



ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS
WORKING PAPER SERIES

Heterogeneous Firms, Productivity, and Poverty Traps

Authors	Levon Barseghyan, and Riccardo DiCecio
Working Paper Number	2005-068B
Revision Date	October 2005
Citable Link	https://doi.org/10.20955/wp.2005.068
Suggested Citation	Barseghyan, L., DiCecio, R., 2005; Heterogeneous Firms, Productivity, and Poverty Traps, Federal Reserve Bank of St. Louis Working Paper 2005-068. URL https://doi.org/10.20955/wp.2005.068

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

The views expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment.

Heterogeneous Firms, Productivity, and Poverty Traps*

Levon Barseghyan
Cornell University

Riccardo DiCecio
Federal Reserve Bank of St. Louis

December, 2006

Abstract

We present a model of endogenous total factor productivity which generates poverty traps. We obtain multiple steady-state equilibria for an arbitrarily small degree of increasing returns to scale. While the most productive firms operate across all the steady states, in a poverty trap less productive firms operate as well. This results in lower average firm productivity and lower total factor productivity. In our model a growth miracle is accompanied by a shift of employment from small to large firms, consistent with the empirical evidence. We calibrate our model and relate entry costs to the price of investment goods. The resulting distributions of output, TFP, and capital-to-output ratio across steady states are similar to the ergodic distributions we estimate from the data.

JEL: L16, O11, O33, O40

Keywords: endogenous productivity, multiple equilibria, poverty traps

*We thank Daron Acemoglu, Costas Azariadis, Larry Blume, Helge Braun, Paco Buera, Jim Bullard, Steve Durlauf, Oded Galor, Espen Henriksen, Nir Jaimovich, Per Krusell, Kiminori Matsuyama, Francesca Molinari, Assaf Razin, Richard Rogerson, Karl Shell, Gustavo Ventura, Ivan Werning, participants at the Cornell-Penn State Macro Workshop, the IFPRI *Conference on Threshold Effects and Non-Linearities in Growth and Development*, the 2006 Federal Reserve System Macroeconomics Meeting, seminars at George Washington University and Duke University, and the NBER Summer Institute for comments. All errors are our own. Any views expressed are our own and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System. Corresponding author: Riccardo DiCecio, dicecio@stls.frb.org.

1 Introduction

We present an endogenous total factor productivity (TFP) model that leads to multiple steady-state equilibria and, hence, poverty traps. Our model is a variant of the neoclassical growth model with increasing returns introduced by Benhabib and Farmer (1994), with firms modeled in the tradition of Lucas (1978), Jovanovic (1982), and Hopenhayn (1992). There are many ex-ante identical potential firms which face an entry cost. Firms that choose to enter are entitled to produce an intermediate good with a productivity level drawn independently across firms from a given distribution. Because firms face a fixed operating cost, the decision to operate depends on the level of the firm's productivity. Productivity must be high enough so that the firm generates enough revenue (net of payments to factor inputs) to cover the operating cost. In other words, the operating cost defines a cutoff: firms with productivity above the cutoff choose to operate, the rest of the firms choose not to. The higher the cutoff, the more productive the average firm is.

The existence of multiple steady states depends on small demand externalities, which imply increasing returns to scale at the aggregate level. One of the main results of our paper is that poverty traps may occur for arbitrarily small increasing returns to scale.¹ Endogenizing TFP allows us to bridge the gap between poverty trap models based on increasing returns and the most recent empirical literature on the degree of returns to scale. An endogenous operating cost provides a powerful amplifying mechanism for increasing returns. We model the operating cost as payments to overhead labor. Since the wage is endogenous, so is the lowest level of productivity used in the economy. This endogeneity may lead to multiple steady states. Consider an economy in a steady state with a high productivity cutoff and a large capital stock. The high cutoff implies that firms' average productivity is high. A large capital stock and high productivity imply that the wage is high, as is the operating cost. A high operating cost makes low productivity firms unprofitable, effectively cleansing the pool of firms. This justifies why the cutoff is high in the first place. Since only high productivity firms are operating, TFP is high. Conversely, in a steady state where capital is low and lower productivity firms are operating (i.e., the cutoff is low), the

¹Galí (1995) obtains multiple equilibria and poverty traps in a model where increasing returns stem from endogenous markups. However, the empirical evidence summarized in Section 4.2 below suggests that the degree of increasing returns is much smaller than the level required to obtain poverty traps in Galí's model.

wage is low. Since the wage is low, lower profits are sufficient to cover the operating cost. That is, the low operating cost sullies the pool of producers, leading to lower TFP and capital. Notice that in a good equilibrium high productivity firms produce more than in a bad equilibrium, despite facing a higher wage and the same interest rate. This is optimal because firms face a higher demand for their goods, which offsets the contractionary pressure of higher factor prices.

An empirical motivation for our work comes from the studies of the determinants of cross-country income differences of Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), and Caselli (2005). These authors find that income differences can be attributed, at least in part, to differences in TFP. Previous studies of poverty trap models with endogenous TFP pointed to the failure of adopting the most productive technology as the cause of low TFP and income in poor countries.² However, there is evidence pointing to the fact that differences in TFP across economies are related to the lowest level of firms' productivity. Comin and Hobijn (2004) take a comprehensive look at the uses of various technologies as determinants of TFP and find that the key is not when new, better technologies are adopted, but when old, obsolete ones are relinquished. Also, the empirical evidence on the importance of international knowledge spillovers summarized in Klenow and Rodriguez-Clare (2005) suggests that all countries can easily access frontier technologies. Banerjee and Duflo (2005) cite the McKinsey Global Institute (2001) report on India, which finds that while larger production units (firms) use relatively new technologies, smaller (in home) production units have low productivity. Finally, Mokyr (1990, 2001) argues that the Industrial Revolution was characterized by a shift from less productive forms of production (workshops) to more productive ones (factories).

A successful model of cross-country income and productivity differences should also provide a plausible story of how a “growth miracle” can occur, i.e., it should be consistent with the transition of a country from low to high output and productivity. In our model economy, a growth miracle is a transition from a bad equilibrium (low productivity cutoff) to a good one (high productivity cutoff). Such a take-off can be triggered by technological progress that makes the highest productivity firms even more productive or

²See, for example, Murphy, Shleifer, and Vishny (1989) and Ciccone and Matsuyama (1996). For comprehensive reviews of the literature on poverty traps, see Matsuyama (2005) and Azariadis and Stachurski (2005).

by a decline in the entry cost. In the first case, the increase in productivity of the best firms makes them more competitive, raising factor prices and driving low productivity firms out of business. In the second case, a decline in the entry cost brings about more competition from entering firms, driving out of the market low productivity firms. In both cases, along the transition path, the economy's TFP, output, capital, and firms' average productivity (and size) rise. An increase in the average firm size, caused by a massive shift of employment from small to large establishments, is a defining feature of the Industrial Revolution. A similar increase is recorded in the case of Japan's growth miracle. Between 1957 and 1969, the employment share of Japan's smallest establishments declined from 41 to 31.5 percent.

A calibrated version of our model provides an explanation to a number of closely related empirical regularities established in the literature, namely the existence of threshold effects, multi-modal long-run distribution of income across countries, and club convergence.³ Both barriers to capital accumulation and to entry have been associated with the underlying state of a country's institutions. Consistently with the findings in Tan (2005), in our model institutions define convergence clubs. Our measure of institutional quality, the price of investment goods,⁴ has a unimodal distribution in the data. TFP, income, and the capital-to-output ratio have multi-modal ergodic distributions. This can be interpreted as evidence of threshold effects. Furthermore, conditional on high values of the price of investment goods, the income distribution is unimodal and concentrated at low levels of income; conditional on low values, it is also unimodal and concentrated at high levels of income. Yet for intermediate values of the price of investment goods, the distribution of income is multi-modal, consistent with the existence of multiple steady states. The long-run distributions of TFP and of the capital-to-output ratio display similar behaviors. In our model there are threshold effects in two parameters: the price of investment goods and the entry cost. A higher price of investment goods discourages capital accumulation and lowers wages and the operating cost. Therefore, it leads to a lower level of the productivity cutoff and TFP. If the density of firms with productivity draws around the cutoff level is high, then even a small decrease in the price of capital will lead to a significant increase in the firms' average productivity and TFP,

³See Durlauf and Quah (1999) and Quah (2006) for extensive reviews of the literature and references.

⁴The idea that differences in the price of capital are driven by institutional features goes back at least to Easterly (1993) and Jones (1994).

i.e., it will have a threshold effect. Economies whose investment goods are priced above the threshold level will have significantly lower TFP and output than economies whose price of investment goods is below the threshold.⁵ Differences in TFP and output will be accompanied by differences in the distribution of labor across firms of different sizes. In economies above the threshold employment will be concentrated in small, low-productivity units; in economies below the threshold, the majority of workers will be employed in larger, more productive units. Such differences in the distribution of labor coincide with the patterns of cross-county differences in the employment distribution in the manufacturing sector reported by Tybout (2000).

Finally, our model's implications are consistent with the literature which explores the effects that various barriers have on productivity: e.g., Parente and Prescott (1994, 2000), Restuccia and Rogerson (2003), Erosa and Hidalgo Cabrillana (2005), Herrendorf and Teixeira (2005a,b), and Barseghyan (2006).

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 studies its steady state, dynamics properties, and some extensions. Section 4 provides an interpretation of growth miracles that arise naturally in the model and discusses the quantitative implications of a calibrated version of the model. Section 5 compares the distribution dynamics of output, TFP, and the capital-to-output ratio for the data and for our calibrated model when heterogeneity across countries is driven by different prices of capital goods. We conclude in Section 6. We provide proofs and describe the data used in the paper in two appendices.

2 The Model

Our model is a variant of the neoclassical growth model. The model departs from the standard framework by having a richer structure of the production side of the economy. Firms are heterogenous: each firm has monopoly power over the good it produces, and firms have different productivity levels. Two features of the production side of the economy are crucial for the results of

⁵For countries with the price of capital at or near the threshold multiple steady states are possible. However, the existence of multiple steady states is not necessary for the threshold effects to exist. Threshold effects may arise even if there are no demand externalities, i.e., under constant returns to scale. See Durlauf and Johnson (1995) for a discussion of the identification issues in the presence of multiple equilibria and threshold effects.

the paper:

1. a sunk entry cost;
2. an operating cost: in addition to capital and labor used directly in production, firms pay for a fixed amount of overhead labor.

A part of the entry costs stems from satisfying different official regulatory requirements (see Djankov, La Porta, Lopez-de-Silanes, and Shleifer, 2002). In addition, in some countries, entry requires significant side payments to local officials.⁶ Entry cost may also include expenses related to acquisition of firm-specific capital,⁷ acquisition of appropriate technology,⁸ and market research.

The operating cost typically refers to overhead labor and expenses that are lumpy in nature (e.g., renting a physical location). According to the findings of Domowitz, Hubbard, and Petersen (1988), in U.S. manufacturing plants, the overhead labor accounts for 31 percent of total labor. Ramey (1991) suggests that overhead labor is about 20 percent. The preferred estimate of overhead inputs in Basu (1996) is 28 percent.

We also assume that firms learn their productivity only after a sunk entry cost is paid. This assumption reflects very high uncertainty faced by entering firms. This is routinely found in the data and documented, for example, by Klette and Kortum (2004) as a stylized fact.

2.1 Households

There is a continuum of households. They supply a fixed amount of labor, consume, invest, and own all firms in the economy. The problem of the representative household is given by

⁶In the case of Peru, this is documented by De Soto (1989).

⁷Ramey and Shapiro (2001) show that in some instances the specificity of firm capital is so extreme that the sale price of such capital after a firm has been dissolved is only a small fraction of the original cost.

⁸See, for example, Atkeson and Kehoe (2005).

$$\begin{aligned}
& \max \sum_{t=0}^{\infty} \beta^t U(C_t), \quad \beta \in (0, 1) \\
& \text{s.t. } C_t + I_t = r_t K_t + w_t + \Pi_t + T_t, \\
& \quad I_t = K_{t+1} - (1 - \delta) K_t,
\end{aligned} \tag{1}$$

where C_t denotes consumption, I_t is investment, K_t denotes the total household capital, r_t is the rental rate on capital, and w_t is the wage.⁹ Π_t is the firms' profits, and T_t is a lump-sum transfer from the government; β and $\delta \in (0, 1)$ are the discount rate and depreciation rate, respectively. We assume a constant elasticity of substitution utility function with elasticity $\sigma > 0$.

2.2 Firms

2.2.1 Final Good Producers

The final consumption good in this economy is produced by perfectly competitive firms, according to the following production function:

$$Y_t = \left[\int_0^{\mu_t} [y_t(i)]^{\frac{1}{\lambda}} di \right]^{\lambda},$$

where μ_t is the number of intermediate goods produced in the economy, λ is a constant which is greater than 1, and $y_t(i)$ is the quantity of the intermediate good i . Let $p_t(i)$ be the price of the i^{th} intermediate good relative to the final good. Then, the maximization problem of the final good producer can be written as

$$\Pi_t^{FF} = \max \left[\int_0^{\mu_t} [y_t(i)]^{\frac{1}{\lambda}} di \right]^{\lambda} - \int_0^{\mu_t} p_t(i) y_t(i) di, \tag{2}$$

and the first-order optimality condition implies that the demand function for the i^{th} intermediate good is given by

$$p_t(i) = \left[\frac{y_t(i)}{Y_t} \right]^{-\frac{\lambda-1}{\lambda}}.$$

⁹We assume that the household inelastically supplies one unit of labor.

2.2.2 Intermediate Goods Producers

A firm in the intermediate goods sector lives for one period and is profit maximizing. All firms are ex-ante identical. There is a sunk entry cost, κ .¹⁰ Once the entry cost is paid, a firm gains the ability to produce an intermediate good. The firm has monopoly power over the good it produces. Next, the firm draws a productivity parameter $A(j)$, where j is drawn from an i.i.d. uniform distribution over $[0, 1]$. The production function for the good j is given by

$$[A(j)]^{1-\gamma} [k(j)^\alpha n(j)^{1-\alpha}]^\gamma,$$

where $k(j)$ and $n(j)$ denote capital and labor, respectively. The productivity parameter differs among the firms. A firm with a higher index has a higher productivity parameter, i.e., $A(j) > A(i)$ for $j > i$. In addition, function $A(j)$ is assumed to be continuous, and $A(0) = 0$. The parameter $\gamma \in (0, \lambda)$ determines the degree of returns to scale in variable inputs.¹¹ The parameter α is between zero and 1.

If a firm decides to produce, it must incur an operating cost in terms of wages paid to ϕ units of overhead labor. Consider the decision of a firm born in time t with a draw j . If it decides to produce, its profits are

$$\begin{aligned} \pi_t^P(j) &= \max_{k_t(j), n_t(j)} p_t(j) y_t(j) - r_t k_t(j) - w_t [n_t(j) + \phi] \\ \text{s.t. } y_t(j) &= [A(j)]^{1-\gamma} [k_t^\alpha(j) n_t^{1-\alpha}(j)]^\gamma, \quad p_t(j) = \left[\frac{y_t(j)}{Y_t} \right]^{\frac{1-\lambda}{\lambda}}, \end{aligned} \quad (3)$$

where r_t denotes the rental rate on capital. Note that $r_t = R_t - (1 - \delta)$, where δ is the depreciation rate of capital used in production. The decision to produce or not depends on whether $\pi_t^P(j)$ is positive. Therefore, the j^{th} firm's profits, $\pi_t^F(j)$, are given by

$$\pi_t^F(j) = \max\{\pi_t^P(j), 0\}. \quad (4)$$

Free entry implies that, in equilibrium, firms' expected profits must be equal to the entry cost, κ :

$$\int_0^1 \pi_t^F(j) dj = \kappa. \quad (5)$$

¹⁰We assume that κ is denominated in consumption units and that all entry-cost payments are rebated to the households in a lump-sum fashion. Alternatively, one can model κ as a sunk investment, i.e., in units of capital. Such a formulation would not change any of our results, but it would make the exposition more cumbersome.

¹¹This is what Lucas (1978) calls managers' span of control.

2.2.3 Firms' average productivity

We derive the equilibrium relationship between the firms' average productivity and the operating cost. First, we determine the lowest productivity level necessary for a firm to decide to produce. The existence of economy-wide competitive factor markets implies that in equilibrium, the gross profits, capital, and labor ratios of any two firms are equal to their (scaled) productivity ratio:

$$\frac{p_t(j)y_t(j)}{p_t(i)y_t(i)} = \frac{k_t(j)}{k_t(i)} = \frac{n_t(j)}{n_t(i)} = \frac{a(j)}{a(i)}, \quad \forall i, j, \quad (6)$$

where $a(j) \equiv A(j)^{\frac{1-\gamma}{\lambda-\gamma}}$. The first-order conditions of problem (3) imply that profits from producing are equal to the firm's share of the gross profits $(1 - \frac{\gamma}{\lambda})$ minus the operating cost¹²:

$$\pi_t^P(j) = (1 - \frac{\gamma}{\lambda})p_t(j)y_t(j) - \phi w_t.$$

Clearly $\pi_t^P(j)$ is increasing in j and, since $a(j) = 0$, there exists a cutoff firm, J_t , which is indifferent between producing or not:

$$(1 - \frac{\gamma}{\lambda})p_t(J_t)y_t(J_t) = \phi w_t. \quad (7)$$

Firms with indices higher than J_t will produce, and those with lower indices will not. Thus, firms' zero profit condition in (5) can be written as

$$\kappa = \phi w_t \int_{J_t}^1 \left[\frac{a(j)}{a(J_t)} - 1 \right] dj. \quad (8)$$

The previous equation defines the cutoff J_t as a function of the operating cost ϕw_t . An increase in the cutoff J_t has two effects: profits of every firm decline, and the number of producing firms as a fraction of entering firms declines. Therefore, the right-hand side of (8) is decreasing in J_t , while it is clearly increasing in the fixed cost (ϕw_t) . Hence, the cutoff is increasing in the operating cost. Therefore, firms' average productivity, $\bar{a}(J_t) = \frac{\int_{J_t}^1 a(j) dj}{1 - J_t}$, is an increasing function of the operating cost.

¹²Later on, with some abuse of terminology, we will refer to $(1 - \frac{\gamma}{\lambda})p_t(j)y_t(j)$ as firms' gross profits.

2.2.4 Entry and the number of operating firms

Entry in this model refers to the number of firms that pay the entry cost, κ . The number of entering firms differs from the number of operating firms because only a fraction of entrants will have productivity high enough to operate: the pool of producers consists only of firms which have an index higher than J_t . In particular, let ν_t denote the number of entering firms and μ_t the number of operating firms. Then

$$\mu_t = \nu_t \int_{J_t}^1 dj. \quad (9)$$

2.3 Aggregate Output and TFP

Let K_t and N_t denote the total amount of capital and labor used by the firms:

$$K_t = \nu_t \int_{J_t}^1 k_t(j) dj, \quad (10)$$

$$N_t = \nu_t \int_{J_t}^1 [n_t(j) + \phi] dj = u_t N_t + \nu_t (1 - J_t) \phi, \quad (11)$$

where u_t is the fraction of labor used in production. Aggregate output can be written as

$$Y_t = \left[(\mu_t \bar{a}(J_t))^{(\lambda-\gamma)} u_t^{(1-\alpha)\gamma} \right] K_t^{\alpha\gamma} (N_t)^{(1-\alpha)\gamma}. \quad (12)$$

Finally, the rental rate on capital the wage and the equation determining the cutoff J_t can be written as

$$\alpha \frac{\gamma}{\lambda} \frac{Y_t}{K_t} = r_t, \quad (13)$$

$$(1 - \alpha) \frac{\gamma}{\lambda} \frac{Y_t}{u_t N_t} = w_t, \quad (14)$$

$$\left(1 - \frac{\gamma}{\lambda}\right) \frac{a(J_t)}{\bar{a}(J_t)} \frac{Y_t}{(1 - u_t) N_t} = \phi w_t. \quad (15)$$

2.4 Closing the Model

The resource constraint is given by

$$C_t + K_{t+1} = Y_t + (1 - \delta) K_t. \quad (16)$$

The only role the government has in the model is to collect the entry fees $\nu_t \kappa$ from firms and rebate them lump-sum to the households:

$$T_t = \nu_t \kappa. \quad (17)$$

Profits and the labor market clearing condition are

$$\Pi_t = \Pi_t^{FF}, \quad (18)$$

$$N_t = 1. \quad (19)$$

The definition of equilibrium is standard.

3 Steady States, Dynamics, and Some Extensions

In this section we analyze the existence and stability of the steady states, and we discuss some extensions to our basic model. The main finding is that there can be multiple stable steady states for an arbitrarily low degree of increasing returns to scale.

Intuitively, if there are multiple steady states, their existence is due to the endogenous productivity mechanism embedded in the model. Equation (8) relates the cutoff J to the operating cost, ϕw_t . The integral on the right-hand side of this equation is decreasing in J_t . Thus, a higher operating cost translates into a higher cutoff and vice versa. In an economy where the operating cost is high, higher (gross) profits are required to cover this cost. Only high productivity firms can generate such profits. Therefore, the lower productivity firms are forced out from the pool of producers. This can be restated in broader terms: as the operating cost increases, the entry cost relative to operating cost falls, allowing more firms to enter. However, out of these firms, only the ones with higher productivity are profitable enough to operate. This relation between the operating cost and the cutoff provides economic intuition for the existence of multiple steady states. If multiple steady states exist, then one steady state has high capital and only high productivity firms are operating. High capital stock and high productivity imply that the wage rate is high, and so is the operating cost. High operating cost, in turn, justifies why only high productivity firms are operating. Finally, since productivity is high, a high capital stock is necessary to equate the

return on capital to $1/\beta$. Conversely, in a “low” steady state, the capital stock and firms’ average productivity are low, and so is the operating cost, allowing lower productivity firms to operate. Since firms’ average productivity is low, the capital stock must be low to have the return on capital equal to $1/\beta$. A firm productive enough to be active in different steady states produces more in a good steady state than in a bad one, despite a higher wage and the same interest rate. This is optimal because it faces a higher demand for its goods, which offsets the contractionary pressure of higher factor prices.

3.1 Steady States

We present the argument formally in Propositions 1 and 2; we provide proofs in Appendix A. First, note that the number of firms is proportional to the total amount of labor used to cover the fixed cost:

$$\mu_t = \frac{1 - u_t}{\phi} N_t.$$

Therefore, aggregate output is given by

$$Y_t = TFP_t K_t^{\lambda s_k}, \quad (20)$$

where $s_k = \frac{\alpha\gamma}{\lambda}$ denotes the capital share of output, and total factor productivity is

$$TFP_t = \left[\phi^{\gamma-\lambda} [\bar{a}(J_t)]^{\lambda-\gamma} (1 - u_t)^{\lambda-\gamma} u_t^{(1-\alpha)\gamma} \right]. \quad (21)$$

There are two components of TFP: firms’ average productivity $(\bar{a}(J_t))^{(\lambda-\gamma)}$ and the term $u_t^{(1-\alpha)\gamma} (1 - u_t)^{\lambda-\gamma}$, which we call the *labor allocation* component. Firms’ average productivity is increasing in J_t . The labor allocation component is a function of J_t as well, though not necessarily monotonic. However, the effect of J_t on average productivity dominates, and TFP_t is increasing in J_t .

The following proposition allows us to present the model economy in a more familiar, neoclassical framework.

Proposition 1 *The aggregate production function in (12) and the total factor productivity in (21) are increasing in the cutoff J_t . The cutoff J_t , the wage w_t , and the aggregate output Y_t are all increasing functions of capital K_t . The rate of return on capital $R_t \equiv (r_t + 1 - \delta)$ is a function of K_t .*

Proof. See Appendix A. ■

The proposition above implies that the dynamics of the economy can be characterized by the following system of difference equations,

$$\begin{aligned} \left(\frac{c_{t+1}}{c_t}\right)^{1/\sigma} &= \beta R(K_{t+1}), \\ C_t + K_{t+1} &= Y(K_t) + (1 - \delta)K_t, \end{aligned} \quad (22)$$

plus a transversality condition. We now turn to the existence and multiplicity of steady states.

Proposition 2 *The economy characterized by the system in (22) generically has an odd number of steady states. For any $\lambda > 1$, there exist a distribution of productivities, $a(j)$, and a value of κ such that the system (22) has multiple steady-state equilibria.*

Proof. (sketch) Some straightforward manipulations of the first order conditions lead to the following relation between the rate of return on capital and the cutoff J :

$$r_t^{\frac{1-\alpha\gamma}{\alpha\gamma}} \kappa = \eta \cdot \Phi(J_t), \quad (23)$$

where

$$\Phi(J) \equiv \left[\frac{\bar{a}(J)}{\bar{a}(J) + \frac{\lambda-\gamma}{(1-\alpha)\gamma} a(J)} \right]^{\frac{\lambda-1}{1-\alpha\gamma}} a(J)^{\frac{\lambda-\gamma}{1-\alpha\gamma}} \int_J^1 \left[\frac{a(j)}{a(J)} - 1 \right] dj \quad (24)$$

and η is a constant. Since $\Phi(J)$ is continuous and $\Phi(0) = \infty$, $\Phi(1) = 0$, there always exists a J^* which satisfies the equation below:

$$[1/\beta - (1 - \delta)]^{\frac{1-\alpha\gamma}{\alpha\gamma}} \kappa = \eta \Phi(J^*). \quad (25)$$

In order for this equation to have more than one solution, it is necessary that the function $\Phi(J)$ be increasing at some point (see Figure 2). In Appendix A we show that there always exists a function $a(j)$ such that this is the case. With a non-monotone $\Phi(J)$, it is trivial to find a value of κ such that equation (23) has multiple solutions. Note that equation (23) implies that if $\Phi(J)$ is increasing, so is $r(K)$: the necessary condition for the existence of multiple steady states is that for some values of K the return on capital must be increasing. The properties of the function Φ mimic those of the firms' expected profits, i.e., the right-hand side of the zero-profit condition

(8). A necessary condition for existence of multiple steady state is that firms' expected profits are increasing in J . An increase in the cutoff J has two opposite effects. On one hand the wage rate increases, increasing expected profits. On the other hand, a higher J implies a lower value of the integral on the right-hand side of (8). For the expected profits to rise with J , the increase in the wage should dominate the fall in the value of the integral. A sufficient condition for this is that $\partial \bar{a}(J) / \partial J \gg \partial a(J) / \partial J$: A relatively high derivative of the average productivity guarantees a strong positive effect on TFP and the wage rate, while a relatively low derivative of the function $a(J)$ implies a mild negative response of the integral term. A function $a(j)$ which is sufficiently on some interval and increases rapidly for higher values of j has this property. ■

Given Propositions 1 and 2 it is easy to establish that the “high J ” economy has higher capital stock, higher output, higher total factor productivity, and higher average productivity for firms.

3.2 Dynamics

The following proposition characterizes the behavior of the economy around the steady state(s).¹³

Proposition 3 *Steady states with an odd index are saddles. Steady states with an even index can be classified as follows:*

1. *source, if $Y' - \delta > \frac{\sigma CR'}{R}$ and $[(Y' - \delta) - \frac{\sigma CR'}{R}]^2 > 4 \frac{\sigma CR'}{R}$;*
2. *unstable spiral, if $Y' - \delta > \frac{\sigma CR'}{R}$ and $[(Y' - \delta) - \frac{\sigma CR'}{R}]^2 < 4 \frac{\sigma CR'}{R}$;*
3. *sink, if $Y' - \delta < \frac{\sigma CR'}{R}$ and $[(Y' - \delta) - \frac{\sigma CR'}{R}]^2 > 4 \frac{\sigma CR'}{R}$;*
4. *stable spiral, if $Y' - \delta < \frac{\sigma CR'}{R}$ and $[(Y' - \delta) - \frac{\sigma CR'}{R}]^2 < 4 \frac{\sigma CR'}{R}$.*

Proof. See appendix A. ■

For the parameter values we consider in the rest of the paper, we obtain three steady states, with the odd steady state unstable (cases 1 and 2 in Proposition 3). In comparing output and TFP across steady states we will focus on the two stable steady states.

¹³An analysis of the global dynamics of our model is beyond the scope of this paper, and we refer the reader to Galí (1995).

3.3 Extending the Basic Model

In this section we consider two possible extensions of our basic model. First, we consider the implications of endogenizing the entry cost. Then, we analyze infinitely lived firms.

3.3.1 Entry and Operating Costs

The key feature of the model that allows for multiple steady-state equilibria is the asymmetry between the entry and the operating cost. While the operating cost is endogenous and changes with the state of the economy, the entry cost is not. One might try to relax this assumption, and allow both the entry and the operating costs to be endogenous. In this case, multiple steady-state equilibria may exist as long as a weaker form of asymmetry is preserved. In particular, the operating cost should be “more” increasing in capital than the entry cost, so that the ratio of the operating cost to the entry cost is increasing in capital. We suggest a simple example, based on Atkeson and Kehoe (2005). Let the entry cost take the form of κ units of entry services, which firms need to purchase to enter. Let the production function of these services be exactly the same as it is for consumption goods, except that it is more or less labor intensive. Then, it can be shown that in a steady state the zero-profit condition in (8) becomes

$$\kappa = \eta w^\nu \int_J^1 \left[\frac{a(j)}{a(J)} - 1 \right] dj, \quad (26)$$

where η and ν are positive constants. When ν is 1, it is the same zero-profit condition as before. When ν is zero, it is the case of Atkeson and Kehoe (2005). As long as ν is not zero, the key relation between wage w and the cutoff J , which leads to multiple steady states, is preserved.

3.3.2 Infinitely Lived Firms

The firms’ productivity changes over time. We consider two opposite cases:

1. firms’ productivity in every period is given by $A(j)$, where j is the original draw.
2. firms draw a new j which is independent of past draws.

We also assume that firms die with constant probability $(1 - p)$. Consider a period- t decision of a firm born in time s with a draw j . The Bellman equation of this firm is

$$V_t(s, j) = \max \left[\pi_t^P(s, j) + \frac{p}{R_{t+1}} E_t V_{t+1}(s + 1, j') \right], \quad (27)$$

where $\pi_t^P(s, j)$ is the profits from producing as defined in Section 2.2.2, and R_{t+1} is the rate of return on capital (i.e., the interest rate).¹⁴ The law of motion of j is $j' = j$ in case 1 and j' is i.i.d. uniform over $[0, 1]$ in case 2. The free entry condition implies that

$$\kappa = \int_0^1 V_t(t, j) dj. \quad (28)$$

Proposition 2 extends to both cases. If firms' productivity is the same as the original draw, the function Φ is unchanged. For the case of i.i.d. draws, the function Φ is replaced by the following:

$$\tilde{\Phi}(J) \equiv \frac{\left[\frac{\bar{a}(J)}{\bar{a}(J) + \frac{(\lambda-\gamma)}{(1-\alpha)\gamma} \left(1 + \frac{p}{R} \int_J^1 \left[\frac{a(j)}{a(J)} - 1 \right] dj \right) a(J)} \right]^{\frac{\lambda-1}{1-\alpha\gamma}} a(J)^{\frac{\lambda-\gamma}{1-\alpha\gamma}} \int_J^1 \left[\frac{a(j)}{a(J)} - 1 \right] dj}{\left(1 + \frac{p}{R} \int_J^1 \left[\frac{a(j)}{a(J)} - 1 \right] dj \right)^{1 - \frac{(\lambda-\gamma)}{1-\alpha\gamma}}} \quad (29)$$

Notice that when $p = 0$, case 2 simplifies to our baseline model in Section 2, i.e., $\tilde{\Phi}(J)|_{p=0} = \Phi(J)$.

4 Properties of the Model

In this section we discuss some qualitative and quantitative properties of our model. First we consider a growth miracle in the model, driven by technological progress or a decrease in entry costs. We argue that the decline in the share of small firms is consistent with Japan's post-World War II experience and with the Industrial Revolution. Then we compute differences in output and TFP in a calibrated version of the model, and we perform

¹⁴Recall that the firms are owned by the households and there is no aggregate uncertainty in the economies we consider. Therefore, $1/R_{t+1}$ is the relevant discount factor.

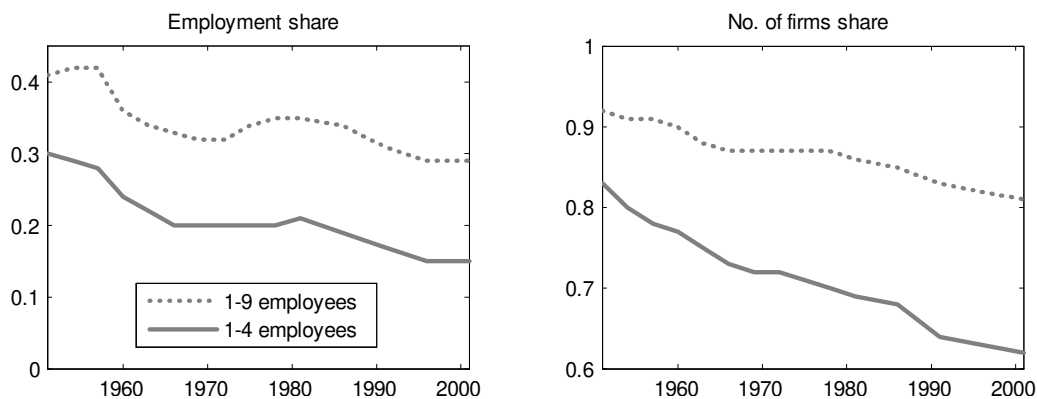


Figure 1: Smallest establishments in Japan: employment and number of firms shares.

sensitivity analysis on the degree of increasing returns to scale and capital share parameters. We conclude that the differences between high and low steady states are sizable.

4.1 Growth Miracles: An Interpretation

A puzzle closely related to cross-country income differences is the question of how and why countries grow and what causes growth miracles. A common view in the literature is that growth miracles are a result of a dramatic shift towards more productive firms and better forms of industrial organization. For example, Mokyr (2001) states that the Industrial Revolution was accompanied by “*the ever-growing physical separation of the unit of consumption (household) from the unit of production (plant),...*” due to “*... concentration of former artisans and domestic workers under one roof (plants), in which workers were more or less continuing what they were doing before, only away from home ...*” and “*... a more radical change in production technique, with substantial investment in fixed capital combined with strict supervision and rigid discipline.*” Thus, plants and factories (i.e., bigger establishments) must have been more productive than “in home” production units (i.e., the smallest establishments), and the Industrial Revolution can be viewed as a shift of resources from smaller, less productive units to larger, more productive ones.

Japan postwar growth miracle is similar in this respect to the Industrial Revolution (see Figure 1): the labor share of the smallest establishments (i.e., establishments with nine employees or fewer) fell by 9 percentage points between 1957 and 1969. The period from 1957 to 1969 was a period of remarkable economic growth, which Parente and Prescott (2005) classify as a period of a growth miracle. Such a shift in our model's framework depends on the properties of the function $a(j)$. If the corresponding probability density function of productivities is one that implies the existence of multiple steady states (i.e., it has a high density somewhere at the lower tail), then a shift from small to the large establishments occurs when the economy moves away from a “low J ” steady state to a “high J ” steady state.

There are two reasons that can cause such a shift. The first one, is a decline in entry barriers, i.e., a decline in the entry cost, κ .

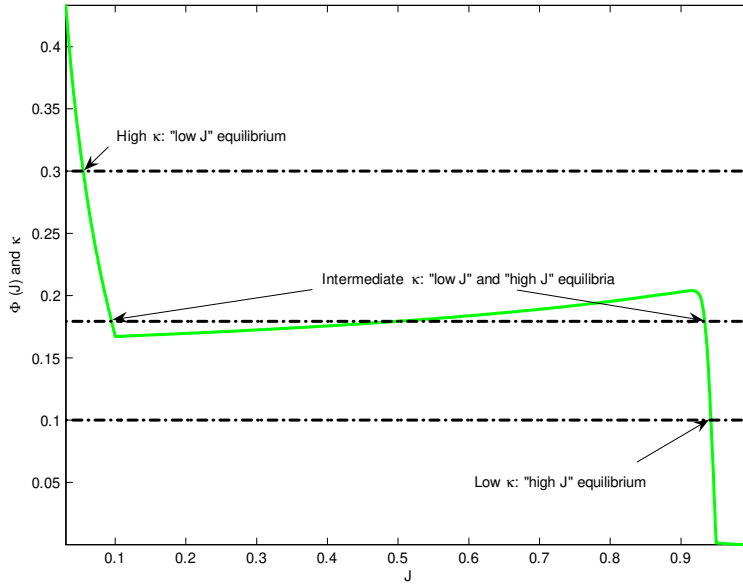


Figure 2: The role of the entry cost, κ .

To illustrate this point, it is useful to start with Figure 2. For larger values of κ , there is a unique, low-cutoff steady state, and for lower κ 's there is a unique steady state, with large J . For intermediate values of κ there can be two steady states. A small change in the value of κ can lead to large differences in J and the corresponding values of capital and output. In our

model economy, the best technologies available are used regardless of the magnitude of the entry cost. The usage of worse technologies, on the other hand, depends on the entry cost. A reduction in the entry cost can cleanse the economy of lower productivity firms, increasing firms' average productivity and TFP. This mechanism of growth miracles shares a common driving force, reduction of barriers, with the one of Parente and Prescott (2000). However, the effect of the reduction of the barriers is different. In their model new, better technologies are not being used because of the barriers. Here, the entry barriers determine not the highest, but the lowest level of technology that is being used in the economy.

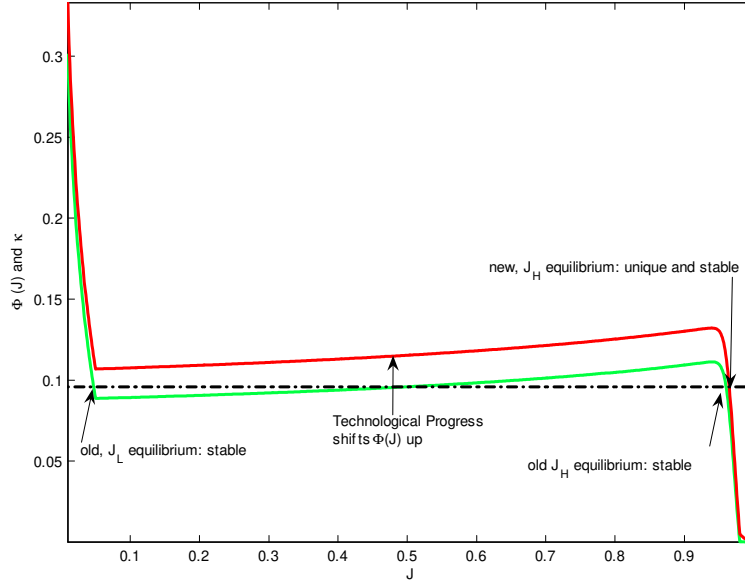


Figure 3: A growth miracle driven by technological progress.

The second reason for a growth miracle is technological progress. A natural way to introduce this into our model is to consider a one-time permanent increase in the function $a(j)$ for values of j close to 1.¹⁵ That is, the best technologies become even better. Mathematically, this can be written, for

¹⁵A better model to address the effect of productivity improvements would be one where the highest level of technology that is available in the economy grows over time. Building and examining such a model is left for future research.

example, as

$$a^{NEW}(j) = \begin{cases} a(\bar{j}) \left[\frac{a^{OLD}(j)}{a(\bar{j})} \right]^q, & \text{if } j \geq \bar{j}, \\ a^{OLD}(j), & \text{otherwise;} \end{cases}$$

where \bar{j} is close to 1, and q is greater than 1. For any $J < \bar{j}$, the change in the function $a(j)$ will cause $\Phi(J)$ to rise. If such a rise is sufficiently large, the “low J ” steady state will disappear (see Figure 3 below), and the economy will start growing toward a “high J ” steady state.

4.2 Model Calibration

Our model contains seven parameters $(\beta, \delta, \lambda, \gamma, \alpha, \phi, \kappa)$, plus any additional parameters determining the function $a(j)$. The model’s implications are robust to the choice of β and δ for the commonly used values of $\beta \in (0.94, 0.99)$ and $\delta \in (0.08, 0.12)$. Therefore, we set $\beta = 0.95$ and $\delta = 0.10$. The parameters λ , γ , and α deserve more consideration.

The first parameter, λ , governs the degree of increasing returns to scale in the economy. There has been a large debate in the recent literature on the magnitude of increasing returns in the economy. While earlier researchers (most notably, Hall, 1988) suggested that there are large increasing returns to scale in the economy, subsequent work has shown that the returns to scale can be best described as constant or at most moderately increasing. The latest estimates of λ are probably those constructed by Laitner and Stolyarov (2004). Their preferred point estimate is $\lambda = 1.1$, with confidence interval $(1.03, 1.2)$. These figures are close to the estimates of Bartelsman, Caballero, and Lyons (1994), Burnside (1996), Burnside, Eichenbaum, and Rebelo (1995), Basu (1996), Basu and Fernald (1997), and Harrison (2003). Hence, we calibrate our model with $\lambda = 1.1$.

The next parameter, γ , represents the share of output that goes to capital and labor used directly in production, for a given value of λ . Note that in the model there is a difference between aggregate returns to scale and firm level returns to scale. While at the aggregate level there are increasing returns to scale, at the firm level, as long as $\gamma < 1$, the returns to scale in *variable* inputs are decreasing. In our model, heterogenous productivity leads to a heterogenous degree of returns to scale in *all* inputs. For firms with higher productivity, the decreasing returns to scale in variable inputs dominate the increasing returns to scale effect of the fixed cost; for the firms with lower productivity, it is the opposite. These observations are broadly consistent

J_1	J_2	b	N_1	N_2	κ
1e-6	0.97	2.13	280.2	104.2	0.50

Table 1: Parameter Values: function $a(j)$ and fixed cost

with empirical findings of Basu (1996), and Basu and Fernald (1997).¹⁶ As a benchmark, we consider $\gamma = 0.85\lambda$, which is the preferred value of Atkeson and Kehoe (2005). This is very close to the estimated value of 0.84 in Basu (1996).

The choice of the next parameter, α , depends on the interpretation of s_k . Interpreted literally, this is the capital share of output. However, if a part of firms' (entrepreneurs') share of output, i.e., $(1 - \gamma/\lambda)$, is interpreted as capital income, then s_k is less than the capital share of output. With this interpretation, one needs to take a stand on how the firm's share of output is divided between capital and labor. A commonly used rule is to split this share so that the capital share of output is α . As a starting point, we set s_k to 0.4. This implies that when γ is set to (0.85λ) , α is equal to 0.47.

We have shown that for some functions $a(j)$ there will be multiple stable steady states. The key property of the function $a(j)$ that generates multiplicity of equilibria is that \bar{a}_J strongly dominates a_J for some J .¹⁷ A function that has this property is one that is sufficiently flat on some interval (J_1, J_2) . The larger this interval is, the farther apart the stable steady states are from each other. In terms of firms' productivity distribution, this translates into the lower steady state having a large number of firms with nearly the same low productivity. Hence, we parameterize the function $a(j)$ as follows:

$$a(j) = \begin{cases} j, & \text{if } j \leq J_1 \\ J_1 + b \left(\frac{j-J_1}{J_2-J_1} \right)^{N_1}, & \text{if } J_1 < j \leq J_2 \\ (J_1 + b) \left(\frac{j}{J_2} \right)^{N_2}, & \text{if } j > J_2. \end{cases} \quad (30)$$

We normalize ϕ to 1,¹⁸ and we choose κ and the five parameters pinning down the productivity distribution (J_1, J_2, N_1, N_2, b) so that the distribution

¹⁶See Kim (2004) for a detailed discussion on different ways of modeling increasing returns to scale.

¹⁷See proof of Proposition 2 in Appendix A.

¹⁸Notice that for our results only $\kappa/\phi^{\frac{\lambda-\gamma}{1-\alpha\gamma}}$ matters (see equations (23) and (42) in Appendix A).

of firms by size implied by our model in the two stable steady state is as close as possible to the distributions of firms by size in the average Least Developed Country (LDC) and in the U.S. (see Tybout, 2000, Table 1).

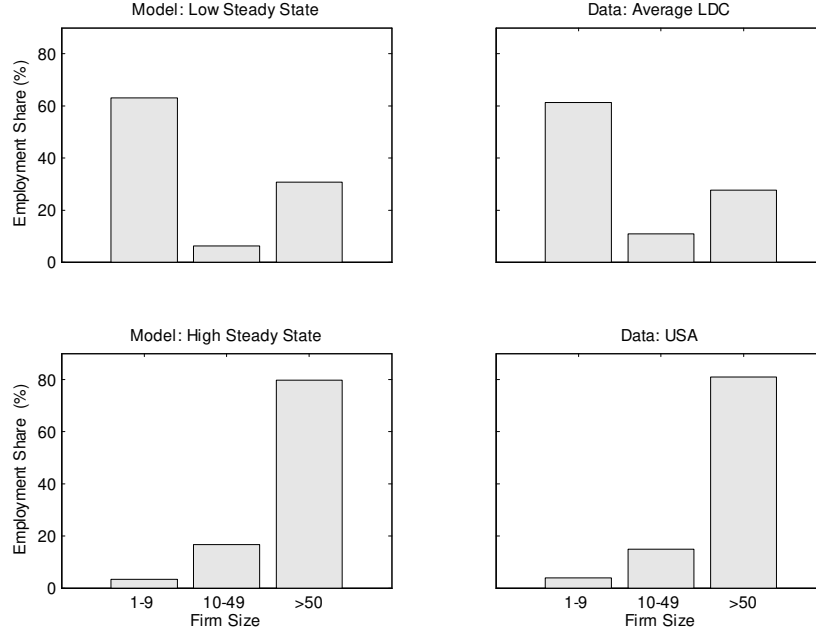


Figure 4: Employment Shares: model versus data.

Figure 4 portrays the distributions of firms by size for the U.S. and the average LDC (right column), together with the distributions for the high and low steady states of our model (left column).

Figure 5 reports the function $a(j)$, which minimizes the distance between the model distributions and their empirical counterparts.

4.3 Sensitivity Analysis

In this section we conduct sensitivity analysis of the baseline calibration by analyzing how varying the degree of increasing returns to scale and the capital share maps into differences across the high and the low steady states of our model.

Tables 2-5 present the ratios of values of output and TFP levels for the two stable steady states for different parameter values. In the first column of

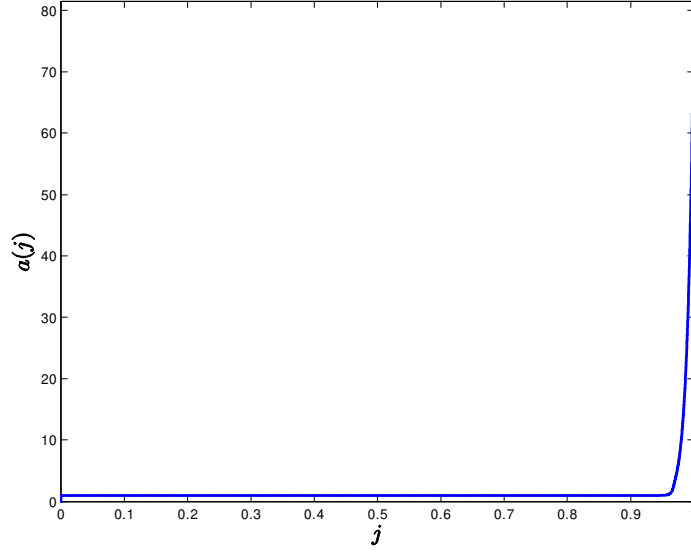


Figure 5: Calibrated function $a(j)$.

Tables 2-5 we report the worst-case scenario of no increasing returns to scale, together with the most favorable function $a(j)$. In the remaining columns we maintain $a(j)$ fixed to the calibrated function discussed above, and we analyze the effect of varying the capital share and the ratio of γ to λ .

		λ						
		1^a	1.01	1.05	1.1	1.15	1.2	1.25
γ	0.95λ	1.05	1.03	1.03	1.03	1.03	1.03	1.04
	0.9λ	1.12	1.06	1.06	1.07	1.07	1.08	1.08
	0.85λ	1.19	1.10	1.10	1.11^b	1.12	1.13	1.14
	0.8λ	1.28	1.14	1.16	1.17	1.18	1.20	1.22

^a Theoretical upper bound for $\lambda \rightarrow 1$. ^b Benchmark calibration.

Table 2: Relative TFP and returns to scale ($s_k = 0.4$)

The limiting case of λ equal to 1 has the least favorable implications for the existence of multiple steady states, because the model essentially collapses to the standard neoclassical model. It is important to see how large the steady state differences can be for λ arbitrarily close to 1. The condition for the existence of multiple steady states translates to $a(J)$ being (almost)

		λ						
		1^a	1.01	1.05	1.1	1.15	1.2	1.25
γ	0.95λ	1.09	1.05	1.05	1.06	1.06	1.07	1.07
	0.9λ	1.20	1.10	1.11	1.12	1.14	1.15	1.17
	0.85λ	1.33	1.17	1.19	1.21 ^b	1.23	1.26	1.30
	0.8λ	1.50	1.25	1.28	1.32	1.37	1.42	1.48

^a Theoretical upper bound for $\lambda \rightarrow 1$. ^b Benchmark calibration.

Table 3: Relative output and returns to scale ($s_k = 0.4$)

		λ						
		1^a	1.01	1.05	1.1	1.15	1.2	1.25
s_k	0.36	1.19	1.10	1.10	1.11	1.12	1.13	1.13
	0.4	1.19	1.10	1.10	1.11 ^b	1.12	1.13	1.14
	0.45	1.19	1.10	1.11	1.12	1.13	1.14	1.15
	0.55	1.20	1.10	1.11	1.12	1.13	1.14	1.16
	0.6	1.20	1.11	1.12	1.13	1.14	1.16	1.18
	0.65	1.22	1.12	1.13	1.15	1.19	1.26	1.45

^a Theoretical upper bound for $\lambda \rightarrow 1$. ^b Benchmark calibration.

Table 4: Relative TFP and capital share ($\gamma = 0.85$)

		λ						
		1^a	1.01	1.05	1.1	1.15	1.2	1.25
s_k	0.36	1.31	1.16	1.17	1.19	1.21	1.23	1.26
	0.4	1.33	1.17	1.19	1.21 ^b	1.23	1.26	1.30
	0.45	1.38	1.19	1.21	1.24	1.28	1.32	1.37
	0.55	1.43	1.22	1.25	1.29	1.34	1.40	1.48
	0.6	1.50	1.26	1.30	1.35	1.43	1.54	1.70
	0.65	1.75	1.39	1.48	1.66	1.99	2.87	7.27

^a Theoretical upper bound for $\lambda \rightarrow 1$. ^b Benchmark calibration.

Table 5: Relative output and capital share ($\gamma = 0.85$)

a constant over some interval. In this case, the extremes of this interval correspond to the two steady-state values of J . This implies that the ratio of total factor productivity, capital, and output levels in the two stable steady

states can be at most:

$$\frac{TFP^{HIGH}}{TFP^{LOW}} \leq \left(\frac{1 - s_k}{\gamma - s_k} \right)^{1 - s_k}, \quad (31)$$

$$\frac{K^{HIGH}}{K^{LOW}} \leq \frac{1 - s_k}{\gamma - s_k}, \quad (32)$$

$$\frac{Y^{HIGH}}{Y^{LOW}} \leq \frac{1 - s_k}{\gamma - s_k}. \quad (33)$$

When our economy approaches constant returns to scale, the endogenous TFP mechanism alone is quite powerful and it can generate differences in TFP and output across steady states of up to 28 and 50 percent, respectively.

In the studies of the long-run behavior of an economy, using the proper measure of capital share of output is of crucial importance. For example, for the unified theory of Parente and Prescott (2005) to be successful, the capital share of output should be between 0.55 and 0.65. The magnitude of this share depends on the definition of investment (capital). In the context of this paper it is proper to define investment as “any allocation of resources that is designed to increase future productivity” (see Parente and Prescott, 2000). That is, investment should include maintenance and repair, research and development, software, investment in organizational capital, and investment in human capital. Parente and Prescott (2000) find that including these items in investment implies that the capital share of output is larger than 1/2 and can reach as high as 2/3.¹⁹

The capital share is important for two reasons. First, there is the standard neoclassical effect: the higher the capital share is, the higher the effect of TFP is on the economy. To see this, note that for two identical economies, differing only in TFP, the steady state capital ratio relates to the TFP ratio as follows:

$$\frac{K^{HIGH}}{K^{LOW}} = \left(\frac{TFP^{HIGH}}{TFP^{LOW}} \right)^{\frac{1}{1 - s_k \cdot \lambda}}.$$

Clearly, the higher the share of capital is, the higher the difference in steady state capital is between the two economies.

¹⁹For details and references see the original paper. A large portion of the unmeasured capital is organization capital. Findings of Atkeson and Kehoe (2005) imply that the value of organizational capital in the US manufacturing sector is *larger* than the value of physical capital.

Second, the capital share directly impacts TFP, because it enters into the definition of TFP in (21) and into the definition of the function $\Phi(J)$ in (24). Because of the highly non-linear nature of TFP and Φ as functions of the cutoff J , it is not possible to derive analytically the effect of an increase in the capital share on the resulting TFP differences across the steady states. However, when λ tends to 1 the theoretical upper bound on these differences gets larger as the capital share grows (see equation (32) above). For all numerical experiments (Table 4) the increase in the capital share of output increases the TFP differences. Combined with the “neoclassical effect” described above, this leads to even larger differences in output and in capital across the steady states (Table 5).

For $s_k = 0.4$ and $\lambda = 1.1$, our baseline calibration, TFP and output differ across steady states by a factor of 1.1 and 1.21, respectively. Differences across steady states increase in λ and s_k . When both s_k and λ are high, the resulting differences in output are large, reaching as much as 627 percent.

5 Price of Capital, Entry Costs, and Distribution Dynamics

A common explanation for why LDC’s are poor is that they have poor institutions which result in high barriers to capital accumulation and entry. These translate in low levels of capital, productivity and output. We proxy barriers to capital accumulation²⁰ and entry²¹ with the relative price of investment goods. Capital goods are more expensive to acquire in poor countries than in rich ones, and this translates into low levels of capital. In our model the effect of differences in the price of capital is magnified, with respect to the neoclassical growth model, because TFP is endogenous. This amplification mechanism helps our model to generate long-run distributions of output, capital-to-output ratio, and TFP across countries similar to their empirical counterparts. We first confirm the multi-modality of these long-run distributions in the data found repeatedly in the literature (Feyrer, 2003; Johnson, 2005). We then compare the empirical distributions with the correspond-

²⁰This is common practice in the literature: see Jones (1994), Restuccia and Urrutia (2001), Restuccia (2004), and Parente and Prescott (1994, 2000, 2005).

²¹Data on more direct measures of entry barriers are available, but only for a limited number of years (see Djankov, La Porta, Lopez-de-Silanes, and Shleifer (2002) and World Bank (2006)).

ing distributions of output, capital-to-output ratio, and TFP across steady states of our calibrated model, where different equilibria are determined by differences in the price of capital goods.

5.1 The Data

In this section we analyze the distribution dynamics of output per capita, TFP, and the capital-to-output ratio, unconditionally and conditionally on the relative price of investment goods.²²

Assume that the distribution of one of the three variables of interest, in logs and relative to the cross-sectional average, evolves according to the following first-order process:

$$f_{t+\tau}(y) = \int_{-\infty}^{+\infty} g(y|x) f_t(x) dx, \quad (34)$$

where f_t denotes the density²³ at time t and g denotes the stochastic kernel relating the time- t and time- $t + \tau$ distributions. The ergodic distribution, f_∞ , solves

$$f_\infty(y) = \int_{-\infty}^{+\infty} g(y|x) f_\infty(x) dx. \quad (35)$$

We estimate f_∞ non-parametrically as follows:

- the joint distribution $g(y, x)$ is estimated by adaptive Epanechnikov kernel smoothing, as described in Silverman (1986, chap. 5);
- the marginal distribution of x is obtained integrating $g(y, x)$:

$$f(x) = \int_{-\infty}^{+\infty} g(y, x) dy;$$

- the conditional distribution $g(y|x)$ is computed as:

$$g(y|x) = \begin{cases} g(y, x) f(x)^{-1} & \text{if } f(x) \neq 0 \\ 0 & \text{otherwise} \end{cases};$$

²²See Appendix B for data sources.

²³We assume that the cross-section distribution can be described in terms of a density. See Quah (2006) or Azariadis and Stachurski (2003) for a discussion of the more general case.

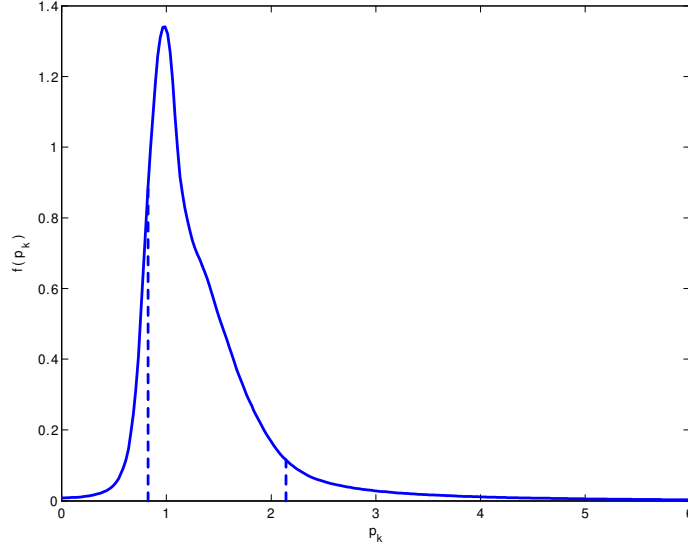


Figure 6: Density of the initial values of p_K . The two dashed lines correspond to the bottom and top decile, respectively.

- the ergodic distribution is computed by discretizing $g(y|x)$, finding the eigenvector associate to the eigenvalue equal to 1, and undoing the discretization.

The ergodic distributions conditional on the initial value of the relative price of capital are constructed in the same way, but considering only the appropriate transitions as determined by the conditioning variable.

We have data on 73 countries for the period 1960-2000,²⁴ and we consider 10-year transitions, i.e., $\tau = 10$. Figure 6 portrays the distribution of the initial values of the relative price of capital goods, p_K . Notice that the distribution has a unique mode, corresponding to $p_K = 0.99$.

The three graphs in the top row of Figure 7 report the ergodic distributions of output, TFP, and the capital-to-output ratio. We report the unconditional ergodic distributions and the ergodic distributions conditional on the initial value of p_K belonging to the top, bottom, and intermediate deciles. The unconditional ergodic distribution of output has three modes, corresponding to levels of output of 0.26, 0.6, and 2.15 times the cross-sectional

²⁴See Appendix B.

average. Multi-modality of the long-run distribution of output across countries is a common finding in the literature, since Quah (1993) introduced distribution dynamics techniques in the empirical growth literature. The unconditional ergodic distributions of TFP and of the capital-to-output ratio are bimodal, with modes at 0.6 and 1.28 and 0.64 and 1.38 times the cross-country average, respectively.

The distributions of output, TFP, and the capital-to-output conditional on the initial value of p_K display similar patterns. To a price of capital in the bottom decile correspond conditional distributions centered around high values; to a high p_K are associated low values of the variables of interest.

5.2 The Model

Introducing the price of capital as an exogenous variable in our model has a direct effect on TFP:

$$\bar{\kappa}p_{K,t} = \phi w_t \int_{J_t}^1 \left[\frac{a(j)}{a(J_t)} - 1 \right] dj, \quad (36)$$

where the entry cost $\kappa_t = (\bar{\kappa}p_{K,t})$ has a fixed component, $\bar{\kappa}$, and responds one-for-one to changes in the price of capital, $p_{K,t}$.

The equation determining the productivity cutoff is modified as follows:

$$[1/\beta - (1 - \delta)]^{\frac{1-\alpha\gamma}{\alpha\gamma}} \bar{\kappa}p_K^{\frac{1}{\alpha\gamma}} = \eta\Phi(J). \quad (37)$$

Notice that a 1 percent change in p_K corresponds to a $1/\alpha\gamma > 1$ percent change in the left-hand side. Consider two different values of the price of capital, say p_K^{HIGH} and p_K^{LOW} , to which correspond two different steady states.²⁵ The ratio of output levels for the two steady states is

$$\frac{Y^{HIGH}}{Y^{LOW}} = \left(\frac{p_K^{HIGH}}{p_K^{LOW}} \right)^{-\frac{s_k\lambda}{1-s_k\lambda}} \left(\frac{TFP^{HIGH}}{TFP^{LOW}} \right)^{\frac{1}{1-s_k\lambda}}.$$

The first term captures the negative effect of a high price of investment goods on capital accumulation and output. This effect is identical to the one arising in the standard neoclassical growth model. The second term captures the indirect effect of the price of capital through our endogenous TFP mechanism: a high p_K leads to low TFP and output. This effect would be present even for values of p_K for which there are multiple equilibria.

²⁵The two different steady states correspond to the high J and low J in Figure 2.

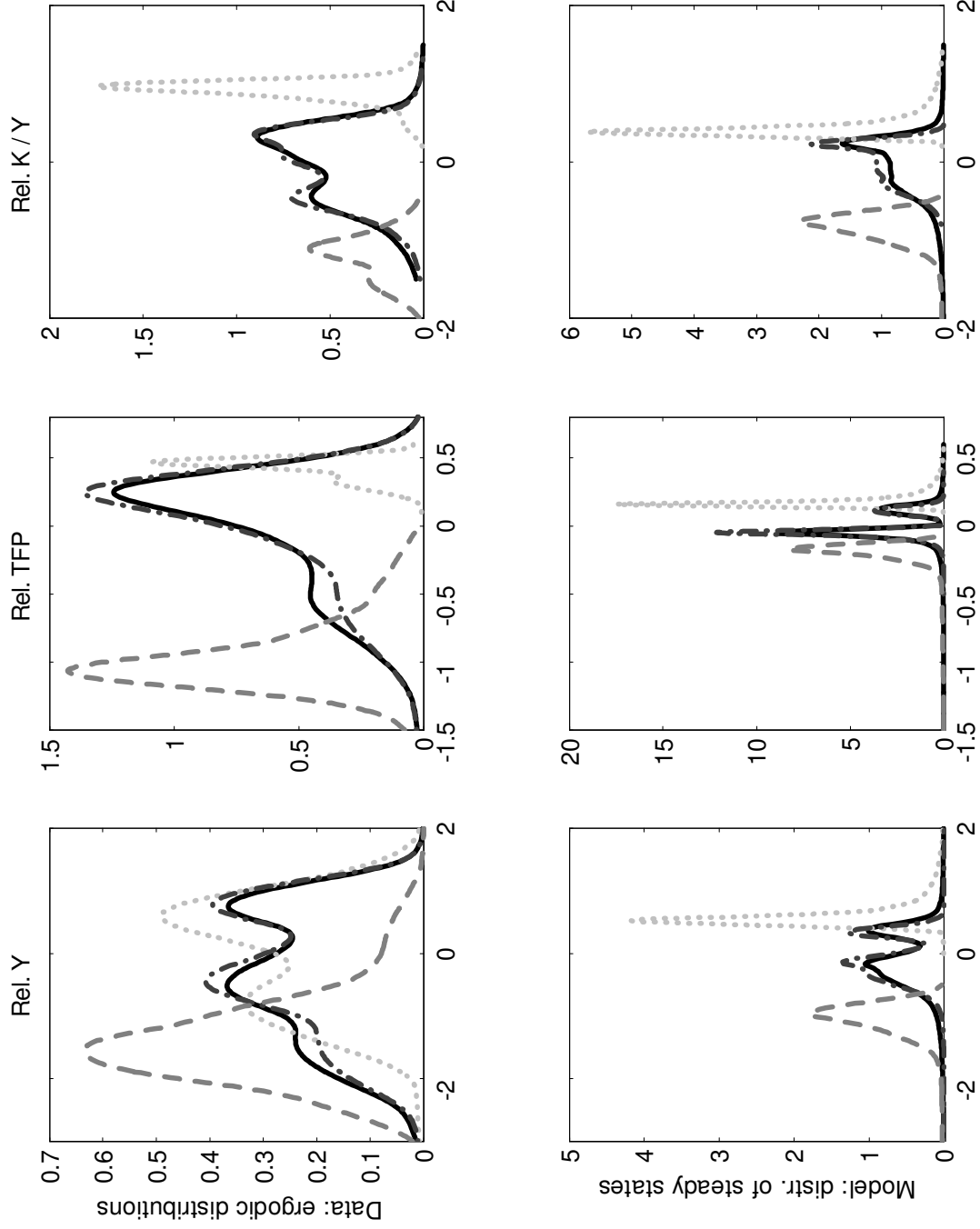


Figure 7: Output, TFP, and capital-to-output ratio: unconditional (black, solid), conditional on $F(p_K) \leq 0.1$ (gray, dashed), conditional on $F(p_K) \in (0.1; 0.9]$ (dark gray, dash-dotted), and conditional on $F(p_K) > 0.9$ (light gray, dotted).

We construct steady-state values of output, TFP, and capital-to-output ratio from our model corresponding to different values of p_K observed in the data, keeping all the other parameters fixed at the benchmark values discussed in Section 4. In the bottom row of Figure 7 we present univariate adaptive kernel-smoothed distributions of steady-state values of Y , K/Y , and TFP in log-deviation from the average.

The unconditional distribution of output is bimodal, with modes at 0.86 and 1.44 times the cross-sectional average. The unconditional distributions of TFP and of the capital-to-output ratio are bimodal as well, with modes at 0.95 and 1.13 and 0.80 and 1.27 times the cross-country average, respectively. Conditioning on a low (high) initial price of capital, all three distributions are unimodal and centered at high (low) values, in a way similar to the corresponding ergodic distributions.

The distributions across the model's steady states are more concentrated than their empirical counterparts. This is due to our model being quite stylized and to the fact that, by comparing steady states, we do not consider the model's transition dynamics. Also, the variation in the price of capital understates the heterogeneity in entry barriers across countries.

The bi-modality of the model's distributions is primarily due to a threshold effect in the price of investment goods. To visualize this, consider Figure 2. Countries with a high price of capital cluster around the low- J steady state,²⁶ while countries with a low price of capital concentrate around the high- J steady state. A high price of investment goods discourages capital accumulation, lowers wages and the operating cost. This leads to a lower level of the productivity cutoff and TFP. When the density of firms with productivity draws around the cutoff level is high, even small variations in the price of capital lead to significant differences in firms' average productivity and TFP across steady states. Economies with a price of capital goods above the threshold level will have significantly lower TFP and output than economies whose price of investment goods is below the threshold.²⁷

²⁶High κ in Figure 2.

²⁷Our calibration implies that the range of the price of capital for which we obtain multiple steady states is relatively small, $p_K \in [0.99, 1.01]$. Hence, our results are robust to which equilibrium is selected in this region.

6 Conclusions

Recent empirical studies attribute a sizable fraction of cross-country income differences to differences in TFP. These differences reflect, in part, the fact that the fraction of low productivity firms in less developed countries is much higher than in industrialized countries. We introduce heterogeneity in productivity across firms in an otherwise standard model. In our model differences in TFP arise endogenously, and we obtain multiple steady-state equilibria for an arbitrarily small degree of increasing returns to scale. If an economy is in a good steady state, only the most productive firms operate, leading to high TFP, capital, and output. In an economy locked in a poverty trap the pool of producers is sullied by low productivity firms, with low TFP, capital and output.

We analyze the qualitative properties of our model by studying a growth miracle. In our model a growth miracle can be induced by technological progress or by a decline in entry barriers and it is accompanied by a shift of employment from small to large firms. This is consistent with the Industrial Revolution and postwar Japan growth experiences.

We calibrate our model using standard parameter values and a distribution of productivity across firms which matches the distribution of firms by size across developed and LDC countries. We feed to our calibrated model the empirical distribution of the price of capital, which determines the level of entry costs. The resulting distributions of output, capital-to-output ratio, and TFP across steady states, unconditional and conditional on the initial value of the price of capital, are qualitatively similar to the ergodic distributions estimated in the data.

References

- ATKESON, A., AND P. J. KEHOE (2005): “Modeling and Measuring Organizational Capital,” *Journal of Political Economy*, 113(5), 1026–53.
- AZARIADIS, C., AND J. STACHURSKI (2003): “A Forward Projection of the Cross-Country Income Distribution,” Discussion Paper No. 570, Institute of Economic Research, Kyoto University.
- (2005): “Poverty Traps,” in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf. Elsevier Science, North-Holland, Amsterdam, New York and Oxford.
- BANERJEE, A., AND E. DUFLO (2005): “Growth Theory Through the Lens of Development Economics,” in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf. Elsevier Science, North-Holland, Amsterdam, New York and Oxford.
- BARSEGHYAN, L. (2006): “Entry Costs and Cross-Country Differences in Productivity and Output,” unpublished manuscript, Cornell University.
- BARTELSMAN, E. J., R. J. CABALLERO, AND R. K. LYONS (1994): “Costomer- and Supplier-Driven Externalities,” *The American Economic Review*, 84(4), 1075–84.
- BASU, S. (1996): “Procyclical Productivity: Increasing Returns or Cyclical Utilization?,” *The Quarterly Journal of Economics*, 111(3), 719–751.
- BASU, S., AND J. G. FERNALD (1997): “Returns to Scale in U.S. Production: Estimates and Implications,” *Journal of Political Economy*, 105(2), 249–283.
- BENHABIB, J., AND R. E. A. FARMER (1994): “Indeterminacy and Increasing Returns to Scale,” *Journal of Economic Theory*, 63(1), 19–41.
- BURNSIDE, C. (1996): “Production Function Regressions, Returns to Scale, and Externalities,” *Journal of Monetary Economics*, 37(2), 177–201.
- BURNSIDE, C., M. EICHENBAUM, AND S. REBELO (1995): “Capital Utilization and Returns to Scale,” in *NBER Macroeconomics Annual 1995*, ed. by B. S. Bernanke, and J. J. Rotemberg, pp. 67–110. MIT Press, Cambridge and London.

- CASELLI, F. (2005): “Accounting for Cross-Country Income Differences,” in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf. Elsevier Science, North-Holland, Amsterdam, New York and Oxford.
- CICCONE, A., AND K. MATSUYAMA (1996): “Start-Up Costs and Pecuniary Externalities as Barriers to Economic Development,” *Journal of Development Economics*, 49(1), 33–59.
- COMIN, D., AND B. HOBIJN (2004): “Neoclassical Growth and the Adoption of Technologies,” NBER Working Paper No. 10733.
- DE SOTO, H. (1989): *The Other Path: The Invisible Revolution in the Third World*. Harper and Row, New York.
- DJANKOV, S., R. LA PORTA, F. LOPEZ-DE-SILANES, AND A. SHLEIFER (2002): “The Regulation of Entry,” *The Quarterly Journal of Economics*, 117(1), 1–37.
- DOMOWITZ, I., R. G. HUBBARD, AND B. C. PETERSEN (1988): “Market Structure and Cyclical Fluctuations in U.S. Manufacturing,” *Review of Economics and Statistics*, 70(1), 55–66.
- DURLAUF, S. N., AND P. A. JOHNSON (1995): “Multiple Regimes and Cross-Country Growth Behavior,” *Journal of Applied Econometrics*, 10(4), 365–84.
- DURLAUF, S. N., AND D. T. QUAH (1999): “The New Empirics of Economic Growth,” in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1A, pp. 235–308. Elsevier Science, North-Holland, Amsterdam; New York and Oxford.
- EASTERLY, W. (1993): “How Much Do Distortions Affect Growth?,” *Journal of Monetary Economics*, 32(2), 187–212.
- EROSA, A., AND A. HIDALGO CABRILLANA (2005): “On Capital Market Imperfections as a Source of Low TFP and Economic Rents,” unpublished manuscript, Universidad Carlos III.
- FEYRER, J. (2003): “Convergence by Parts,” unpublished manuscript, Dartmouth College.

- GALÍ, J. (1995): “Product Diversity, Endogenous Markups, and Development Traps,” *Journal of Monetary Economics*, 36(1), 39–63.
- HALL, R. E. (1988): “The Relation Between Price and Marginal Cost in U.S. Industry,” *Journal of Political Economy*, 96(5), 921–47.
- HALL, R. E., AND C. I. JONES (1999): “Why Do Some Countries Produce So Much More Output Per Worker Than Others?,” *Quarterly Journal of Economics*, 114(1), 83–116.
- HARRISON, S. G. (2003): “Returns to Scale and Externalities in the Consumption and Investment Sectors,” *Review of Economic Dynamics*, 6(4), 963–76.
- HERRENDORF, B., AND A. TEIXEIRA (2005a): “Barriers to Entry and Development,” unpublished manuscript, Arizona State University.
- (2005b): “How Barriers to International Trade Affect TFP,” *Review of Economic Dynamics*, 8(4), 866–76.
- HOPENHAYN, H. A. (1992): “Entry, Exit, and Firm Dynamics in Long Run Equilibrium,” *Econometrica*, 60(5), 1127–1150.
- JOHNSON, P. A. (2005): “A Continuous State Space Approach to “Convergence by Parts”,” *Economics Letters*, 86(3), 317–21.
- JONES, C. I. (1994): “Economic Growth and the Relative Price of Capital,” *Journal of Monetary Economics*, 34, 359–382.
- JOVANOVIĆ, B. (1982): “Selection and the Evolution of Industry,” *Econometrica*, 50(3), 649–670.
- KIM, J. (2004): “What Determines Aggregate Returns to Scale?,” *Journal of Economic Dynamics and Control*, 28(8), 1577–94.
- KLENOW, P. J., AND A. RODRIGUEZ-CLARE (1997): “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?,” in *NBER Macroeconomics Annual 1997*, ed. by B. S. Bernanke, and J. J. Rotemberg, pp. 73–103. MIT Press, Cambridge and London.

- (2005): “Externalities and Growth,” in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf. Elsevier Science, North-Holland, Amsterdam, New York and Oxford.
- KLETTE, T. J., AND S. S. KORTUM (2004): “Innovating Firms and Aggregate Innovation,” *Journal of Political Economy*, 112(5), 986–1018.
- LAITNER, J., AND D. STOLYAROV (2004): “Aggregate Returns to Scale and Embodied Technical Change: Theory and Measurement Using Stock Market Data,” *Journal of Monetary Economics*, 51(1), 191–233.
- LUCAS, JR., R. E. (1978): “On the Size Distribution of Business Firms,” *The Bell Journal of Economics*, 9(2), 508–523.
- MATSUYAMA, K. (2005): “Poverty Trap,” in *New Palgrave Dictionary of Economics*, ed. by L. Blume, and S. Durlauf. Palgrave Macmillan, London and New York, 2nd edn.
- MCKINSEY GLOBAL INSTITUTE (2001): “India: The Growth Imperative,” available at <http://www.mckinsey.com/mgi/reports/pdfs/india/India.pdf>.
- MOKYR, J. (1990): *The Lever of Riches*. Oxford University Press, New York, New York.
- (2001): “The Rise and Fall of the Factory System: Technology, Firms, and Households Since the Industrial Revolution,” *Carnegie-Rochester Conference Series on Public Policy*, 55, 1–45.
- MURPHY, K. M., A. SHLEIFER, AND R. W. VISHNY (1989): “Industrialization and the Big Push,” *Journal of Political Economy*, 97(5), 1003–26.
- PARENTE, S. L., AND E. C. PRESCOTT (1994): “Barriers to Technology Adoption and Development,” *Journal of Political Economy*, 102, 298–321.
- (2000): *Barriers to Riches*. MIT Press, Cambridge and London.
- (2005): “A Unified Theory of the Evolution of International Income Levels,” in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf. Elsevier Science, North-Holland, Amsterdam, New York and Oxford.

- QUAH, D. (1993): “Empirical Cross-Section Dynamics in Economic Growth,” *European Economic Review*, 37(2-3), 426–34.
- (2006): “Growth and Distribution,” unpublished manuscript, London School of Economics.
- RAMEY, V. A. (1991): “Comment on "Markups and the Business Cycle" by Michael Woodford and Julio Rotemberg,” in *NBER Macroeconomics Annual 1991*, ed. by O. J. Blanchard, and S. Fisher. MIT Press, Cambridge and London.
- RAMEY, V. A., AND M. D. SHAPIRO (2001): “Displaced Capital: A Study of Aerospace Plant Closings,” *Journal of Political Economy*, 109(5), 958–992.
- RESTUCCIA, D. (2004): “Barriers to Capital Accumulation and Aggregate Total Factor Productivity,” *International Economic Review*, 45(1), 225–38.
- RESTUCCIA, D., AND R. ROGERSON (2003): “Policy Distortions and Aggregate Productivity with Heterogenous Plants,” unpublished manuscript, University of Toronto.
- RESTUCCIA, D., AND C. URRUTIA (2001): “Relative Prices and Investment Rates,” *Journal of Monetary Economics*, 47(1), 93–121.
- SILVERMAN, B. W. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.
- STATISTICAL RESEARCH AND TRAINING INSTITUTE (various years): *Japan Statistical Yearbook*. Ministry of Internal Affairs and Communications, Tokyo, Japan.
- TAN, C. M. (2005): “No One True Path: Uncovering the Interplay Between Geography, Institutions, and Fractionalization in Economic Development,” Tufts University Economics Working Paper No. 2005-12.
- TYBOUT, J. R. (2000): “Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?,” *Journal of Economic Literature*, 38(1), 11–44.
- WORLD BANK (2006): *Doing Business 2007: How to Reform*. the World Bank and the International Finance Corporation, Washington, D.C.

A Proofs of Propositions

A.1 Proof of Proposition 1

Equations (14) and (15) imply that the fraction of labor used in production u_t is a function only of the cutoff J_t :

$$u_t = \frac{\bar{a}(J_t)}{\bar{a}(J_t) + \frac{\lambda-\gamma}{(1-\alpha)\gamma}a(J_t)}.$$

Substituting this expression of u_t into equation (21), we obtain:

$$\begin{aligned} TFP(J_t) &= \phi^{\gamma-\lambda} \left[\frac{\frac{\lambda-\gamma}{(1-\alpha)\gamma}a(J_t)}{\bar{a}(J_t) + \frac{\lambda-\gamma}{(1-\alpha)\gamma}a(J_t)} \right]^{\lambda-\gamma} \times \\ &\times \left[\frac{\bar{a}(J_t)}{\bar{a}(J_t) + \frac{\lambda-\gamma}{(1-\alpha)\gamma}a(J_t)} \right]^{(1-\alpha)\gamma} [\bar{a}(J_t)]^{\lambda-\gamma} \end{aligned} \quad (38)$$

Differentiating the previous expression,

$$\text{signum}(TFP_J) = \text{signum} \left[\frac{\frac{(\lambda-\gamma)(1-\alpha)\gamma}{(1-\alpha)\gamma\bar{a} + (\lambda-\gamma)a} \frac{a_J}{a} (\bar{a} - a) +}{\bar{a}_J \left(\frac{1}{\bar{a}} - \frac{1}{\bar{a} + \frac{\lambda-\gamma}{(1-\alpha)\gamma}a} \right)} \right], \quad (39)$$

where

$$TFP_J = \frac{\partial TFP(J)}{\partial J}, \quad a_J = \frac{\partial a(J)}{\partial J}, \quad \bar{a}_J = \frac{\partial \bar{a}(J)}{\partial J}.$$

The terms in parenthesis in (39) are positive and they are multiplied by positive terms. Hence, $TFP_J > 0$.

Using the firms' first-order condition in (14) and the zero profit condition in (8) we get that the following relation between the cutoff J_t and capital K_t :

$$\begin{aligned} \kappa &= (1-\alpha) \frac{\gamma}{\lambda} \left[\frac{\lambda-\gamma}{(1-\alpha)\gamma} \right]^{\lambda-\gamma} \left[\frac{\bar{a}(J_t)}{\bar{a}(J_t) + \frac{\lambda-\gamma}{(1-\alpha)\gamma}a(J_t)} \right]^{(1-\alpha)\gamma + (\lambda-\gamma) - 1} \\ &\quad [a(J_t)]^{\lambda-\gamma} K_t^{\alpha\gamma} \left[\frac{1}{a(J_t)} \int_{J_t}^1 a(j) dj - (1 - J_t) \right] \end{aligned}$$

For a given K_t the left-hand side of this equation varies with J_t from $+\infty$ to zero. Moreover, one can easily show that the left-hand side is decreasing in J_t . Thus, there exists a unique J_t which solves the equation. In addition, it is increasing in K_t . Because J_t is increasing in K_t , so is output Y_t and wage w_t . In addition, since, for a given K_t , output Y_t is uniquely determined, so is the R_t ; i.e., R_t is a function of K_t . ■

A.2 Proof of Proposition 2

Use equations (13) and (20) to express K_t as a function of r_t and J_t . By substituting this expression of K_t into equation (38) and by using equation (14), we get

$$r_t^{\frac{1-\alpha\gamma}{\alpha\gamma}} \kappa = \eta \cdot \Phi(J_t) \quad (40)$$

where

$$\Phi(J) \equiv \left[\frac{\bar{a}(J)}{\bar{a}(J) + \frac{\lambda-\gamma}{(1-\alpha)\gamma} a(J)} \right]^{\frac{\lambda-1}{1-\alpha\gamma}} a(J)^{\frac{\lambda-\gamma}{1-\alpha\gamma}} \int_J^1 \left[\frac{a(j)}{a(J)} - 1 \right] dj, \quad (41)$$

and η is a constant:

$$\eta = \phi^{\frac{\gamma-\lambda}{1-\alpha\gamma}} \left(1 - \frac{\gamma}{\lambda} \right) \left(\frac{\alpha\gamma}{\lambda} \right)^{\frac{\alpha\gamma}{1-\alpha\gamma}} \left[\frac{\lambda-\gamma}{(1-\alpha)\gamma} \right]^{\frac{\lambda-\gamma}{1-\alpha\gamma}-1}. \quad (42)$$

Since $\Phi(J)$ is continuous and $\Phi(0) = \infty$, $\Phi(1) = 0$, there always exists a J^* that satisfies the equation below:

$$[1/\beta - (1-\delta)]^{\frac{1-\alpha\gamma}{\alpha\gamma}} \kappa = \eta \cdot \Phi(J^*). \quad (43)$$

We now have to show that for any J^* satisfying equation (43) there exists a pair (c^*, K^*) , both positive, such that $R(K^*) = 1/\beta$ and $c^* = Y(K^*) - \delta K^*$. This is an immediate consequence of proposition 1.

If there is more than one J^* satisfying equation (43), then there will be multiple steady states. Note that for given parameters λ, γ , and α , the shape of the function $\Phi(J)$ is entirely determined by the shape of function $a(j)$: If $a(j)$ is such that $\Phi_J > 0$ then (43) can have multiple solutions. To conclude the proof, we must show that there exists a function $a(j)$ such that $\Phi_J > 0$.

The sign of Φ_J can be checked as follows:

$$\text{signum}(\Phi_J) = \text{signum} \left\{ \begin{aligned} & \frac{\lambda-1}{1-\alpha\gamma} \left[\frac{\bar{a}_J}{\bar{a}} - \frac{\bar{a}_J + \frac{\lambda-\gamma}{(1-\alpha)\gamma} a_J}{\bar{a} + \frac{\lambda-\gamma}{(1-\alpha)\gamma} a} \right] \\ & + \left[\frac{\lambda-\gamma}{1-\alpha\gamma} - \frac{1-J}{\int_J^1 (a(j)-a(J))dj} \right] \frac{a_J}{a} \end{aligned} \right\}.$$

Consider a function

$$a(j) = \begin{cases} j, & \text{if } j \leq J_1 \\ J_1 + b_1 \left(\frac{j-J_1}{J_2-J_1} \right)^{N_2}, & \text{if } J_1 < j \leq J_2 \\ b_2 \left(\frac{j}{J_2} \right)^{N_2}, & \text{if } j > J_2. \end{cases} \quad (44)$$

where $b_i, N_i > 0$. The constant and b_2 must satisfy the following restrictions to guarantee continuity:

$$b_2 = (J_1 + b_1).$$

Taking limits for $N_1 \rightarrow \infty$:

$$\begin{aligned} \lim_{N_1 \rightarrow \infty} a(J) &= J_1 \\ \lim_{N_1 \rightarrow \infty} a_J(J) &= 0 \\ \lim_{N_1 \rightarrow \infty} \bar{a}(J) &= \frac{1}{1-J} \left[(J_2 - J) J_1 + \frac{J_1 + b_1}{1 + N_2} \left(\frac{1}{(J_2)^{N_2}} - J_2 \right) \right] > J_1 > 0 \\ \lim_{N_1 \rightarrow \infty} \bar{a}_J(J) &= \frac{\lim_{N_1 \rightarrow \infty} \bar{a}(J) - J_1}{(1-J)} > 0 \end{aligned}$$

Therefore, as long as $\lambda > 1$, $\lim_{N_1 \rightarrow \infty} \Phi_J > 0$. It follows that there exists a finite N_1 for which $\Phi_J > 0$. If $\Phi_J > 0$, then Φ will have at least one local minimum and one local maximum, say $\underline{\Phi}$ and $\bar{\Phi}$. For any $\kappa \in \left([1/\beta - (1-\delta)]^{\frac{\alpha\gamma-1}{\alpha\gamma}} \eta \underline{\Phi}, [1/\beta - (1-\delta)]^{\frac{\alpha\gamma-1}{\alpha\gamma}} \eta \bar{\Phi} \right)$ our model has multiple steady-state equilibria. ■

A.3 Proof of Proposition 3

Linearizing (22) about a steady state:

$$\begin{bmatrix} \hat{K}_{t+1} \\ \hat{C}_{t+1} \end{bmatrix} = \begin{bmatrix} Y' + 1 - \delta & -1 \\ \frac{\sigma R'}{R} C (Y' + 1 - \delta) & 1 - \frac{\sigma R'}{R} C \end{bmatrix} \begin{bmatrix} \hat{K}_t \\ \hat{C}_t \end{bmatrix}$$

Condition	Steady state stability
$R' < 0$	saddle
$Y' - \delta > \frac{\sigma CR'}{R} > 0$ $\left[(Y' - \delta) - \frac{\sigma CR'}{R}\right]^2 > 4\frac{\sigma CR'}{R}$	source
$Y' - \delta > \frac{\sigma CR'}{R} > 0$ $\left[(Y' - \delta) - \frac{\sigma CR'}{R}\right]^2 < 4\frac{\sigma CR'}{R}$	unstable spiral
$\frac{\sigma CR'}{R} > Y' - \delta > 0$ $\left[(Y' - \delta) - \frac{\sigma CR'}{R}\right]^2 > 4\frac{\sigma CR'}{R}$	sink
$\frac{\sigma CR'}{R} > Y' - \delta > 0$ $\left[(Y' - \delta) - \frac{\sigma CR'}{R}\right]^2 < 4\frac{\sigma CR'}{R}$	stable spiral

Table A.1: Steady-state stability for different parameters configurations

The eigenvalues of the transition matrix are given by:

$$\xi_{1,2} = 1 + \frac{(Y' - \delta) - \frac{\sigma CR'}{R} \pm \sqrt{\left[(Y' - \delta) - \frac{\sigma CR'}{R}\right]^2 - 4\frac{\sigma CR'}{R}}}{2}.$$

If $R' < 0$ (odd steady states) both eigenvalues are real and $\xi_1 < 1 < \xi_2$.
If $R' > 0$ (even steady states) there are four possible cases:

1. $Y' - \delta > \frac{\sigma CR'}{R} \wedge \left[(Y' - \delta) - \frac{\sigma CR'}{R}\right]^2 > 4\frac{\sigma CR'}{R} \Rightarrow \xi_{1,2} \in \mathbb{R}, \|\xi_{1,2}\| > 1;$
2. $Y' - \delta > \frac{\sigma CR'}{R} \wedge \left[(Y' - \delta) - \frac{\sigma CR'}{R}\right]^2 < 4\frac{\sigma CR'}{R} \Rightarrow \xi_{1,2} \in \mathbb{C}, \|\xi_{1,2}\| > 1;$
3. $Y' - \delta < \frac{\sigma CR'}{R} \wedge \left[(Y' - \delta) - \frac{\sigma CR'}{R}\right]^2 > 4\frac{\sigma CR'}{R} \Rightarrow \xi_{1,2} \in \mathbb{R}, \|\xi_{1,2}\| < 1;$
4. $Y' - \delta < \frac{\sigma CR'}{R} \wedge \left[(Y' - \delta) - \frac{\sigma CR'}{R}\right]^2 < 4\frac{\sigma CR'}{R} \Rightarrow \xi_{1,2} \in \mathbb{C}, \|\xi_{1,2}\| < 1.$

■

B Data Sources

The data used in the paper are available from the following sources:

1. Data on the establishments in Japan: *Japan Statistical Yearbook*, edited by the Statistical Training Institute and published by the Statistics Bureau, both under the Ministry of Internal Affairs and Communications, various issues. Data is available every three years for the period 1951-1981 and every five years subsequently. The data are for establishments. An establishment is a single physical location where the business is conducted or where services or industrial operations are performed.
2. Labor share of smallest manufacturing plants in the US and LDC's: Tybout (2000).
3. Data for the distribution dynamics analysis in Section 5: Klenow and Rodriguez-Clare (2005). We consider the 73 countries listed in Table B.1, which have complete data for the period 1960-2000. The price of capital is measured as the ratio of the price levels of investment and GDP.

Code	Country	Code	Country	Code	Country
ARG	Argentina	IDN	Indonesia	PAK	Pakistan
AUS	Australia	IND	India	PAN	Panama
BEL	Belgium	IRL	Ireland	PER	Peru
BGD	Bangladesh	IRN	Iran	PHL	Philippines
BOL	Bolivia	ISL	Iceland	PRT	Portugal
BRA	Brazil	ISR	Israel	PRY	Paraguay
BRB	Barbados	ITA	Italy	ROM	Romania
CAN	Canada	JAM	Jamaica	SEN	Senegal
CHE	Switzerland	JOR	Jordan	SLV	El Salvador
CHL	Chile	JPN	Japan	SWE	Sweden
CMR	Cameroon	KEN	Kenya	SYR	Syria
COL	Colombia	KOR	Korea, Republic of	TGO	Togo
CRI	Costa Rica	LKA	Sri Lanka	THA	Thailand
DNK	Denmark	LSO	Lesotho	TTO	Trinidad & Tobago
DOM	Dominican Republic	MEX	Mexico	TUR	Turkey
ECU	Ecuador	MLI	Mali	TZA	Tanzania
ESP	Spain	MOZ	Mozambique	UGA	Uganda
FIN	Finland	MUS	Mauritius	URY	Uruguay
FRA	France	MWI	Malawi	USA	USA
GBR	United Kingdom	MYS	Malaysia	VEN	Venezuela
GHA	Ghana	NER	Niger	ZAF	South Afriva
GRC	Greece	NLD	Netherlands	ZAM	Zambia
GTM	Guatemala	NOR	Norway	ZWE	Zimbabwe
HKG	Hong Kong	NPL	Nepal		
HND	Honduras	NZL	New Zealand		

Table B.1: List of countries used in the distribution dynamics analysis.