital and by the various final goods producers. The rental price is v_t and labor trades at a wage w_t. Both are expressed in units of the composite commodity per unit of time. Investors in capital goods maximize profits:

$$\max_{K,L} q_t F(K, \Lambda_t L) - v_t K - w_t L$$

Since F exhibits constant returns, these profits will be zero and factor prices and the inputs $K_{I,t}$ and $L_{I,t}$ must satisfy:

$$\mathrm{D}_K F\left(1, \frac{\Lambda_t L_{I,t}}{K_{I,t}}\right) = \frac{v_t}{q_t} \tag{4}$$

8

$$D_L F\left(1, \frac{\Lambda_t L_{I,t}}{K_{I,t}}\right) = \frac{w_t}{q_t \Lambda_t} \tag{5}$$

On a balanced growth path, the fraction of the aggregate labor supply assigned to investment is constant. The aggregate labor supply grows at a rate η, and so $L_{I,t} = L_I e^{\eta t}$ for some L_I. Balanced growth requires a constant ratio of capital over effective labor used in investment. Therefore $K_{I,t} = K_I e^{(\lambda+\eta)t}$ for some K_I. It follows from (4)-(5) that the rental price of capital will grow at the same rate as the price of capital, and that wages will grow at the rate at which the price of capital grows, plus the growth rate of labor-augmenting technical progress.

The price of capital in terms of the composite final good depends on its marginal product in the final goods producing sector. This marginal product will turn out to grow at a rate κ that is determined in the next section. As a result, $q_t = q e^{\kappa t}$, $v_t = v e^{\kappa t}$ and $w_t = w e^{(\kappa+\lambda)t}$. On the balanced growth path, per capita consumption of the composite commodity will grow at a rate $\kappa + \lambda$. The real interest rate in units of the composite commodity is therefore:

$$r = \rho + \alpha(\kappa + \lambda) \tag{6}$$

The payoffs to owning capital consist of increases in its relative price minus physical depreciation, plus the revenues from renting out capital. Since there is no aggregate risk, these payoffs must be equal to rq_t per unit of time. The rental price of capital is therefore determined by:

$$\frac{v}{q} = r + \delta - \kappa \tag{7}$$

Equations (6) and (7) determine r and v/q in terms of exogenous parameters. Together with (4) this determines the ratio $\Lambda L_I / K_I$, and (5) then determines w/q.

The total capital stock is given by $K_t = K e^{(\lambda+\eta)t}$ along the balanced growth path. The evolution of the capital stock described in (2) therefore implies that:

$$F(K_I, \Lambda L_I) = (\lambda + \eta + \delta)K \tag{8}$$

along the balanced growth path. This pins down the ratio K_I/K of capital used in the investment sector over the total capital stock. The following assumption ensures that the technology for reproducing capital is sufficiently productive to make balanced growth possible.

Assumption 1: *Preference parameters and growth rates satisfy:*

$$\rho + \alpha(\kappa + \lambda) > \kappa + \lambda$$

9

The production function for capital accumulation is such that:

$$D_K F(k, 1) = \rho + \alpha(\kappa + \lambda) + \delta - \kappa$$

implies:

$$F(k, 1) > (\lambda + \eta + \delta)k$$

The first part of this assumption is needed to ensure that the labor income of consumers is finite. The second part implies that $F(1, \Lambda L_I/K_I)/(\lambda + \eta + \delta)$ is greater than one when evaluated at the solution to (4) and (6)-(7). This then guarantees that the amount of capital required to make the capital stock grow at the required rate is less than the aggregate amount of capital available.

2.3   Final Goods Production

A typical firm can produce a unique final good using capital, variable labor, and a certain amount of overhead labor. The firm shuts down permanently if the required overhead labor is not employed. A firm may also be shut down as a result of exogenous shocks that occur with a probability ϵ per unit of time. This exogenous exit probability may be zero, but allowing for it to be positive is important empirically to account for exit that is not perfectly related to profitability.

At age a, a firm that entered at time t with fixed characteristics x uses capital $K_{t,a}$ and production labor $L_{t,a}$ to produce a quantity $y_{t,a}$ of the final good of type $u_{t,a}$ given by:

$$y_{t,a} = z_{t,a} G_{[x]}(K_{t,a}, \Lambda_{t,a} L_{t,a}) \tag{9}$$

The overhead labor requirement of the firm is denoted by $M[x]$.[4] Given a price $p_{t,a}$, revenues of the firm are given by $R_{t,a} = p_{t,a} y_{t,a}/P_t$ in units of the composite commodity. The demand function for type-$u_{t,a}$ goods then implies that revenues of the firm are:

$$R_{t,a} = C_{t+a}^{1-\omega} \left[Z_{t,a} G_{[x]}(K_{t,a}, \Lambda_{t,a} L_{t,a}) \right]^{\omega} \tag{10}$$

where the state variable $Z_{t,a}$ combines the state of preferences and technology as follows:

$$Z_{t,a} = \left(u_{t,a}^{1-\omega} z_{t,a}^{\omega} \right)^{\frac{1}{\omega}}$$

---

[4]An easy extension is to introduce decreasing returns by raising $G_{[x]}(K_{t,a}, \Lambda_{t,a} L_{t,a})$ to a power between 0 and 1. This will generate a version of the managerial span-of-control model of Lucas (1978). The main effect is to replace ω by a fraction of ω.

A firm's revenues vary with changes in the type $u_{t,a}$ of its output and its productivity level $z_{t,a}$. Note that $Z_{t,a} = z_{t,a}$ if $u_{t,a}$ is constant and normalized to 1. In much of what follows, this paper ignores prices and quantities of the individual final goods, and focuses on the properties of equilibrium revenues and factor inputs. For this, it is not necessary to disentangle preference and technology shocks. With some abuse of terminology, $Z_{t,a}$ will be referred to as firm total factor productivity or simply firm productivity.

The initial values of the other technology parameters of a firm entering at time t are determined by:

$$(\Lambda_{t,0}, Z_{t,0}) = \left(e^{\lambda t}, e^{\zeta t} Z[x]\right)$$

The vector x of fixed characteristics is drawn from an exogenously given time-invariant distribution with density g. As the firm ages, $\Lambda_{t,a}$ and $Z_{t,a}$ evolve according to:

$$\mathrm{d}\begin{bmatrix} \ln(\Lambda_{t,a}) \\ \ln(Z_{t,a}) \end{bmatrix} = \begin{bmatrix} \lambda \\ \zeta_* \end{bmatrix} \mathrm{d}a + \begin{bmatrix} 0 \\ \sigma_* \end{bmatrix} \mathrm{d}W_{t,a}$$

where $W_{t,a}$ is a standard Brownian motion. The standard deviation σ_* is taken to be positive. A final goods producer enters with an idiosyncratic level of labor input requirements that depend on its vintage. After entry, the efficiency units of labor in any given final goods firm continue to grow at the same rate as they do in any other final goods firm, and in the investment sector. Instead, the combined preference and technology state variable evolves according to a geometric random walk, with a drift ζ_* that will be required to be not too large relative to ζ. The Brownian motions $W_{t,a}$ are assumed to be independent across firms.[5]

Measured in units of the composite good, the value $V_t[x]$ of a firm entering at time t with fixed characteristics x is given by:

$$V_t[x] = \max_{K,L,Y,\tau} \mathrm{E}_t \left[\int_0^\tau e^{-(r+\epsilon)a} \left(R_{t,a} - v_{t+a} K_{t,a} - w_{t+a} \left(L_{t,a} + M[x]\right)\right) \mathrm{d}a \right] \qquad (11)$$

The maximization is subject to (10) and the restriction that choice variables only depend on the available information.

The aggregate labor supply is growing at a rate η, and every firm requires a minimum fixed amount of labor to operate. The number of firms can therefore not grow exponentially at any rate above η. Along the balanced growth path, the number of firms will grow exactly at the rate η, so that $N_t = Ne^{\eta t}$ for some N. Observe that firm revenues

---

[5]Much of what follows can be generalized by allowing ζ_* and σ_* to be functions of the underlying fixed characteristics x. This may be an empirically important generalization.

can be written in terms of the amount of aggregate consumption per firm C_{t+a}/N_{t+a} and $N_{t+a}^{(1-\omega)}Z_{t,a}^{\omega}$. Growth in the number of firms acts like an improvement in productivity. In the equilibrium constructed below, aggregate productivity in the final goods sector itself grows at the rate ζ. As a result, the effective growth rate of productivity in this sector will be:

$$\kappa = \frac{1}{\omega}\left[\omega\zeta + (1-\omega)\eta\right] \tag{12}$$

along the balanced growth path. This growth rate exceeds ζ if utility is increasing in the number of goods and the population growth rate is positive. Aggregate consumption at time t will be $C_t = Ce^{(\kappa+\lambda+\eta)t}$ along the balanced growth path.

2.3.1 Variable Input Choices

Capital and variable labor inputs are chosen to maximize variable revenues at a point in time. For a firm with fixed characteristics x, the resulting ratio of capital over effective labor is a constant:

$$\frac{K_{t,a}}{\Lambda_{t,a}L_{t,a}} = k[x]$$

where $k[x]$ solves:

$$\frac{w}{v} = \frac{D_L G_{[x]}\left(k[x],1\right)}{D_K G_{[x]}\left(k[x],1\right)} \tag{13}$$

This is a direct consequence of the fact that the rate of labor-augmenting technological progress of the firm is equal to the growth rate λ of the ratio of rental prices w_t/v_t. The ratio w_t/v_t is already determined by balanced growth considerations in the capital accumulation sector, and so (13) pins down the capital-labor ratios for all firms. The optimal variable revenues and factor expenditures are given by:

$$R_{t,a} = C_{t+a}\left[\left(\frac{Z_{t,a}\Lambda_{t,a}}{w_{t+a}}\right)\omega D_L G_{[x]}\left(k[x],1\right)\right]^{\omega/(1-\omega)} \tag{14}$$

$$\begin{bmatrix} w_{t+a}L_{t,a} \\ v_{t+a}K_{t,a} \end{bmatrix} = \frac{\omega R_{t,a}}{G_{[x]}\left(k[x],1\right)}\begin{bmatrix} D_L G_{[x]}\left(k[x],1\right) \\ k[x]D_K G_{[x]}\left(k[x],1\right) \end{bmatrix} \tag{15}$$

Note that the ratios of factor payments over revenues are constant. Variable profits are a fraction $1-\omega$ of revenues. Using the link (12) between productivity growth, the growth rate of the number of firms, and the growth rate of wages, one can simplify the expression for firm revenues to:

$$R_{t,a} = \left(\frac{C_{t+a}}{N_{t+a}}\right)N\left[\frac{e^{-\zeta a}Z_{t,a}}{Z_{t,0}}\left(\frac{\omega Z[x]}{w}\right)D_L G_{[x]}\left(k[x],1\right)\right]^{\omega/(1-\omega)}$$

Aggregate consumption per firm grows at the same rate as wages. The evolution over time of firm revenues and factor payments relative to wages is therefore determined by the variable $e^{-\zeta a}Z_{t,a}/Z_{t,0}$. This is the key variable determining the dynamics of firm size. Observe that it only depends on a and not on t. The cross-sectional distribution of this variable will be constant along the balanced growth path. As a result, the population average of firm revenues $R_{t,a}$ will have the same trend as aggregate consumption per firm.

To determine the value of a firm, it will be convenient to write firm profits as:

$$R_{t,a} - (v_{t+a}K_{t,a} + w_{t+a}L_{t,a} + w_{t+a}M[x]) = w_{t+a}M[x]\left(e^{s_a} - 1\right)$$

where s_a is defined by:

$$s_a = S[x] + \frac{\omega}{1-\omega}\left[\ln\left(\frac{Z_{t,a}}{Z_{t,0}}\right) - \zeta a\right] \tag{16}$$

and where $S[x]$ is given by:

$$e^{S[x]} = \frac{(1-\omega)C}{wM[x]}\left[\left(\frac{\omega Z[x]}{w}\right)D_L G_{[x]}\left(k[x],1\right)\right]^{\omega/(1-\omega)} \tag{17}$$

The variable s_a measures profitability, and the initial profitability of a new firm with fixed characteristics x is measured by $S[x]$. If $s_a = 0$, then variable revenues just cover overhead, and flow profits are zero. Define the value function $V(\cdot)$ as:

$$V(s) = \max_{\tau} E\left[\int_0^{\tau} e^{-(r+\epsilon-[\kappa+\lambda])a}\left(e^{s_a} - 1\right)da \,\middle|\, s_0 = s\right]$$

for any s. The evolution of the profitability of a firm set up at time t is described by $ds_a = \mu da + \sigma dW_{t,a}$, where, by (16):

$$\begin{bmatrix} \mu \\ \sigma \end{bmatrix} = \frac{\omega}{1-\omega}\begin{bmatrix} \zeta_* - \zeta \\ \sigma_* \end{bmatrix} \tag{18}$$

The profitability of a firm, as measured by s_a, has a negative drift when productivity inside the firm is expected to grow more slowly than the productivity of new entrants. Note that the differences in these growth rates and the variance of productivity shocks are magnified significantly when the differentiated goods are close substitutes. The value of a firm entering at time t with fixed characteristics x can now be written as:

$$V_t[x] = w_t M[x]V(S[x])$$

$V(s)$ is the value of a firm relative to its overhead expenses when the current state of profitability is s. The value function $V(\cdot)$ only depends on the interest rate and fixed parameters of the economy.

2.3.2 Exit Decisions

The presence of a fixed cost implies a minimum size. Firms with very low productivity will choose to exit. The value of a firm must be finite in any equilibrium. The following assumption makes sure that this is the case.

Assumption 2: *The productivity and exogenous exit parameters satisfy:*

$$\rho + \alpha(\kappa + \lambda) + \epsilon > \kappa + \lambda + \mu + \frac{1}{2}\sigma^2$$

where κ is given by (12) and μ and σ are defined in terms of ζ, ζ_* and σ_* by (18).

Assumption 1 and the fact that ϵ is non-negative already imply that $\rho + \alpha(\kappa + \lambda) + \epsilon > \kappa + \lambda$, which ensures that the fixed costs incurred by the firm have a finite present value even if the firm is never shut down for other than exogenous reasons. Assumption 2 implies that the present value of profits before fixed costs is also finite for such a shut-down policy.

The value function $V(s)$ must satisfy the following Bellman equation in the range of s where a firm is not shut down:

$$rV(s) = (\kappa + \lambda - \epsilon)V(s) + \mathcal{A}V(s) + e^s - 1 \tag{19}$$

where:

$$\mathcal{A}V(s) = \mu \mathrm{D}V(s) + \frac{1}{2}\sigma^2 \mathrm{D}^2 V(s)$$

The return to owning a firm consists of a capital gain $(\kappa + \lambda - \epsilon)V(s) + \mathcal{A}V(s)/V(s)$ and a dividend yield $(e^s - 1)/V(s)$. It is optimal to shut down a firm when its profitability s falls below some threshold b. Given that the firm is shut down at b, it must be that the value of a firm is zero at that point. This implies the boundary condition $V(b) = 0$. The optimal threshold must be such that V is differentiable at b, and so $\mathrm{D}V(b) = 0$. A further boundary condition follows from the fact that the value function cannot exceed the value of a firm that operates without fixed costs. This implies that $V(s) \leq e^s/(r + \epsilon - [\kappa + \lambda + \mu + \sigma^2/2])$.

With these boundary conditions, the Bellman equation (19) has only one solution. For $\sigma > 0$, the solution is by $V(s) = 0$ for $s \leq b$, and by:

$$V(s) = \frac{1}{r + \epsilon - [\kappa + \lambda]} \left[-1 + \left(\frac{\xi}{1 + \xi}\right) e^{s-b} + \left(\frac{1}{1 + \xi}\right) e^{-\xi(s-b)} \right] \tag{20}$$

for all $s \geq b$. The coefficient ξ is given by:

$$\xi = \frac{1}{\sigma^2} \left[\mu + \sqrt{\mu^2 + 2\sigma^2 \left(r + \epsilon - [\kappa + \lambda]\right)} \right]$$

and the exit barrier b satisfies:

$$\frac{e^b}{r + \epsilon - \left(\kappa + \lambda + \mu + \frac{1}{2}\sigma^2\right)} = \left(\frac{\xi}{1+\xi}\right)\frac{1}{r + \epsilon - [\kappa + \lambda]} \tag{21}$$

The first two terms in (20) represent the present value of $\{e^{s_a} - 1\}_{a \geq 0}$ given $s_0 = s$. The third term represents the added value that comes from the possibility of shutting down the firm when variable revenues are low. Note that $r + \epsilon - [\kappa + \lambda]$ implies $\xi > 0$. The third term in (20) therefore converges to zero as s_a gets large. It follows that $V(s)$ is strictly increasing on (b, ∞), with an asymptote equal to the present value of $\{e^{s_a} - 1\}_{a \geq 0}$. The fact that ξ is positive also implies that the factor multiplying the present value $(r + \epsilon - [\kappa + \lambda])^{-1}$ in (21) is positive and smaller than one. The left-hand side of (21) is the present value of $\{e^{s_a}\}_{a \geq 0}$ conditional on $s_0 = b$. Thus production will be stopped when s_a reaches the point where the present value of continuing production forever falls below $-(r + \epsilon - [\kappa + \lambda])^{-1}(1 + \xi)^{-1}$.

The construction of $V(s)$ in (20)-(21) holds for $\sigma > 0$. The case of $\sigma = 0$ can be computed directly or by taking limits in the above expressions.

2.3.3 Entry Decisions

Setting up a firm requires a blueprint. A blueprint with fixed characteristic x is defined by the production function $G_{[x]}(\cdot)$ and the fixed cost and productivity parameters $(M[x], Z[x])$. New blueprints can be obtained by drawing a characteristic x from a distribution with density $g(x)$. The rate at which such draws can be obtained is equal to $1/\Gamma$ times the amount of labor assigned to the development of blueprints. Blueprints can only be used at the time they are developed. Along the balanced growth path to be constructed, the number of firms grows at the same rate as the population and firms exit at a constant rate. Thus new firms must be set up at a constant rate as well. A positive amount of labor must therefore be assigned to the development of new blueprints. Since anyone can enter the business of developing new firms, profits must be zero. This implies:

$$\Gamma = \int M[x]V(S[x])g(x)\mathrm{d}x \tag{22}$$

This zero-profit condition depends implicitly on the interest rate r, via $V(\cdot)$, and on steady-state rental prices and aggregate consumption, via $S[\cdot]$. The following assumption is used to make sure that the expected value of entry is finite.

Assumption 3: *The entry density g of fixed characteristics is such that:*

$$\int M[x]g(x)\mathrm{d}x \quad < \quad \infty$$

$$\int \left[Z[x]\mathrm{D}_L G_{[x]}\left(k[x],1\right)\right]^{\omega/(1-\omega)} g(x)\mathrm{d}x \quad < \quad \infty$$

The expected fixed cost and productivity of a potential entrant cannot be too high. The fact that this suffices to ensure that the right-hand side of (22) is finite follows from (17) and (20).

If Assumption 3 holds, then the zero-profit condition (22) implies a unique equilibrium value for $C/w^{1/(1-\omega)}$. In turn this pins down the equilibrium value of the function $S[\cdot]$. To see this, recall from (17) that $e^{S[x]-b}$ is equal to the product of $C/w^{1/(1-\omega)}$ and a function that only depends on fixed characteristics x and the capital-labor ratios $k[x]$. These capital-labor ratios only depend on fixed characteristics and the ratio of rental prices v/w. This ratio was determined by the balanced growth requirements (4)-(6). It then follows from (20) that (22) uniquely determines the equilibrium value of $C/w^{1/(1-\omega)}$. It is not difficult to see that this ratio, and thus $e^{S[x]-b}$ for all x, is increasing in Γ. In equilibrium, the more difficult it is to enter, the higher must be the initial profitability of all entering firms.

3. The Stationary Distribution of Firm Characteristics

Every final goods firm needs some amount of overhead labor to produce, and the aggregate labor supply grows at a non-negative rate η. Along the balance growth path to be constructed below, the number of firms also grows at the rate η. The following assumption will be imposed to guarantee a stationary distribution of firm characteristics with a mean profitability that is finite.

Assumption 4: *The productivity parameters satisfy:*

$$\epsilon + \eta > \mu + \frac{1}{2}\sigma^2$$

where μ and σ are defined in (18).

Note that $\mu + \sigma^2/2$ is the drift of the profitability measure e^{s_a}. Thus Assumption 4 requires that the profitability of a typical incumbent firm is not expected to grow faster than the sum of the population growth rate and the exogenous exit rate. The growth

rate of profitability among surviving firms will be greater than $\mu + \sigma^2/2$. If $\epsilon + \eta$ is zero then μ must be negative, but it can be positive otherwise. Section 5 offers two examples of environments in which μ is negative.

3.1 Fixed Characteristics, Age, and Profitability

The state of a firm is defined by its fixed characteristics x and the state variable s. Although the age of the firm is irrelevant for its operations, knowing the joint distribution of age and other firm characteristics is useful when comparing the model with data.[6] The following therefore describes the joint distribution of age a and characteristics (x, s).

Along the balanced growth path, there will be a measure of firms growing at a rate η, and defined on the set of possible ages, fixed characteristics, and profitability levels. The density of this measure at date t can be written as $Ne^{\eta t}p(a, x, s)$, where p is a probability density. This density must satisfy the following version of the Kolmogorov forward equation:

$$\mathrm{D}_a p(a, x, s) = -(\epsilon + \eta)p(a, x, s) - \mu \mathrm{D}_s p(a, x, s) + \frac{1}{2}\sigma^2 \mathrm{D}_{ss} p(a, x, s) \qquad (23)$$

for all $a > 0$ and $s > b$. The first term on the right-hand side of (23) reflects the exogenous exit of firms and the fact that the measure of firms grows over time. The remaining two terms describe how the density $p(a, x, s)$ evolves as a result of changes in the profitability of individual firms.

The boundary value $p(0, x, s)$ is determined by the rate at which new firms enter. Let $Ee^{\eta t}$ be the number of blueprints developed at time t per unit of time. That is, $Ee^{\eta t}$ represents the number of draws obtained from the density g per unit of time. Of these "entry attempts" only those for which $S[x] > b$ lead to entry. This implies that at the $a = 0$ boundary:

$$p(0, x, s) = \begin{cases} g(x)E/N & s = S[x] > b \\ 0 & \text{otherwise} \end{cases} \qquad (24)$$

A further boundary condition is given by the requirement that $p(a, x, b) = 0$ for all $a > 0$. This condition arises from the fact that firms exit at b while none enter starting from profitability levels below b. To ensure that p is a proper density, it must also be the case that $p(a, x, s)$ goes to zero for large (a, x, s).

---

[6]For example, the U.S. Census of Manufactures is taken once every five years, generating five-year age cohorts. The stationary distribution of age and firm characteristics can be used to infer the composition of such cohorts and this can then be used to compute subsequent cohort growth and exit rates implied by the model. See Luttmer (2004). The fact that the distribution is known analytically greatly fascilitates these computations.

The partial differential equation (23)-(24) can be viewed as a collection of independent partial differential equations in (a, s), one for each x. This is a consequence of the assumption that x is constant over time, and the result that exit decisions only depend on the profitability level s. The solution of (23)-(24) is given by:

$$p(a, x, s) = e^{-(\epsilon+\eta)a}\psi(a, s|S[x])g(x)E/N \tag{25}$$

for all $a > 0$, all x, and all $s > b$, where:[7]

$$\psi(a, s|s_0) = \frac{1}{\sigma\sqrt{a}}\left[\phi\left(\frac{s - s_0 - \mu a}{\sigma\sqrt{a}}\right) - e^{-\frac{2\mu}{\sigma^2}(s_0-b)}\phi\left(\frac{s + s_0 - 2b - \mu a}{\sigma\sqrt{a}}\right)\right]$$

By differentiating, one can check that the two terms that define $e^{-(\epsilon+\eta)a}\psi(a, s|s_0)$ both satisfy (23). For small values of a, the first term approximates a normal probability density that puts almost all probability close to $s = s_0$. The second term converges to zero as a goes to zero, since $s + s_0 > 2b$. As a result, $p(a, x, s)$ converges to the value required by the boundary condition (24), as a goes to zero. The fact that $\psi(a, s|b) = 0$ ensures that the boundary condition $p(a, x, b) = 0$ holds.

3.2 Profitability Conditional on Fixed Characteristics– The Power Law

To complete the description of the equilibrium and to analyze the properties of the size distribution of firms, it will be convenient to have expressions for the density of age and profitability, as well as profitability, conditional on fixed characteristics.

Define the parameters θ and θ_* as follows:

$$\theta = \frac{1}{\sigma^2}\left(-\mu + \sqrt{\mu^2 + 2\sigma^2[\epsilon + \eta]}\right), \theta_* = \frac{1}{\sigma^2}\left(\mu + \sqrt{\mu^2 + 2\sigma^2[\epsilon + \eta]}\right)$$

Assumption 4 implies that $\theta > 1$ and the fact that $\epsilon + \eta$ is non-negative implies that θ_* is non-negative as well. If $\epsilon + \eta = 0$, then θ is simply equal to $-2\mu/\sigma^2$ and $\theta_* = 0$. By integrating $e^{-(\epsilon+\eta)a}\psi(a, s|s_0)$ one can verify that the density of age and profitability among firms with the same fixed characteristics and an initial profitability $s_0 = S[x]$ is:

$$\pi(a, s|s_0) = \left(\frac{1 - e^{-\theta_*(s_0-b)}}{\epsilon + \eta}\right)^{-1} e^{-(\epsilon+\eta)a}\psi(a, s|s_0) \tag{26}$$

---

[7]The standard normal density and distribution functions are denoted by ϕ and Φ, respectively.

for all $s \geq b$. The implied density of profitability given an initial profitability $s_0 > b$ is:[8]

$$\pi(s|s_0) = \frac{\theta}{\theta + \theta_*} \frac{\min\left\{e^{[\theta+\theta_*](s-b)}, e^{[\theta+\theta_*](s_0-b)}\right\} - 1}{e^{\theta(s-b)}\left[\frac{e^{\theta_*(s_0-b)}-1}{\theta_*}\right]} \tag{27}$$

for all $s \geq b$. An example is shown in Figure 2. The kink in this density at $s = s_0$ is a result of the entry that takes place at s_0. Conditional on $s \geq s_0$, the density of e^s implied by (27) is a Pareto density with tail probabilities of the form $e^{-\theta s}$. This is the power law found in firm size data. The parameter θ will be referred to as the "tail index" of the profitability distribution. In the case of Axtell (2001) and Figure 1 above, θ is just greater than 1, as required by Assumption 4.
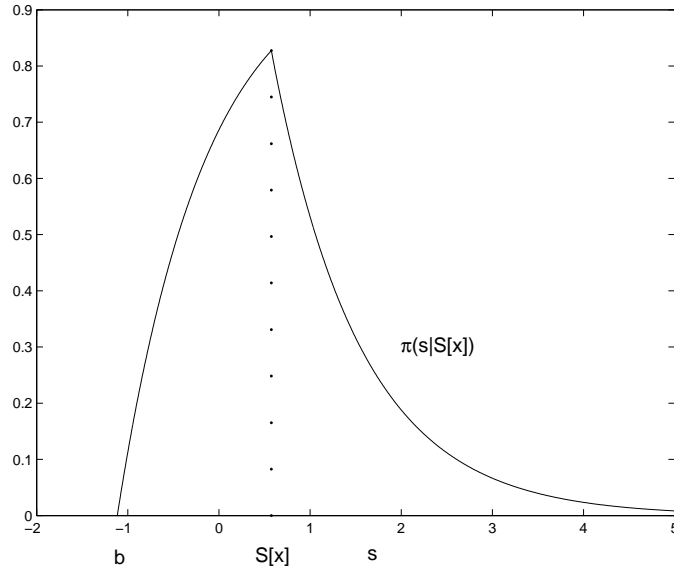

Figure 2: Density of Size Conditional on Fixed Characteristics

The size of a firm measured by revenues or variable labor inputs is proportional to a factor depending on fixed characteristics, and profitability. The model described here can therefore account for a power law in firm size data if profitability is the main determinant of size. Unlike the Pareto density, and like the log-normal density often considered as an alternative, the density (27) is upward sloping for low size levels. Unlike the log-normal, the support of (27) is bounded below, as a result of the fixed costs that cause low-productivity firms to exit.

---

[8]Alternatively, note that by integrating out a from (23)-(24) one obtains an inhomogeneous second-order differential equation in s with a characteristic equation $\epsilon + \eta + \mu z - \sigma^2 z^2/2 = 0$. The roots of this equation are θ and θ_*. The differential equation is solved by (27).

The mean of e^{s-b} given s_0 is:

$$\int_b^\infty e^{s-b}\pi(s|s_0)\mathrm{d}s = \frac{\theta}{\theta-1}\left[\frac{\frac{1-e^{-(1+\theta_*)(s_0-b)}}{1+\theta_*}}{\frac{1-e^{-\theta_*(s_0-b)}}{\theta_*}}\right]e^{s_0-b} \tag{28}$$

Given that $\epsilon + \eta$ is non-negative, Assumption 4 is exactly equivalent to $\theta > 1$. Thus an approximate version of Zipf's law can hold in this economy, but it cannot hold exactly. The stationary density (27) would be well defined even for $\theta > 0$, but without $\theta > 1$ the mean of e^{s-b} would not be finite and aggregate quantities would not be well defined. The average profitability in a large cross-section of firms would not converge as the number of firms grows without bound. The abstraction of an economy with balanced growth and a continuum of infinitesimal firms is not well defined in that case.[9] [10]

The right-hand side of (28) is greater than $\theta/(\theta-1)$. Thus the ratio of the profitability of the average firm relative to an exiting firm is bounded below by $\theta/(\theta-1)$. When the productivity growth rate of existing firms is close to that of new entrants, θ will be close to 1 and the lower bound $\theta/(\theta-1)$ will be a very large number. In such an environment, even the average firm is much larger and more profitable than the firms that exit. If entry is easy, then $S[x] - b$ will be small for most x. In that case, entering firms will also be small relative to the average firm, and the power law will apply over much of the range of firm sizes.

3.3   The Rate of Entry

The solution p to (23)-(24) given in (25) is a probability density only for a particular value of E/N. This can be used to determine the amount of entry that must take place relative to the number of existing firms. Since π is a probability density, it follows from (25) and (26) that the steady state rate of attempted entry must be given by:

$$\frac{E}{N} = \left[\int_{S[x]>b}\left(\frac{1-e^{-\theta_*(S[x]-b)}}{\epsilon+\eta}\right)g(x)\mathrm{d}x\right]^{-1} \tag{29}$$

---

[9]Gabaix (2003) has proposed to move away from this abstraction and to consider an economy with large firms to generate aggregate shocks.

[10]Suppose the exogenous exit and population growth rates are zero. Consider the limiting distribution obtained by letting s_0 go to b. This turns the profitability process of a dynasty of firms into a Brownian motion with a negative drift and a reflecting barrier at b. The resulting distribution for e^s is a Pareto distribution on $e^s \geq e^b$ with mean $e^b\theta/(\theta-1)$. In Gabaix (1999), e^s is the size of a city relative to the average city size. This must have mean 1, and so $\theta = 1/(1 - e^b)$. The explanation given in Gabaix (1999) for Zipf's law for relative city sizes is that b must be very small.

This expression simplifies when there is no exogenous exit or population growth. Note that $[1 - e^{-\theta_*(s-b)}]/(\epsilon + \eta)$ converges to $-(S[x] - b)/\mu$ as $\epsilon + \eta$ goes to zero. This is the average life span of a firm with fixed characteristics x that only exits when its profitability reaches b. The entry rate E/N therefore converges to the reciprocal of the average life span $-\mu^{-1} \int_{S[x]>b}(S[x] - b)g(x)\mathrm{d}x$ among entering firms as the exogenous exit and population growth rates go to zero. Absent initial heterogeneity the amount of entry per unit of time is simply the reciprocal of the expected life span of an entering firm, as might be expected.

The assumed linear technology for generating entry attempts implies that the amount of labor used for entry at time t is $L_E e^{\eta t}$, where $L_E/N = \Gamma E/N$. The equilibrium value of this ratio is therefore determined by the entry rate (29). A decrease in the entry cost Γ implies more entry attempts per unit of labor and lowers $e^{S[x]-b}$ by a factor that is common to all x, by the zero profit condition. It follows from (29) that the entry rate E/N is decreasing in the entry cost Γ.

3.4 Firm Characteristics and Selection on Profitability

The stationary population of firm characteristics is generated by a distribution of characteristics among potential entrants, together with a process of pre- and post-entry selection on profitability. The resulting (x, s) marginal of the stationary density (25) is:

$$p(x, s) = \pi(s|S[x]) \left[\frac{\left(\frac{1-e^{-\theta_*(S[x]-b)}}{\epsilon+\eta}\right)g(x)}{\int_{S[x]>b}\left(\frac{1-e^{-\theta_*(S[x]-b)}}{\epsilon+\eta}\right)g(x)\mathrm{d}x} \right] \tag{30}$$

for all $s \geq b$ and $S[x] \geq b$. The second factor in this expression is the marginal $p(x)$ of the stationary density. Selection determines how this marginal differs from the density of fixed characteristics among potential entrants, and from the density of fixed characteristics among actual entrants. Only firms with a sufficiently high initial profitability will enter, and the entering firms with an initial profitability $S[x]$ close to the exit barrier b are more likely to exit over time. As a result, $g(x)$ is truncated at $S[x] = b$ and down-weighted for small $S[x] - b$. As expected, this last effect is strongest when $\epsilon + \eta$ is small. The resulting density is then approximately proportional to $(S[x]-b)g(x)$. On the other hand, if the random exit rate or the population growth rate are large, then $e^{-\theta_*(S[x]-b)}$ is small and the stationary density of fixed characteristics is determined essentially by pre-entry selection. Post-entry selection will be strong also if Γ is sufficiently small so that entry is very easy. This implies that $S[x] - b$ is small for all x, and the resulting stationary density is then again approximately proportional to $(S[x] - b)g(x)$.

21

The initial profitability $S[x]$ depends on the technology $G_{[x]}(\cdot)$ and the overhead and productivity parameters $M[x]$ and $Z[x]$. Other things equal, the process of selection makes firms with low fixed costs $M[x]$, high initial productivity $Z[x]$, and high marginal products of labor $D_L G_{[x]}(k[x], 1)$ more prevalent in the overall population than among entrants or potential entrants.

To emphasize the importance of random productivity growth in shaping the distribution of firm characteristics, it is instructive to consider what happens as the variance of productivity shocks, σ^2, goes to zero. For simplicity, suppose that $\epsilon + \eta = 0$. Assumption 4 then requires $\mu < 0$ and at $\sigma^2 = 0$ one obtains $\xi = (r + \epsilon - [\kappa + \lambda])/|\mu|$ and $b = 0$. Firms exit immediately when they no longer break even since there is no option value to continuing operations. An entering firm starts with profitability s_0, and this profitability will then decline linearly to 0, at which point the firm exits. As σ^2 goes to 0, the tail-index θ grows without bound. Using (27) one can verify that the stationary profitability distribution $\pi(\cdot|S[x])$ converges to a uniform distribution on $(0, S[x])$. In this limiting economy, the largest and most profitable firm conditional on fixed characteristics is the most recent entrant. This is in sharp contrast to what is found in the data. The randomness in productivity growth generates a selection mechanism by which the typical firm can be much larger and productive than recent entrants.

4. The Balanced Growth Path

4.1 Market Clearing

The decision rules (13)-(17) imply that the revenues of a firm with fixed characteristic x in state s at time t are given by $w_t M[x] e^s/(1 - \omega)$. Aggregate output of the composite consumption good is determined by aggregating these revenues using the stationary density of firm characteristics. Physical capital and variable labor inputs for a firm of type (x, s) are also implied by (13)-(17). Recall that per-capita consumption grows at a rate $\kappa + \lambda$ along the balanced growth path. Per capita variable labor inputs must be constant. Let $L_C e^{\eta t}$, $K_C e^{(\lambda + \eta)t}$ and $Y_C e^{(\kappa + \lambda + \eta)t}$ be aggregate variable labor, capital, and aggregate output in the final goods sector. Since $w_t = we^{(\kappa + \lambda)t}$ and $v_t = ve^{\kappa t}$, labor and capital shares are constant and equal to wL_C/Y_C and vK_C/Y_C, respectively. Labor productivity grows at a rate $\kappa + \lambda$ in the final goods sector. The levels of the balanced

22

growth paths of variable inputs and output of final goods are therefore determined by:

$$\frac{1}{wN}\begin{bmatrix} Y_C \\ vK_C \\ wL_C \end{bmatrix} = \frac{1}{1-\omega}\int_b^\infty \int_{S[x]>b}\frac{M[x]e^s}{G_{[x]}(k[x],1)}\begin{bmatrix} G_{[x]}(k[x],1) \\ \omega D_L G_{[x]}(k[x],1) \\ \omega D_K G_{[x]}(k[x],1) \end{bmatrix}p(x,s)\mathrm{d}x\mathrm{d}s \quad (31)$$

The aggregate quantity of overhead labor is denoted by $L_M e^{\eta t}$, where:

$$\frac{L_M}{N} = \int_b^\infty \int_{S[x]>b} M[x]p(x,s)\mathrm{d}x\mathrm{d}s \quad (32)$$

Recall that the ratio of rental prices v/w and the ratios L_I/K_I and K_I/K are determined by balanced growth considerations in the capital accumulation sector. This determines $k[x]$ for all firm types, and together with the zero-profit condition this then determines the function $S[\cdot]$. The equilibrium firm averages $[Y_C/w, K_C, L_C, L_M]/N$ are therefore implied by (31) and (32).

The balanced growth path is now determined by imposing market clearing conditions. For capital and labor, these can be written as:

$$\frac{K_I}{N} = \left[\frac{K}{K_I} - 1\right]^{-1}\frac{K_C}{N} \quad (33)$$

and:

$$\frac{H}{N} = \frac{L_E}{N} + \frac{L_M}{N} + \frac{L_C}{N} + \frac{L_I}{K_I}\frac{K_I}{N} \quad (34)$$

respectively. The ratio K_I/N follows immediately from (33) and the value of K_C/N implied by (31). The labor market clearing condition (34) can then be used to solve for the number of firms N, using the allocation of labor $[L_E, L_M, L_C]/N$ implied by (29) and (31)-(32).

When measured in units of labor, average firm output $Y_C/(wN)$ follows from (31). Together with the market clearing condition $Y_C = C$ and the definition of $S[\cdot]$, this can be used to solve for the level of wages:

$$w = \left(N\int_b^\infty \int_{S[x]>b}\left[\omega Z[x]D_L G_{[x]}(k[x],1)\right]^{\frac{\omega}{1-\omega}}e^{s-S[x]}p(x,s)\mathrm{d}x\mathrm{d}s\right)^{\frac{1-\omega}{\omega}} \quad (35)$$

The equilibrium rental price of capital follows from this and the ratio v/w. This establishes the following proposition.

Proposition 1: *If Assumptions 1-4 hold, then there exists a balanced growth path.*

23

4.2   Market Size, Fixed and Entry Costs, and Per Capita Consumption

The economy described here has implications for the effects of market size and of fixed and entry costs that are exactly the same as those obtained in a static economy of this type, such as Krugman (1979).

Measure market size by population, and consider variation in H. Recall that $k[\cdot]$ and $S[\cdot]$ are determined by balanced growth and zero-profit conditions that do not depend on H. The right-hand side of (31) does not, therefore, vary with H, and (34) then implies that the number of firms and final goods must be proportional to H. It follows from (35) that the elasticity of w and v with respect to the size of the population is $(1-\omega)/\omega$, as in (12). This is also the elasticity of per-capita consumption with respect to H, since $Y_C/H = (Y_C/N)(N/H)$ varies in direct proportion to the wage. The aggregate value of physical capital in the economy is $v_t K_t / r$, and the aggregate value of firms not including physical capital, or "intangible capital," is equal to $N_t w_t$ times the average of $M[x]V(s)$ in the population. Changes in H do not affect the ratios of these aggregates over aggregate consumption.

The effects of entry and fixed costs are easiest to examine by considering what happens when the entry costs Γ and fixed costs $M[\cdot]$ are scaled down by a common factor. The zero-profit condition only depends on the ratio $M[\cdot]/\Gamma$, and so profitability at entry $S[x] - b$ is not affected. It follows from (29), (31)-(32) and (33) that $[L_E, L_M, L_C, L_I]/N$ declines by the same factor as $(\Gamma, M[\cdot])$. The labor market clearing condition then implies that the number of firms must rise by the same factor. A reduction in entry and fixed costs has exactly the same effect on the number of firms as an increase in the population. The elasticity of wages with respect to fixed and entry costs is therefore $-(1 - \omega)/\omega$. The resulting elasticity of per capita consumption is also $-(1 - \omega)/\omega$, since the increase in N and the reduction in $M[\cdot]$ exactly cancel each other in (31). The aggregate allocation of labor and, in particular, the aggregate expenditure on entry and fixed costs measured in units of labor do not change when entry and fixed costs per firm decline. As expected, the effect on final goods consumption can be large if goods are very imperfect substitutes.

4.3   Entry Costs and Per Capita Consumption

The effects of varying the entry cost Γ only while keeping the fixed costs $M[\cdot]$ constant are more complicated than those of a parallel shift in $(\Gamma, M[\cdot])$. An increase in Γ implies that $C/w^{1/(1-\omega)}$ and the function $e^{S[x]-b}$ are scaled up. Entrants must earn higher profits to warrant the higher expense of entry. The shift in $e^{S[x]-b}$ will in general affect

the distribution of characteristics among entrants, and this complicates tracing out the effects of an increase in Γ. If there is no heterogeneity in fixed characteristics, then the entry rate E/N will unambiguously go down. New entrants will enter farther away from the exit boundary, and this implies that average life spans will be longer. As a result, it takes less entry to maintain a given population of firms. Per firm, the amount of labor required for entry is $L_E/N = \Gamma(E/N)$. Using the zero profit condition (22) and the equation (29) for E/N, one can verify that L_E/N increases with an increase in Γ. The reduction in the entry rate E/N does not off-set the increase in Γ. Since $e^{S[x]-b}$ rises, firms will on average be larger, and one can use (31) and (32) to show that labor and capital inputs per firm will rise. Together with (33)-(34) this implies that the number of firms will decline. Note from (35) that wages are an increasing function of N times the population average of $e^{s-S[x]}$. This is the ratio of average firm profitability over the profitability of entrants. An increase in $S[x]$ causes this ratio to fall, by (28). It follows that wages fall as a result of an increase in entry costs. Aggregate consumption is equal to the population average of $M[x]e^s$ times wN. Although the population average of $M[x]e^s$ increases with an increase in Γ, both and w and N decline. It is possible to construct examples in which aggregate consumption is an increasing function of the entry cost over some range.[11]

4.4 The Limiting Case of Zipf's Law

Absent any heterogeneity in fixed characteristics, Zipf's law corresponds to $\theta = 1$. This case is ruled out in Assumption 4. But the balanced growth path is well defined for any θ greater than 1. To examine the nature of the balanced growth path for economies for which Zipf's law is a very good approximation, consider what happens if θ approaches 1 from above as ζ_* increases to the boundary of the region defined by Assumption 4. The parameter ζ_* does not affect the balanced growth rates or the interest rate. The zero-profit condition and the resulting scale of $e^{S[x]-b}$ vary continuously with θ, and are well defined at $\theta = 1$. The entry rate relative to the number of firms defined in (29) converges to a finite and positive number. The same is true for the fixed costs given in (32). Thus L_E/N and L_M/N will converge to finite and positive limits. But, as can be seen from (28) and (31), the level of output per firm measured in units of labor grows without

---

[11]Lagos (2001) has emphasized selection effects in determining the relation between the magnitude of search frictions and the level of aggregate productivity. The mechanism here is similar and the economy described here might be useful to further examine the effects of allocation frictions on aggregate productivity.

bound, and so do the required capital and labor inputs. It follows from the capital and labor market clearing conditions (33)-(34) that the number of firms must converge to zero. An immediate implication is that $L_E + L_M$ goes to zero. For θ close to one, almost all labor is allocated as variable labor to operate a small set of large firms, and almost no labor is used to cover fixed and entry costs. It may be hard to obtain precise empirical analogues to the fixed and entry costs as they are defined in this economy. But it would seem that the implication of Zipf's law that aggregate expenditures on fixed and entry costs are essentially zero is too extreme.

If $\epsilon + \eta$ is zero, then $\theta = 1$ implies that $\mu + \sigma^2/2$ is zero, and thus $\mu < 0$ as long as $\sigma^2 > 0$ and Assumption 4 is required to hold. That is, incumbent firms must experience slower productivity growth than potential entrants. If firm productivity growth is deterministic, then the opposite case of $\mu > 0$ can be analyzed separately. A balanced growth path for such an economy has no entry and an initial population of firms that stays in business forever. Aggregate consumption and wages grow at a rate $\zeta_* + \lambda$ instead of $\zeta + \lambda$. The resulting lack of entry and exit is at odds with the data.[12] The properties of economies in which Assumption 4 fails and in which productivity growth is stochastic remain to be examined.

4.5   A Possibly Anomalous Implication

The price of capital in terms of consumption goods is given by $q_t = qe^{\kappa t}$ along the balanced growth path. The parameter κ is defined in (12) and measures the growth rate of total factor productivity in the final goods sector. Many authors have argued that the price of capital has declined in terms of consumption at a steady rate in post-war U.S. data (see Gordon (1990) and Greenwood, Hercowitz and Krusell [1997]). One possibility is to interpret this decline as an adjustment to the balanced growth path, with the marginal product of capital falling over time as more and higher quality capital is used in the production of consumption goods. But the adjustment would have to be extremely slow. Using the economy described here, the decline in the price of capital can be interpreted as a balanced growth phenomenon by taking ζ negative. This means that total factor productivity in the final goods sector, accounting for both quantity and quality changes, grows at a negative rate. This is probably not very plausible.

An alternative explanation is that the growth rate of the measured consumer price

---

[12]Dunne, Roberts and Samuelson (1988) report 5-year entry and exit rates for manufacturing firms on the order of .5. That is, roughly 10% of the population of manufacturing firms exits every year and is replaced by new firms.

index significantly over-estimates the growth rate of the actual cost of a constant-utility bundle of goods. For recent discussions of this issue, see Gordon (2000) and Klenow (2003). If an appropriately measured consumer price index grows more slowly than the one commonly used, then the measured price of capital would not decline as much or even rise in terms of consumption. The composite commodity described in this paper involves a constantly changing set of goods with quality weights that are constantly changing as well. Even complete data on quantities and prices of all final goods consumed in this economy would not allow one to measure the composite commodity since the utility weights are unobservable. Further identifying assumptions would be needed.

A related interpretation can be constructed by amending the model with a retail sector that converts labor and the differentiated output of firms into consumption. Specifically, take actual consumption to be $C_t^\varrho L_{R,t}^{1-\varrho}$ for some $\varrho \in (0,1)$. C_t is the composite commodity used throughout the paper, and $L_{R,t}$ is labor used in the retail sector. Note that there is no labor-augmenting technological progress in the retail sector. The Cobb-Douglas assumption is necessary for balanced growth in this modified economy. The composite commodity is appropriately interpreted as an intermediate good and consumption is the output of the retail sector. Along the balanced growth path, $L_{R,t}$ is constant and the price of consumption in terms of the intermediate good will rise at a rate $(1-\varrho)(\kappa+\lambda)$. As a result, the price of capital in terms of consumption grows at a rate $\varrho(\kappa+\lambda)-\lambda$. This will be negative if λ is a sufficiently large part of the growth rate $\kappa+\lambda$ of intermediate output, and if the share of labor in the retail sector is sufficiently large.

The key to this interpretation is a slower rate of labor-augmenting technological progress in the retail sector than in the up-stream sectors of the economy. This is also the essence of the explanation offered in Greenwood, Hercowitz and Krusell (1997). It relies heavily on the Cobb-Douglas assumption to generate balanced growth.

5. Two Interpretations

The main assumption used to construct a stationary distribution of firm size with a finite mean is that ζ is sufficiently large relative to ζ_*. In particular, if $\epsilon = \eta = 0$, then $\theta > 1$ means that $\zeta > \zeta_* + \frac{1}{2}\sigma_*^2\omega/(1-\omega)$. The productivity growth rate for new entrants must exceed that of incumbent firms by an amount that can be large when differentiated final goods are close substitutes or when the variance σ_*^2 is large.

This section presents two possible interpretations of the assumption that ζ is suffi-

ciently large relative to ζ_*. In one, the adaptation by existing producers to new and improved inputs is imperfect, while new entrants are perfectly adapted. In the other, some of the information generated as a result of learning-by-doing by incumbents can be used by new entrants. If this information spillover is not too small, then ζ will exceed ζ_* by enough to generate a stationary distribution of firm characteristics with a finite mean.

5.1   Limited Adaptation

Suppose that the "quality" of capital at time t is measured by Q_t and let K_t/Q_t denote the constant-quality stock of capital at time t. If $Q_t = Qe^{\varkappa t}$, then the technology for accumulating constant-quality capital is the same as (2), with Λ_t replaced by Λ_t/Q_t and δ replaced by $\delta + \varkappa$. If capital inputs for final goods producing firms are measured in constant-quality units, then $z_{t,a}$ and $\Lambda_{t,a}$ in (9) must be replaced by $z_{t,a}Q_{t+a}$ and $\Lambda_{t,a}/Q_{t+a}$, respectively. The growth rates κ and λ change to $\kappa + \varkappa$ and $\lambda - \varkappa$. Of course, the growth rate of output is not affected by the units in which capital is measured, but the decomposition of this growth rate into total factor productivity growth and labor-augmenting technical progress is. The technology (2) can be given a vintage capital interpretation as in Solow (1960), by taking $F(K_{I,t}/Q_t, (\Lambda_t/Q_t)L_{I,t})$ to be the output of capital of vintage t, with quality level Q_t.

To describe limited adaptation by incumbent firms, consider a final goods producer with fixed characteristics x who entered at time t, and modify the technology for this producer as follows:

$$y_{t,a} = z_{t,a}^* G_{[x]}(Q_{t,a}K_{t,a}^*, \Lambda_{t,a}L_{t,a}) \tag{36}$$

where $K_{t,a}^* = K_{t,a}/Q_{t+a}$ represents the quantity of constant-quality capital used by the firm at time $t+a$. That is, constant-quality capital $K_{t,a}^*$ and hours of labor $L_{t,a}$ translate into firm-specific efficiency units determined by $Q_{t,a}$ and $\Lambda_{t,a}$, respectively. Suppose it is the case that $Q_{t,a}/Q_{t+a} = \Lambda_{t,a}/\Lambda_{t+a}$. Then the modified technology (36) corresponds to the original technology (9) with $z_{t,a}$ defined as:

$$z_{t,a} = \frac{z_{t,a}^* Q_{t,a}}{Q_{t+a}} = \frac{z_{t,a}^* \Lambda_{t,a}}{\Lambda_{t+a}}$$

If $Q_{t,a}/Q_{t+a} = \Lambda_{t,a}/\Lambda_{t+a} < 1$, then this provides an interpretation of the assumption that firm productivity grows more slowly on average than aggregate productivity. Although the quality of capital and labor inputs available to the firm improves over time, the firm can only make use of these quality improvements to a certain extent. If there is no

improvement in the way in which the firm can make use of firm-specific efficiency units, then $z_{t,a}^*$ does not depend on age a and the discrepancy $\zeta_* - \zeta$ measures precisely the extent to which a typical firm cannot take advantage of the quality improvements of its inputs.

5.2 Learning-by-Doing with Spillovers

Firms generate improvements in their firm-specific total factor productivity at an average rate ζ_*, as long as they pay the fixed cost M. This can be interpreted as learning-by-doing, and the rate at which learning takes place is taken as given. Firms do face an extensive margin: by staying in business a firm chooses to continue to improve the technology available to the firm. But what determines the rate at which the technology of potential entrants improves? In the model described so far, potential entrants invest in new products and techniques by sampling from a distribution, and this distribution improves for reasons that are left unexplained.

One possibility is that every potential entrant has to start from scratch. In the context of the economy described here, this can be interpreted to mean $\zeta = 0$. If $\zeta_* > 0$, this violates the conditions for a stationary distribution of firms to emerge. But one can guess what will happen: the average growth of incumbents and the selection that arises from $\sigma_* > 0$ will generate a population of incumbents that is farther and farther removed from the technology that is available to potential entrants. As a result, the value of trying to enter from scratch will become very small: the technology likely to be obtained by an entrant will become more and more inefficient relative to the technologies of the surviving incumbents. Very little if any entry will occur. Attempting to enter the car market in 2004 is not likely to be profitable for an entrant who first needs to re-invent the wheel.

To generate a steady-state level of successful entry, some amount of copying by potential entrants must be possible. This will be modeled here by assuming that potential entrants can start with an initial technology that is related to the population of technologies already in use. This generates an externality that is well known to be a potential source of growth (Lucas [1988]). In Eaton and Eckstein (1997), spillovers across cities are the mechanism by which the distribution of human capital across cities is kept from spreading out too fast.

For simplicity, heterogeneity in fixed characteristics x is ruled out in the following. All x arguments are therefore omitted. Let X_t be the average of $Z_{t-a,a}$ in the population

at time t. It follows from (16) and (30) that:

$$X_t = Z_{t,0} \int_b^\infty e^{(s-S)(1-\omega)/\omega} \pi(s|S) \mathrm{d}s \tag{37}$$

where $Z_{t,0}$ is the initial productivity level available to any date-t entrant. Suppose now that an entrant at time t can start with a level of productivity that is a certain fraction ϕ of the average productivity in the population:

$$Z_{t,0} = \varphi X_t \tag{38}$$

There is no need for φ to be greater than one, and X_t can turn out to be far below the productivity of the set of firms that account for most of aggregate output. Even if φ is not large, the specification (38) will ensure that no entrant has to start from scratch after a certain amount of growth in X_t. There will however be a lower bound on the set of φ that are consistent with a steady level of entry and a stationary distribution of firm size.

Combining (37) and (38) we obtain:

$$1 = \varphi \int_b^\infty e^{(s-S)(1-\omega)/\omega} \pi(s|S) \mathrm{d}s \tag{39}$$

This is a new equilibrium condition that can be used to solve for the growth rate ζ. The right-hand side of (39) depends directly on ζ via the definition of $\pi(\cdot|S)$. Furthermore, the growth rate κ of productivity in the final goods sector depends on ζ via (12). In turn, κ determines the interest rate r and ratio of factor prices v/w via the balanced growth conditions (4)-(8). These prices influence the profitability of entry and the value of continuing operations as an incumbent, and thereby the values of S and b.

Proposition 2: *Suppose the utility function is logarithmic and that the price elasticity of the demand for differentiated goods is greater than two. Then there exists a minimum spillover parameter $\underline{\varphi} > 0$ so that for all $\varphi > \underline{\varphi}$ the economy has a balanced growth path with an average technology growth rate ζ determined by (39). The growth rate ζ and the tail index θ are increasing functions of φ.*

The assumption on the price elasticity ensures that the mean productivity level is finite whenever mean profitability is finite, so that it makes sense to have spillovers of the form (38). Because preferences over the composite commodity are taken to be logarithmic, Assumption 1 does not depend on κ, and therefore not on ζ either. This means that balanced growth requirements do not restrict the range of ζ. Furthermore, $r - \kappa = \rho + \lambda$

and so the discount rate used to define the value function $V(\cdot)$ also does not depend on κ. The parameter ξ is then unambiguously decreasing in ζ. As a result, an increase in ζ causes $V(S)$ to decline for any fixed $S - b$. Faster technology improvements outside the firm will make incumbents at a fixed distance above the exit barrier b less profitable. The zero-profit condition $\Gamma = V(S)$ then implies that $S - b$ must be an increasing function of ζ. An increase in ζ also increases θ and reduces θ_*. Using the stationary density $\pi(\cdot|S)$ one can verify that the mean of $e^{(s-S)(1-\omega)/\omega}$ declines as a function of ζ. The equilibrium condition (39) then implies the results of Proposition 2.

An example of the relation between spillovers, growth and the thickness of the tail of the size distribution is given in Figure 3. The example assumes utility is logarithmic and the price elasticity of demand is 2.5. There is no population growth, no exogenous exit by firms, and no labor-augmenting technological progress. Furthermore, $\Gamma = M/\rho$, so that the entry cost is the same as the present value of the fixed costs for a firm that is never shut down. The subjective discount rate ρ is 3%, and the average growth of productivity inside the firm is zero. Its standard deviation is 20%. In this economy, all growth is due to spillovers in the final goods sector, and the figure shows that the smallest spillover consistent with a stationary equilibrium generates a growth rate in total factor productivity of 3%.
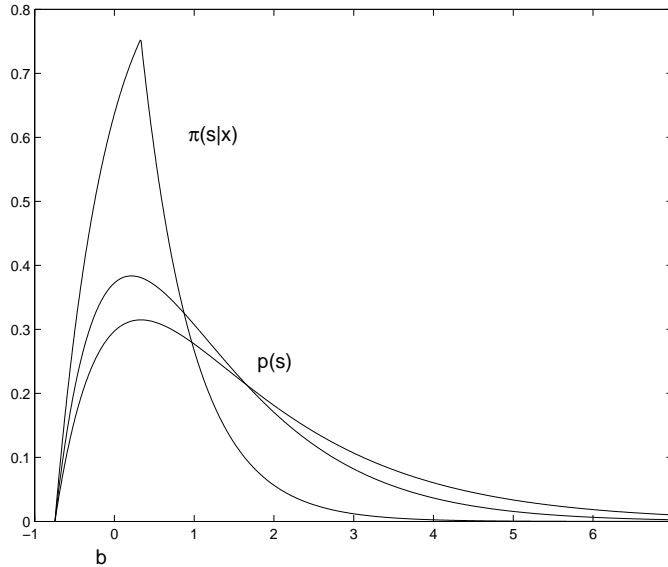

Figure 3: Learning-by-Doing and Spillovers

In this economy, firms do not have an intensive margin along which they can affect the growth rate of their productivity. Making it easier to copy from existing firms leads

unambiguously to faster growth. Because of the shorter average life-span of firms, firms will not have a chance to grow large and this makes for a size distribution of firms with a thin tail. The thick tail observed in firm size data indicates that φ cannot be too large. Organization capital must be relatively hard to copy.

6. Concluding Remarks

Growth in the economy described in the last section is generated by learning-by-doing in production, and spillovers of the knowledge created in the process. Average information about the operations of existing production units can be used by new entrants. Entry drives up wages and thus the fixed cost needed to keep existing production units in operation. This induces exit of the least efficient production units. Low-productivity production units are replaced by production units that are closer to the average in terms of their productivity. This replacement is a source of growth beyond the growth of productivity that occurs in existing production units. The fat tail of the size distribution of U.S. firms implies that the growth generated by entry and exit can account for only a relatively modest part of the growth of aggregate output.

It is not difficult to calculate exit rates and growth rates conditional on survival for the economy described here. For firms with a particular set of fixed characteristics, exit rates are nearly zero soon after entry. They then increase to a maximum and decline to a long-run constant exit rate. In contrast, the data suggest that exit rates of firms decline monotonically with age. This monotonically declining hazard rate can be captured in the model described here by averaging over firms with heterogeneous initial profitability levels. Heterogeneity in firm characteristics is also needed, in the context of the economy described here, to interpret the empirical observation that firm exit rates are not only correlated with size but also with age. Luttmer (2004) characterizes the heterogeneity and the average growth and variability of productivity shocks needed to fit the joint distribution of age and size, exit rates conditional on age and size, and survivor growth rates conditional on age and size, as reported for U.S. manufacturing plants by Dunn, Roberts and Samuelson (1989). The average growth and variability of productivity growth are intimately linked to the value of a firm. Given data on firm value, this provides an important over-identifying restriction for the economy described here. Examining this restriction is the subject of ongoing research.

A potentially important extension of the economy described here is to allow for capacity constraints. One can suppose that entrants not only pay entry costs and draw

fixed cost and input requirement parameters, but also a maximum scale. This is an extension that remains analytically tractable and that is likely to be useful when examining plant-level data. In particular, this extension may help in interpreting differences between the size distributions of firms and plants. It also allows one to examine an economy in which there are categories of final goods that are perfect substitutes.

References

[1] Atkeson, A. and P.J. Kehoe (2002): "Measuring Organization Capital," Federal Reserve Bank of Minneapolis, Research Department Staff Report 291.

[2] Axtell, R.L. (2001): "Zipf Distribution of U.S. Firm Sizes," Science, vol. 293, pp. 1818-1820.

[3] Dixit, A.K, and J.E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," American Economic Review, vol. 67, no. 3, pp. 297-308.

[4] Dixit, A.K. and R.S. Pindyck (1994): Investment under Uncertainty, Princeton University Press.

[5] Dunne, T., M.J. Roberts, and L. Samuelson (1988): "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," RAND Journal of Economics, vol. 19, no. 4, pp. 495-515.

[6] Dunne, T., M.J. Roberts, and L. Samuelson (1989): "The Growth and Failure of U.S. Manufacturing Plants," Quarterly Journal of Economics, vol. 104, no. 4, pp. 671-698.

[7] Eaton, J. and Z. Eckstein (1997): "Cities and Growth: Theory and Evidence from France and Japan," Regional Science and Urban Economics, vol. 27, pp. 443-474.

[8] Gabaix, X. (1999): "Zipf's Law for Cities: An Explanation," Quarterly Journal of Economics, vol. 114, pp. 739-767.

[9] Gabaix, X. (2004): "Power Laws and the Origins of the Aggregate Fluctuations," MIT working paper.

[10] Gibrat, R. (1931): Les Inégalités économiques, Paris.

[11] Gordon, R.J. (1990): The Measurement of Durable Goods Prices, NBER, University of Chicago Press.

[12] Gordon, R.J. (2000): "The Boskin Commission Report and Its Aftermath," NBER working paper no. 7759.

[13] Greenwood, J., Z. Hercowitz and P. Krusell (1997): "Long-Run Implications of Investment-Specific Technological Change," American Economic Review, vol. 87, no. 3, pp. 342-362.

[14] Greenwood, J. and B. Jovanovic (2001): "Accounting for Growth," in New Developments in Productivity Analysis, E.R. Dean, M.J. Harper and C.R. Hulten eds., NBER and University of Chicago Press.

[15] Harrison, J.M. (1985): Brownian Motion and Stochastic Flow Systems, Wiley.

[16] Klenow. P.J. (2003): "Measuring Consumption Growth: The Impact of New and Better Products," Federal Reserve Bank of Minneapolis Quarterly Review, vol. 27, no. 2, pp. 10-23.

[17] Krugman, P.R. (1979): "Increasing Returns, Monopolistic Competition, and International Trade," Journal of International Economics, vol. 9, pp. 469-479.

[18] Hopenhayn, H. (1992): "Entry and Exit and Firm Dynamics in Long Run Equilibrium," Econometrica, vol. 60, no. 5, pp. 1127-1150.

[19] Ijiri, Y., and H.A. Simon (1964): "Business Firm Growth and Size," American Economic Review, vol. 54, pp. 77-89.

[20] Jovanovic, B. (1982): "Selection and the Evolution of Industry," Econometrica, vol. 50., no. 3, pp. 649-670.

[21] Karlin, S. and H.M. Taylor (1981): A Second Course in Stochastic Processes, Academic Press.

[22] Klette, T. and S. Kortum (2004): "Innovating Firms and Aggregate Innovation," Journal of Political Economy, forthcoming.

[23] Krugman, P. (1979): "Increasing Returns, Monopolistic Competition, and International Trade," Journal of International Economics, vol. 9, pp. 469-480.

[24] Lagos, R. (2001): "A Model of TFP," New York University working paper.

[25] Lucas, R.E. Jr. (1978): "On the Size Distribution of Business Firms," Bell Journal of Economics, vol. 9, no. 2, pp. 508-523.

[26] Lucas, R.E. Jr. (1988): "On the Mechanics of Economic Developement," Journal of Political Economy, vol. 22, no. 1, pp. 3-42.

[27] Luttmer, E.G.J. (2004): "On the Age and Size Distribution of Business Firms," work in progress.

[28] Melitz, M. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," Econometrica, vol. 71, no. 6, pp. 1695-1725.

[29] Parente, S.L. and E.C. Prescott (2000): Barriers to Riches, MIT press.

[30] Rebelo, S. (1991): "Long-Run Policy Analysis and Long-Run Growth," Journal of Political Economy, vol. 99, pp. 500-521.

[31] Romer, P. (1990): "Endogenous Technological Change," Journal of Political Economy, vol. 98, no. 5, pp. 71-102.

[32] Rossi-Hansberg, E. and M.L.J. Wright (2003): "Urban Structure and Growth," Stanford University.

[33] Rossi-Hansberg, E. and M.L.J. Wright (2004): "Firm Dynamics in the Aggregate Economy," Stanford University.

[34] Simon, H.A., and C.P. Bonini (1958): "The Size Distribution of Business Firms," American Economic Review, vol. 48, pp. 607-617.

[35] Solow, R.M. (1960): "Investment and Technical Progress," in Mathematical Methods in the Social Sciences, 1959, Proceedings of the First Stanford Symposium, edited by K.J. Arrow, S. Karlin and P. Suppes, Stanford University Press.

[36] Sutton, J. (1997): "Gibrat's Legacy," Journal of Economy Literature, vol. 35, pp. 40-59.

[37] Uzawa, H. (1961): "On a Two-Sector Model of Economic Growth," Review of Economic Studies, vol. 29, no. 1, pp. 40-47.