DISTANCE TO FRONTIER, SELECTION, AND ECONOMIC GROWTH

Daron Acemoglu
Massachusetts Institute of Technology

Philippe Aghion
Harvard University

Fabrizio Zilibotti
Institute for International Economic Studies

Abstract
We analyze an economy where firms undertake both innovation and adoption of technologies from the world technology frontier. The selection of high-skill managers and firms is more important for innovation than for adoption. As the economy approaches the frontier, selection becomes more important. Countries at early stages of development pursue an investment-based strategy, which relies on existing firms and managers to maximize investment but sacrifices selection. Closer to the world technology frontier, economies switch to an innovation-based strategy with short-term relationships, younger firms, less investment, and better selection of firms and managers. We show that relatively backward economies may switch out of the investment-based strategy too soon, so certain policies such as limits on product market competition or investment subsidies, which encourage the investment-based strategy, may be beneficial. However, these policies may have significant long-run costs because they make it more likely that a society will be trapped in the investment-based strategy and fail to converge to the world technology frontier. (JEL: O31, O33, O38, O40, L16)

In a number of important historical instances industrialization processes, when launched at length in a backward country, showed considerable differences with more advanced countries, not only with regard to the speed of development (the rate of industrial growth) but also with regard to the productive and organizational structures of industry.... These differences in the speed and character of industrial development were to a considerable extent the result of application

of institutional instruments for which there was little or no counterpart in an established industrial country. Gerschenkron (1962, p. 7)

1. Introduction

In his famous essay, *Economic Backwardness in Historical Perspective*, Gerschenkron (1962) argued that relatively backward economies such as Germany, France, and Russia during the nineteenth century could rapidly catch up to more advanced economies by undertaking large investments and adopting frontier technologies. He emphasized that certain "non competitive" arrangements, including long-term relationships between firms and banks, as well as large firms and state intervention, might facilitate such convergence. If this assessment is correct, then the institutions and policies that are appropriate to relatively backward nations should encourage investment and technology adoption, even if this comes at the expense of various market rigidities and a less competitive environment. Implicit in this argument, and in the use of the term "appropriate," is also the notion that such arrangements are not beneficial for more advanced economies.

In this paper, we construct a simple endogenous growth model where certain relatively rigid arrangements emerge in equilibrium at early stages of development and disappear as the economy approaches the world technology frontier. We also use this framework to investigate how certain policies that might initially increase growth and the speed of convergence could then lead to slower growth.

To understand the main mechanism in our model, imagine a stylized economy with three key features: (i) entrepreneurs are either high skill or low skill; (ii) there are credit constraints restricting the amount of investment; and (iii) entrepreneurs engage both in innovation and adoption of existing technologies from the world technology frontier. If an entrepreneur is successful and is revealed to be high skill, he will continue to operate. If he is revealed to be low skill, the firm can dismiss him and replace him by a new draw, who will on average have higher skills. However, because of credit-market imperfections, the retained earnings of insider entrepreneurs enable them to undertake greater investments. Consequently, the decision to retain unsuccessful entrepreneurs creates a trade-off between investment and selection.

It is also plausible that skills (or match quality between firms and their activities) and the selection of the right entrepreneurs are more important for innovation than for adoption of existing technologies: Adoption and imitation are relatively straightforward activities compared to innovation. This leads to a key implication of our model: Retaining unsuccessful entrepreneurs is more costly, and less likely to arise in equilibrium, when innovation is more important. A corollary is that, as an economy approaches the world technology frontier and there remains less room for adoption and imitation, retention of unsuccessful entrepreneurs becomes less likely.

A likely equilibrium sequence is for an economy to start with an *investment-based strategy*, relying on long-term relationships between entrepreneurs (or managers) and firms (or financiers) in order to maximize investment. Intuitively, this strategy corresponds to an equilibrium where selection is less important, insiders are protected, and savings are channeled through existing firms in an attempt to achieve rapid investment growth and technology adoption. As the economy approaches the world technology frontier, lack of selection becomes more costly and there is typically a switch to an *innovation-based strategy*, where contracts with less successful entrepreneurs are terminated.

Furthermore, as suggested by Gerschenkron, government intervention to encourage the investment-based strategy might be useful because that strategy may otherwise fail to emerge even when it would be good for growth or welfare. This is due to the standard *appropriability effect* in models with monopolistic competition (as in most endogenous technical change models): Greater investment leads to greater productivity and output, but monopolists appropriate only part of these gains while bearing the investment costs in full. This creates a bias against large investments and hence against the investment-based strategy. Investment subsidies or anticompetitive policies which increase the amount of the productivity gains that monopolists can appropriate encourage the investment-based strategy and may increase the equilibrium growth rate.

Nevertheless, our analysis also reveals that the investment-based strategy can be socially costly in the long run. Countering the appropriability effect is the *rent-shield effect*: the cash (rents) in the hands of insiders creates a shield that protects them from more efficient newcomers. This effect can outweigh the appropriability effect and imply that an economy may stay in the investment-based strategy too long. Delayed switch to the innovation-based strategy reduces growth because the economy is not making best use of innovation opportunities. More important, there exists a level of development (distance to frontier) such that, if an economy does not switch out of the investment-based strategy before this threshold, it will be stuck in a *non-convergence trap* where convergence to the frontier stops.

An implication of this discussion is a new theory of leapfrogging. Economies pursuing policies encouraging the investment-based strategy may initially grow faster than others, but then become stuck in a non-convergence trap and be leapfrogged by the initial laggards. This is a different view of leapfrogging from the standard approach (e.g., Brezis, Krugman, and Tsiddon 1994), which is based on comparative advantage and learning-by-doing.

This analysis raises another important question: Why do governments not choose institutions and policies that favor the investment-based strategy when the country is in its early stages of development and then switch to policies supporting innovation and selection as the country approaches the frontier? The answer lies in the political economy of government intervention. Policies that favor the investment-based strategy create and enrich their own supporters. When economic

power buys political power, it becomes difficult to reverse policies that have an economically and politically powerful constituency. Consequently, societies may remain trapped with "inappropriate institutions" and relatively backward technologies.

The rest of the paper is organized as follows. Section 2 presents some motivating evidence and discusses the related literature. Section 3 outlines the basic model, and Section 4 characterizes the equilibrium. Section 5 discusses government policy and the possibility of political economy traps. Section 6 concludes.

2. Motivating Evidence and Related Literature

2.1. Evidence

The main assumption of our analysis is that innovation and selection become more important as an economy approaches the world technology frontier. We first investigate the plausibility of this assumption by looking at the correlation between distance to frontier and research and development (R&D) intensity at the industry level using data from the OECD sectoral database previously analyzed by (among others) Griffith, Redding, and Van Reenen (2004).

A natural (inverse) proxy for distance to frontier is "proximity to frontier" for an industry, PTF_{ict}, defined as TFP (total factor productivity) in industry i in country c at time t divided by the highest TFP in industry i at time t in the sample. We obtain estimates of proximity to frontier as well as data on R&D intensity, RD_{ict} (R&D divided by sales), for the years 1974–1990.[1]

Table 1 reports the correlation between these two measures with or without controlling for country and industry effects. The first three columns use a measure of PTF_{ict} without correcting for differences in skills and hours, and the last three columns use a measure that corrects for these differences (see Griffith, Redding, and Van Reenen 2004). All columns show the same pattern of a statistically significant positive correlation between proximity to frontier and R&D intensity: industries closer to their respective frontier are more R&D intensive. Moreover, as an industry approaches the world technology frontier more rapidly than others, it becomes relatively more R&D-intensive.

Naturally, these regressions do not show a causal effect of distance to frontier on R&D (e.g., past investments in R&D will affect productivity and distance to the frontier). Nevertheless, they are consistent with the view that

---

TABLE 1. Innovation and distance to frontier.

	No correction for difference in skills and hours			Correction for difference in skills and hours		
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to frontier	0.031	0.018	0.009	0.034	0.018	0.008
	(0.006)	(0.004)	(0.004)	(0.006)	(0.004)	(0.004)
Year dummies	YES	YES	YES	YES	YES	YES
Country dummies	NO	YES	YES	NO	YES	YES
Industry dummies	NO	YES	YES	NO	YES	YES
Country-industry dummies	NO	NO	YES	NO	NO	YES
Number of observations	1801	1801	1801	1801	1801	1801

Notes: Standard errors are in parentheses. The dependent variable is the ratio of R&D over value added at the 2/3 digit level. The independent variable Distance to frontier is the inverse of TFP in each industry relative to frontier (see Griffith, Redding, and Van Reenen 2004) and is defined as decreasing in the distance to frontier. The mean of the dependent variable is 0.033 and its standard deviation is 0.045. The mean of the independent variable is 0.729 (0.705 in columns 4, 5, 6) and its standard deviation is 0.196 (0.203 in columns 4, 5, 6).

R&D is more important in industries or countries closer to the world technology frontier.

Our analysis also implies that barriers to competition should have limited costs (or even benefits) when countries are far from the world technology frontier but should become much more costly near the frontier. Although this implication appears to be consistent with the experiences of many Latin American countries as well as with those of Korea and Japan,[2] we are not aware of any systematic empirical investigation. To take a first step in this direction, we consider a sample of non-OECD countries including those that joined the OECD in the 1990s, such as Korea and Mexico, but excluding former socialist countries. The sample is chosen so as to approximate "follower" countries, which are significantly behind the world technology frontier and therefore provide us with an opportunity to investigate convergence patterns.

We split the sample into low-barrier and high-barrier countries according to the "number of procedures to open a new business" variable from Djankov et al. (2002); the results are similar using the two other measures of barriers to entry from Djankov et al. Countries are classified into the *low-barrier* group if the number of such procedures is smaller or equal to ten and into the *high-barrier* group otherwise. This implies that 19 countries are classified as high barrier

---

2.   A number of Latin American countries, most notably Brazil, Mexico, and Peru, which grew relatively rapidly with import substitution and protectionist policies until the mid 1970s, stagnated and were taken over by other economies with relatively more competitive policies, such as Hong Kong or Singapore. The experiences of Korea and Japan are also consistent with this story. Though in many ways more market friendly than the Latin American countries, for much of the post-war period both Korea and Japan achieved rapid growth and convergence relying on high investment, large conglomerates, government subsidies, and relatively protected internal markets. Convergence and growth came to an end in the mid 1980s in Japan and during the Asian crisis in Korea (though Korea appears to have adopted a number of important reforms and resumed growth rapidly after the crisis).

and 23 countries as low barrier.[3] The barrier measures are time invariant, and so is our classification. Proximity to frontier is defined as the ratio of the country's GDP to the U.S. GDP at the beginning of the sample (from the Penn data set). For the cross-sectional regressions, per capita GDP growth rates are for 1965–1995, and the initial data are for 1965.

We define the two dummy variables *HB* and *LB*, where *HB* is equal to 1 for high-barrier countries and *LB* takes the value 1 for the low-barrier countries. We also control for a dummy for sub-Saharan Africa because countries in this region have experienced much slower growth than the rest of the world during this time period but we do not think that this is related to the mechanisms emphasized here (see Acemoglu, Johnson, and Robinson (2001) on the role of institutions, and see Easterly and Levine (1997) on the role of ethnolinguistic fragmentation in explaining low growth in Africa). Thus the estimating equation is

$$g_{i,65-95} = \alpha_{0,HB} LB_i + \alpha_{0,LB} LB_i + \alpha_{1,HB} \left(\frac{y_{i,65}}{y_{US,65}} HB_i \right)$$
$$+ \alpha_{1,LB} \left(\frac{y_{i,65}}{y_{US,65}} LB_i \right) + \alpha_2 SA_i + \varepsilon_i,$$

where $g_{i,65-95}$ is growth in GDP per capita in country i between 1965 and 1995, $y_{i,65}$ is GDP per capita in country i in 1965, $y_{US,65}$ is GDP per capita in the United States in 1965, SA_i is a dummy for sub-Saharan Africa, and HB_i and LB_i are the low- and high-barrier dummies defined previously. The coefficients of interest are the *convergence* coefficients, $\alpha_{1,HB}$ and $\alpha_{1,LB}$. A more negative estimate for $\alpha_{2,LB}$ implies that high-barrier countries do relatively well far from the frontier but worse closer to the frontier.

Column 1 of Table 2 (panel (A)) reports the estimates of $\alpha_{1,HB}$ and $\alpha_{1,LB}$ from this cross-sectional regression, which are also shown as the slope coefficients in Figures 1a and 1b. As can be seen in these figures, there is a stronger negative relationship between proximity to frontier ($y_{i,65}/y_{US,65}$) and growth for high-barrier countries.[4] For example, $\alpha_{1,LB}$ is estimated as -0.028 (S.E. $= 0.029$), and $\alpha_{1,HB}$ is -0.078 (S.E. $= 0.028$). Whereas there is a strong negative relationship between growth and proximity to frontier for countries with high barriers, the

---

3.   Our *low-barrier* group includes: Chile, Ghana, Hong Kong, India, Israel, Jamaica, Malaysia, Nigeria, Pakistan, Peru, Singapore, South Africa, Sri Lanka, Taiwan, Thailand, Tunisia, Uruguay, Zambia, and Zimbabwe; the *high-barrier* group includes Argentina, Bolivia, Brazil, Burkina Faso, Colombia, Dominican Republic, Ecuador, Egypt, Indonesia, Jordan, Kenya, Korea, Madagascar, Malawi, Mali, Mexico, Morocco, Mozambique, the Philippines, Senegal, Tanzania, Uganda, and Venezuela.

   The median number of procedures is 11 and four countries have exactly 11 procedures: Egypt, Indonesia, Kenya, and Uganda. The results are robust to classifying these countries into the low-barrier group.

4.   The vertical axes in the figures show country growth rates after the effect of the sub-Saharan Africa dummy, estimated in the corresponding multivariate regression, is taken out.

TABLE 2. Growth, distance to frontier, and barriers to entry.

	(1)	(2)	(3)	(4)	(5)	(6)
	Cross-section 1965–95	Panel regression (5-year averages, 1960–95)				
		OLS			IV	
Panel A: Dep. variable is growth rate of GDP per worker (annual average)						
High barriers (main effect)	0.040	0.039	0.021			
	(0.009)	(0.005)	(0.009)			
Low barriers (main effect)	0.036	0.029	0.011			
	(0.008)	(0.005)	(0.009)			
Distance to frontier * high barriers	−0.078	−0.062	−0.072	−0.109	−0.214	−0.245
	(0.028)	(0.013)	(0.016)	(0.047)	(0.072)	(0.080)
Distance to frontier * low barriers	−0.028	0.009	−0.018	−0.039	−0.035	−0.046
	(0.029)	(0.017)	(0.025)	(0.037)	(0.049)	(0.052)
p-value difference interaction coeffs.	0.219	0.002	0.061	0.237	0.041	0.035
Dummy sub-saharan Africa	YES	YES	YES	NO	NO	NO
Time dummies	NO	YES	YES	YES	YES	YES
Country fixed effects	NO	NO	NO	YES	YES	YES
Control for education	NO	NO	YES	NO	NO	YES
Number of observations	43	300	262	300	290	255
R^2	0.65	0.44	0.47	0.49	0.21	0.20
Panel B: Dep. variable is growth rate of GDP per worker (annual average)						
Distance to frontier (main effect)	−0.016	−0.045	0.011	0.048	0.195	0.195
	(0.055)	(0.036)	(0.047)	(0.074)	(0.105)	(0.111)
Barriers (main effect)	−0.001	0.000	0.000			
	(0.001)	(0.001)	(0.001)			
Distance to frontier * barriers	−0.004	−0.007	−0.005	−0.011	−0.030	−0.032
	(0.005)	(0.003)	(0.004)	(0.007)	(0.011)	(0.012)
Dummy sub-saharan Africa	YES	YES	YES	NO	NO	NO
Time dummies	NO	YES	YES	YES	YES	YES
Country fixed effects	NO	NO	NO	YES	YES	YES
Control for education	NO	NO	YES	NO	NO	YES
Number of Observations	43	300	262	300	290	255
R^2	0.52	0.29	0.29	0.46	0.17	0.14

Note: Standard errors are in parentheses. The regressions in column 1 are cross-sectional with one observation per country, and the dependent variable is the annual average growth rate of GDP per worker (in PPP terms) over the period 1965–1995. The remaining columns describe panel regressions, and the dependent variable is the average growth for five-years intervals, 1960–1965, 1965–1970,…,1990–1995. The independent variable Distance to frontier is the ratio of the country's GDP per worker to the GDP per worker in the U.S., both calculated at the beginning of each period. The independent variable Barriers in panel (B) is the "procedure measure" from Djankov et al. (2002), which measures the number of procedures necessary to open a business. The independent variable High barriers (Low barriers) in panel (A) is a dummy variable taking the value 1 for countries with a number of procedures larger or equal to (smaller than) 11, and zero else. The control variable for education is the average years of schooling in the male population over 25 at the beginning of each period.

In panel (A), columns 5 and 6, the interactions between distance to frontier and the dummy for high- and low-barriers are instrumented using one-period lags of the same variables. In panel (B), columns 5 and 6, both the main effect of the distance to frontier and its interaction with barriers are instrumented using one-period lags of the same variables.

relationship is much weaker for countries with low barriers. In other words, high-barrier countries do relatively well (i.e., converge rapidly) when they are far from the frontier but slow down significantly near the frontier, whereas low-barrier countries grow almost equally successfully near or far from the frontier. This is consistent with the notion that barriers to entry are more harmful to growth
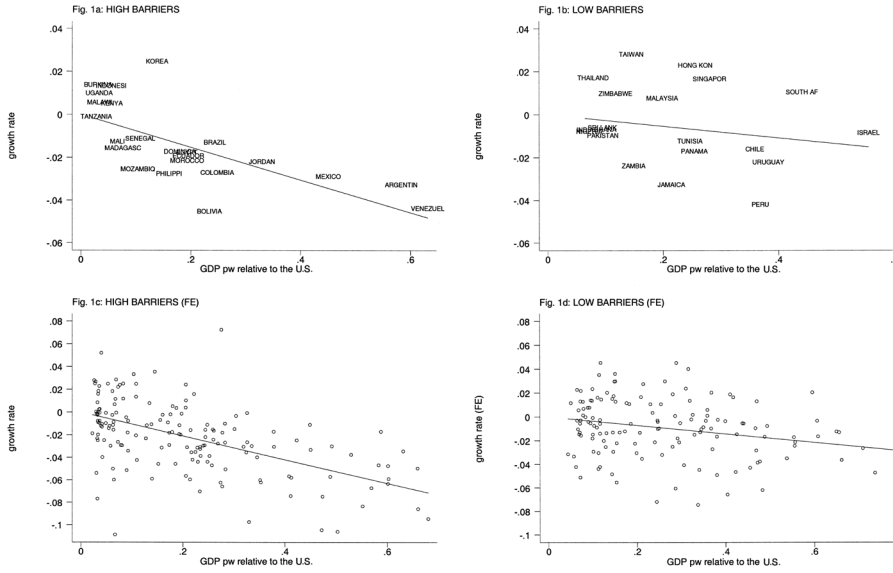
FIGURE 1. Growth and proximity to frontier in countries with high and low barriers to entry.

closer to the frontier, although this cross-country relationship may be driven by other omitted cross-country differences. Nevertheless, the *p*-value at the bottom of the table shows that in this case we cannot reject the hypothesis that these two coefficients are equal given the standard errors in this cross-sectional regression.

The cross-sectional regression does not exploit all of the relevant information, however, for the implication of our approach is that, at any point in time, there should be a stronger relationship between proximity to frontier and growth for the high-barrier countries than for the low-barrier countries. To investigate this issue, we next estimate regressions of the form

$$g_{i,t} = \alpha_{0,HB} LB_i + \alpha_{0,LB} LB_i + \alpha_{1,HB} \left(\frac{y_{i,t-1}}{y_{US,t-1}} HB_i \right)$$

$$+ \alpha_{1,LB} \left(\frac{y_{i,t-1}}{y_{US,t-1}} LB_i \right) + \alpha_2 SA_i + f_t + \varepsilon_{it},$$

where $g_{i,t}$ is the growth rate in country i between $t-1$ and t, $y_{i,t-1}$ is GDP per capita in country i at date $t-1$, $y_{US,t-1}$ is GDP per capita in the United States at date $t-1$, f_t denotes a full set of time effects, and we take the time intervals to be 5 years.[5] The results are reported in column 2 of Table 2 (panel (A)) and show

---

5. The sample for the five-year regressions is not balanced, and we extend the sample back to 1960 for some countries. The results are very similar if we start in 1965 for all countries.

a similar pattern, with no relationship between growth and proximity to frontier for low-barrier countries, yet with a strong negative relationship for the high-barrier countries. For example, $\alpha_{1,LB}$ is now estimated to be 0.009 (S.E. = 0.017), and $\alpha_{1,HB}$ is -0.062 (S.E. = 0.013). The difference between these two coefficients is now statistically significant at the 1% level. The next column adds controls for years of schooling (we use male use of schooling from Barro and Lee 1996), and the pattern is unchanged.

The patterns shown in columns 1–3 of Table 2 (panel (A)) could be driven by some omitted country characteristics. A stronger test of the implication of our model would be to see whether growth in high-barrier countries slows down more significantly as they approach the frontier relative to growth in low-barrier countries. In order to investigate this, in column 4 of Table 2 (panel (A)) we augment the specification to include a full set of country fixed effects. Hence we are now investigating whether the same pattern holds when we look at deviations from the country's "usual" growth rate. The results confirm the pattern shown in the previous columns: $\alpha_{1,LB}$ is estimated to be -0.039 (S.E. = 0.037), and $\alpha_{1,HB}$ is estimated at -0.109 (S.E. = 0.047). Figures 1d and 1e show the convergence patterns captured by these *within* regressions. Near the frontier, a country with high barriers grows less than its usual growth rate. Therefore, as implied by our model, countries with high barriers slow down more significantly as they approach the frontier. But the difference between the coefficients is once again statistically insignificant.

Nevertheless, the results in column 4 of Table 2 are difficult to interpret because of the standard bias in models with fixed effects and lagged dependent variables (see, e.g., Wooldridge (2002), chap. 10). Proximity to frontier is correlated with the lags of the dependent variable, because $g_{i,t} \approx (y_{i,t} - y_{i,t-1})/y_{i,t}$. This creates a bias in the estimation of the fixed effects and therefore in the estimates of the α_1. To deal with this problem, in columns 5 and 6 of Table 2 we report regressions where proximity to frontier is instrumented by its one-period lag. The results are similar to those reported in column 4 the estimate of $\alpha_{1,LB}$ is -0.035 (S.E. = 0.049), and $\alpha_{1,HB}$ is estimated at -0.214 (S.E. = 0.072) (e.g., in column 5). The difference between the coefficients $\alpha_{1,HB}$ and $\alpha_{1,LB}$ is statistically significant at the 5% level without years of schooling and at the 1% level with years of schooling in the regression.

Finally, in the bottom panel of Table 2 we report regressions that interact the barrier variable with proximity to frontier:

$$g_{i,t} = \beta_0 + \beta_1 B_i + \beta_2 \frac{y_{i,t-1}}{y_{US,t-1}} + \beta_3 \left(\frac{y_{i,t-1}}{y_{US,t-1}} B_i\right) + d_i + f_t + \varepsilon_{it};$$

here B_i denotes the level of barriers in country i, and β_3 is the coefficient of interest. The results are consistent with those reported in the top panel. In column 2, the interaction term β_2, which now captures the difference between $\alpha_{1,HB}$ and $\alpha_{1,LB}$, is

estimated to be negative and statistically significant (-0.007, with S.E. $= 0.003$). This again implies that high-barrier countries slow down more the closer they get to the frontier. Other results, with the exception of those in column 3, also show a significant interaction term.

2.2. Related Literature

Our paper relates to a number of different literatures. First, the notion that skills are more important for innovation than for adoption is closely related to the role of human capital in technological progress emphasized in the seminal paper by Nelson and Phelps (1966), as well as to the emphasis in Galor and Tsiddon (1997) and Hassler and Rodriguez Mora (2000) on the importance of ability and skill in times of economic change and turbulence. Nelson and Phelps, for example, rank activities by the degree to which they require adaptation to change. They write: "At the bottom of this scale are functions that are highly routinized.…In the other direction on this scale we have, for example, innovative functions which demand keeping abreast of improving technology" (p. 69). Nelson and Phelps argue that the importance of human capital increases with the innovative content of the tasks performed or with the extent to which it is necessary to follow and to understand new technological developments.

Second, our focus is related to work on technological convergence, including Barro and Sala-i-Martin (1997), Aghion and Howitt (1992, 1998), Zeira (1998), Howitt (2000), and especially to Howitt and Mayer (2002), who investigate how some countries may stagnate while others converge to an income level below the world technology frontier.

Third, our model is related to work on finance and growth, including Greenwood and Jovanovic (1990), King and Levine (1993), and Acemoglu and Zilibotti (1997). More closely related is Tong and Xu (2004), who extend the model by Dewatripont and Maskin (1995) to compare "multi-financier" and "single-financier" credit relationships, emphasizing that multi-financier relationships become more beneficial at later stages of development when selecting good R&D projects becomes more important. None of these papers, however, investigates how certain arrangements that are at first growth enhancing can later reduce growth and cause non convergence traps. The only exception is Rajan and Zingales (1999), who suggest that the same practices that were useful for the success of East Asian economies may have also been responsible for the East Asian crisis which is similar to our argument that certain social arrangements are first beneficial and then become costly. Nevertheless, Rajan and Zingales neither develop this point formally nor provide empirical evidence supporting this claim.

3. The Model

3.1. *Agents and Production*

The model economy is populated by overlapping generations of risk-neutral agents, who live for two periods and discount the future at the rate r. The population is constant. Each generation consists of a mass $1/2$ of *capitalists* with property rights on "production sites" but with no skills or other wealth, and a mass $(N + 1)/2$ of workers who are born with no wealth but are endowed with skills. Property rights are transmitted within dynasties. All workers supply their labor inelastically and are equally productive in production tasks, but they have heterogeneous productivity in entrepreneurial activities. In particular, we assume that each worker is a high-skill (ability) entrepreneur with probability λ and a low-skill entrepreneur with probability $1 - \lambda$.

There is a unique final good in the economy that is also used as an input to produce intermediate goods. We take this good as the numéraire. The final good is produced competitively using labor and a continuum one of intermediate goods as inputs with the aggregate production function

$$y_t = \frac{1}{\alpha} N_t^{1-\alpha} \left(\int_0^1 (A_t(v))^{1-\alpha} x_t(v)^\alpha \, dv \right), \tag{1}$$

where $A_t(v)$ is productivity in sector v at time t, $x_t(v)$ is the flow of intermediate good v used in final-good production at time t, N_t is the number of production workers at time t, and $\alpha \in (0, 1)$.

In each intermediate sector v, one production site has access to the most productive technology, $A_t(v)$, so this "leading firm" will enjoy monopoly power. Each leading firm has access to a technology capable of transforming one unit of the final good into one unit of intermediate good with productivity $A_t(v)$. A fringe of additional firms can "steal" this technology and produce the same intermediate good, with the same productivity $A_t(v)$, without using the production site or an entrepreneur. But this fringe faces higher costs of production and needs χ units of the final good to produce one unit of the intermediate, where $1/\alpha \geq \chi > 1$ (naturally, these firms will not be active in equilibrium). The parameter χ captures technological factors as well as government regulation affecting entry. A higher χ corresponds to a less competitive market. The fact that $\chi > 1$ implies that the fringe is less productive than the incumbent producer, while $\chi \leq 1/\alpha$ implies that this productivity gap is sufficiently small that the incumbent will be forced to charge a limit price in order to prevent entry by the fringe. Naturally, this limit price will be equal to the marginal cost of the fringe,

$$p_t(v) = \chi. \tag{2}$$

The final-good sector is competitive, so each intermediate-good producer v at date t faces the inverse demand schedule $p_t(v) = (A_t(v)N_t/x_t(v))^{1-\alpha}$. This equation together with (2) gives equilibrium demands: $x_t(v) = \chi^{-(1/(1-\alpha))} A_t(v)N_t$, with equilibrium profits therefore equal to

$$\pi_t(v) = [p_t(v) - 1]x_t = \delta A_t(v)N_t, \tag{3}$$

where $\delta \equiv (\chi - 1)\chi^{-(1/(1-\alpha))}$ is monotonically increasing in χ (because $\chi \leq 1/\alpha$). Thus, a higher δ corresponds to a less competitive market and implies higher profits for the leading firms.

Equation (1) gives aggregate output as $y_t = \alpha^{-1}\chi^{-(\alpha/(1-\alpha))} A_t N_t$, where

$$A_t \equiv \int_0^1 A_t(v)dv \tag{4}$$

is the average level of technology in the economy at time t. The market-clearing wage level is equal to the marginal product of labor in production:

$$w_t = (1-\alpha)\alpha^{-1}\chi^{-[\alpha/(1-\alpha)]} A_t. \tag{5}$$

Finally, let y_t^{net}, the net output, be final output minus the cost of intermediate production. Then

$$y_t^{net} = y_t - \int_0^1 x_t(v)dv = \zeta A_t N_t, \tag{6}$$

where $\zeta \equiv (\chi - \alpha)\chi^{-(1/(1-\alpha))}/\alpha$ is monotonically decreasing in χ. Thus, for given average technology A_t, both total and net output are decreasing in the extent of monopoly power (i.e., in χ) because of standard monopoly distortions. Note also that net output, (6), and profits, (3), have identical forms except that net output has the term ζ instead of $\delta < \zeta$. This reflects an *appropriability effect*: Monopolists capture only a fraction of the greater productivity in the final-good sector (or of the consumer surplus) created by their production.

3.2. Technological Progress and Productivity Growth

Each leading firm (capitalist) requires an entrepreneur (or manager) to operate the firm. This leaves $N_t = N$ production workers (recall that the total size of the worker population is $N + 1$). This implies that $\pi_t(v) = \delta A_t(v)N$ for all v and t.[6]

---

6. This expression shows that profits are increasing in the size of the population, N, and, as in most models of endogenous growth, this creates a scale effect here. This scale effect plays no role in any of our results, and we do not emphasize the comparative statics with respect to this variable. The scale effect can be removed by modifying the model, for example, by introducing a maximal span of control at the firm level and by introducing free entry to determine the number of firms.

Firm productivity is determined by entrepreneurial skill and by the size of the project that the entrepreneur operates. To simplify the discussion, we assume that there are two possible project sizes, "small" and "large." Running a project requires an additional investment, which is naturally greater for the large project than for the small project. The investment cost can be financed either through the retained earnings of the entrepreneur or by the capitalist who owns the firm. At the beginning of the period, capitalists can borrow from a set of competitive intermediaries who collect funds from consumers. Intermediation is without any cost, and there is free entry into this activity. Moreover, because intermediation takes place within a period, there are no interest costs to be covered.

Entrepreneurial skills, which affect productivity growth, are initially unknown but are revealed after an agent works as an entrepreneur for the first time. Entrepreneurs perform two important tasks: (i) they engage in *innovation*, and entrepreneurial skills are important for success in this activity, (ii) they also *adopt* technologies from the frontier, and here skills play a less important role than in innovation. This assumption captures the notions that relatively backward economies can grow by adopting already well-established technologies and that entrepreneurial selection is less important for adoption than for innovation.

Let us denote the growth rate of the world technology frontier, $\bar{A}_t$, by g, so that

$$\bar{A}_t = \bar{A}_0(1 + g)^t. \tag{7}$$

We shall return shortly to the determination of this growth rate. All countries have a state of technology A_t, as defined by (4), that is less than the frontier technology. In particular, for the representative country, we have $A_t \leq \bar{A}_t$.

The productivity of intermediate good v at time t is expressed as

$$A_t(v) = s_t(v)\left[\eta\bar{A}_{t-1} + \gamma_t(v)A_{t-1}\right]; \tag{8}$$

here $s_t(v) \in \{\sigma, 1\}$ denotes the size of the project, with $s_t(v) = \sigma < 1$ corresponding to a small project and $s_t(v) = 1$ corresponding to a large project. The term $\gamma_t(v)$ denotes the skill level of the entrepreneur. Equation (8) captures the two dimensions of productivity growth: adoption and innovation. By adopting existing technologies, firms benefit from the state of world technology in the previous period, $\bar{A}_{t-1}$, irrespective of the skill of the entrepreneur. In addition there is productivity growth due to innovation building on the body of local knowledge, A_{t-1}, and success in innovation depends on entrepreneurial skills as captured by the term $\gamma_t(v)$. This feature introduces the assumption that entrepreneurial skills are more important for innovation than for imitation; put differently, innovation relies on entrepreneurial selection more than does imitation. Finally, equation (8) also implies that greater investment (the large project) leads to higher productivity growth.

Rearranging (8) and using the definition in (4) yields the growth rate of aggregate technology,

$$\frac{A_t}{A_{t-1}} \equiv \frac{\int_0^1 A_t(v)dv}{A_{t-1}} = \int_0^1 s_t(v)\left[\eta\frac{\bar{A}_{t-1}}{A_{t-1}} + \gamma_t(v)\right]dv. \tag{9}$$

Equation (9) shows the importance of *distance to frontier*, as captured by the term $\bar{A}_{t-1}/A_{t-1}$. When this term is large, the country is far from the world technology frontier and the major source of growth is the adoption of already well-established technologies, as captured by the $\eta\bar{A}_{t-1}/A_{t-1}$ term. When $\bar{A}_{t-1}/A_{t-1}$ becomes close to unity, so that the country is close to the frontier, innovation matters relatively more and growth is driven by the $\gamma_t(v)$ term. Consequently, as the country develops and approaches the world technology frontier, innovation and entrepreneurial selection become more important.

For simplicity, we assume that $\gamma_t(v) = 0$ for a low-skill entrepreneur and denote the productivity of a high-skill entrepreneur by $\gamma_t(v) = \gamma > 0$. To guarantee a decreasing speed of convergence to the world technology frontier, we also assume that $\lambda\gamma < 1$ (recall that, λ is the fraction of high-skill agents in the population).

Finally, the cost of investment for the small and large projects is, respectively,

$$k_t(v \mid s) = \begin{cases} \phi\kappa\bar{A}_{t-1} & \text{if } s = \sigma, \\ \kappa\bar{A}_{t-1} & \text{if } s = 1, \end{cases} \tag{10}$$

where $\phi \in (0, 1)$. The assumption that investment cost is proportional to $\bar{A}_{t-1}$ ensures balanced growth.[7] Intuitively, an important component of entrepreneurial activity is to undertake imitation and adaptation of already existing technologies from the world frontier. As this frontier advances, entrepreneurs need to incur greater costs to keep up with and make use of these technologies, hence investment costs increase with $\bar{A}_{t-1}$.

3.3. Contracts, Incentive Problems, and Credit Constraints

Capitalists have deep pockets (they can borrow from competitive intermediaries at the exogenous interest rate r) and make contract offers to a subset of workers to become entrepreneurs, specifying the loan amount from intermediaries, as well as payments to entrepreneurs and the level of investment. Investment costs are

---

7.   Alternatively, investment costs of the form $k_t(v) = \kappa\bar{A}_{t-1}^{\rho}A_{t-1}^{1-\rho}$ for any $\rho \in [0, 1]$ would ensure balanced growth. We choose the formulation in the text with $\rho = 1$ because it simplifies some of the expressions, without affecting any of our major results. See the NBER working paper version of this paper (Acemoglu, Aghion, and Zilibotti 2002) for the expressions when $\rho < 1$.

financed either through the retained earnings of entrepreneurs or by capitalists. To simplify the discussion, we also assume that young capitalists (new firms) cannot hire old entrepreneurs (e.g., because old cohort skills are not adaptable to the new vintage of technologies); thus a new firm (young capitalist) necessarily employs a young entrepreneur.

Entrepreneurs engaged in innovative activities, or even simply entrusted with managing firms, are difficult to monitor. This creates a standard moral hazard problem, which we formulate in the simplest possible way: We assume that an entrepreneur can divert a fraction μ of the returns for his own use and will never be prosecuted. The parameter μ measures the extent of the incentive problems, or, equivalently, the severity of the credit-market imperfections resulting from these incentive problems. Moral hazard plays two important roles in our model: first, it creates credit constraints and so restricts investment, especially for young entrepreneurs who do not have any retained earnings; second, via this channel, it enables the retained earnings of old entrepreneurs (or, equivalently, the cash in the hands of existing businesses) to shield them against the threat of entry by new entrepreneurs.

To specify the incentive compatibility constraints more formally, define $\pi_t(v \mid s, e, z)$ as the ex post cash flow generated by firm v at date t as a function of the size of the project, $s \in \{\sigma, 1\}$, and of the entrepreneur's age, $e \in \{Y, O\}$ and skill level, $z \in \{L, H\}$; here Y denotes young, O old, L low skill, and H high skill. Observe that $\pi_t(v \mid s, e, z)$ is simply given by the expression in (3) with $A_t(v)$ substituted from (8) as a function of s, e and z. For the entrepreneur not to divert revenues, the following incentive compatibility constraint must be satisfied:[8]

$$S_t(v \mid s, e, z) - \mu \pi_t(v \mid s, e, z) \geq 0, \tag{11}$$

where $S_t(v \mid s, e, z)$ is the payment to an entrepreneur of age e and skill z who is running a project of size s. Thus, incentive compatibility requires the manager to be paid a certain fraction of the ex post profits.

Let us define $\widehat{RE}_t(v \mid s, e, z) \leq k_t(v \mid s)$ as the retained earnings injected by an entrepreneur to finance part of the investment costs and define $RE_t(v \mid s, e, z)$ as total retained earnings.[9] Naturally, we have $0 \leq \widehat{RE}_t(v \mid s, e, z) \leq RE_t(v \mid s, e, z)$ and, because young entrepreneurs have no funds to contribute,

---

8. This specification rules out long-term contracts where the payment to an old entrepreneur is conditioned on whether or not he has diverted funds in the first period. Such long-term contracts would require a commitment technology on the part of capitalists, which we assume is not present in this economy. Introducing credible long-term contracts does not affect the main results of the analysis.

9. This inequality implies that side payments from entrepreneurs to capitalists are not possible. This assumption can be motivated by various arguments. For example, firms may be unable to commit to employ an entrepreneur after receiving the side payment (note that, provided a new entrepreneur must incur the investment cost again, this commitment problem would not rule out the use of retained earnings to finance part of the investment costs).

$\widehat{RE}_t(\nu \mid s, e = Y, z) = 0$. Let us next define

$$V_t(\nu \mid s, e, z) = \pi_t(\nu \mid s, e, z) - S_t(\nu \mid s, e, z) - \left[k_t(\nu \mid s) - \widehat{RE}_t(\nu \mid s, e, z)\right] \tag{12}$$

as the value of a capitalist with a project of size s run by an entrepreneur of age e and skill z, and define

$$s^*(e, z) \in \arg\max_s E_t V_t(\nu \mid s, e, z) \tag{13}$$

as the profit-maximizing project size choice for capitalists given entrepreneur of age e and of skill z, where E_t is the expectations operator at time t (this applies in the case of young entrepreneurs whose skills are yet unknown).

Capitalists maximize their expected returns as given in (12) subject to the incentive compatibility constraints in (11) and a set of participation constraints for the entrepreneurs. These participation constraints are given in Appendix A, where we also show that, as long as $\sigma < 1/(1+r)$ and N is sufficiently large, all of the participation constraints are slack even if entrepreneurs inject all of their retained earnings to finance part of the cost of investment. In the text, we therefore ignore these constraints.

Finally, let us denote the maximized value of the capitalists by

$$E_t V_t^*(e, z) = E_t V_t[\nu \mid s^*(e, z), e, z], \tag{14}$$

with managerial payments $S_t(\nu \mid s, e, z)$ satisfying the incentive compatibility constraints in (11).

Because the participation constraint is slack, there will be an excess supply of young agents willing to become entrepreneurs. Hence young entrepreneurs will be paid the lowest salary consistent with incentive compatibility, (11). The same holds for old low-skill entrepreneurs (because these entrepreneurs cannot work in young firms, old capitalists will make take-it-or-leave-it offers to them, forcing them down to their incentive compatibility constraint).[10] But there will typically be an excess demand for old entrepreneurs who are revealed to be high skill. Competition between old capitalists then implies that[11]

$$V_t^*(e = O, z = H) \leq \max[V_t^*(e = O, z = L), E_t V_t^*(e = Y)]. \tag{15}$$

Suppose this condition did not hold. Then an old capitalist currently working with either an old low-skill entrepreneur or a young entrepreneur could deviate

---

10. If young firms could also hire old entrepreneurs, competition between young and old firms would generate rents for old low-skill entrepreneurs. This would complicate the expressions, without affecting any of the qualitative results.

11. With a slight abuse of notation, we denote the expected value of the firm with a young entrepreneur $E_t V_t^*(e = Y)$ (i.e., we omit the argument z). Similarly, we denote the optimal size of a firm with a young entrepreneur by $s_t^*(e = Y)$.

offering a higher salary to attract an old high-skill entrepreneur and thus increase his profits. To rule out such deviations, (15) must hold.

4. Equilibrium

4.1. Definition of Equilibrium

To define an equilibrium, let us first introduce the notation

$$a_t \equiv \frac{A_t}{\bar{A}_t} \tag{16}$$

as proximity to frontier, an inverse measure of the country's distance to frontier. This variable will summarize the state of the economy.

The key decisions in this economy are the level of investment (project size) with various types of entrepreneurs and whether to terminate an entrepreneur and replace him with a new one. It is clear that high-skill entrepreneurs will always be retained, so the crucial choice is whether a low-skill entrepreneur will be retained or not. We denote the retention decision by $R_t(\nu) \in \{0, 1\}$, with $R_t = 0$ corresponding to termination and $R_t = 1$ corresponding to retention.

A *static equilibrium* (given the state a_t of the economy) is then a set of intermediate good prices, $p_t(\nu)$, that satisfy (2), profit levels given by (3), a wage rate w_t given by (5), project size choices $s^*(e, z)$ given by (13), and a continuation decision with low-skill entrepreneurs R_t, such that

$$R_t = \begin{cases} 0 & \text{if } E_t V_t^*(e = Y) \geq V_t^*(e = O, z = L), \\ 1 & \text{if } E_t V_t^*(e = Y) < V_t^*(e = O, z = L). \end{cases}$$

A *dynamic equilibrium* is obtained by piecing together static equilibria as defined in this section through the law of motion of aggregate productivity as given by (9). We provide the equilibrium law of motion in greater detail as follows.

4.2. Equilibrium Investment and Refinancing decision

In this section we characterize the equilibrium investment (project size) and refinancing decisions. Even when, absent moral hazard, it would be profitable for firms to pay the investment cost and operate the large project, credit-market imperfections and moral hazard can lead to underinvestment in equilibrium (i.e., firms may run small projects even though large projects are socially more efficient). To understand why, note that because of the incentive compatibility constraint (11) profits must be shared between the capitalist and the entrepreneur, with respective shares $1 - \mu$ and μ, after production is realized. Entrepreneurs, however,

are credit-constrained, and this forces capitalists to bear the bulk (sometimes the whole) of the investment cost, although they appropriate only a fraction $1 - \mu$ of the returns. Hence, underinvestment can occur in equilibrium.

The underinvestment problem tends to be more severe when firms hire young entrepreneurs who have no wealth and so force capitalists to bear the entire investment cost. Old entrepreneurs, in contrast, can bear part of the cost by injecting their retained earnings. To see why, consider the extreme case in which an old entrepreneur is willing to cover the entire investment cost out of his retained earnings. Then the firm will necessarily prefer to operate the large project with this entrepreneur, and there will be no underinvestment. This example illustrates why retained earnings mitigate the underinvestment problem but also reduces selection by making old low-skill entrepreneurs more attractive to firms.

Although different equilibrium configurations are possible, we now restrict attention to the region of the parameter space where the trade-off emphasized in this paper emerges more clearly: Refinancing an old entrepreneur who has proven to be of low skill may be profitable because it mitigates the credit-market imperfection.

LEMMA 1.   *Let*

$$\delta_L \equiv \frac{\kappa}{N} \frac{1}{(1 - \mu)(1 - \sigma)\eta + \frac{1+r}{1+g}\mu\sigma\eta},$$

$$\delta_H \equiv \frac{\kappa}{N} \frac{1 - \phi}{(1 - \mu)(1 - \sigma)(\eta + \lambda\gamma)},$$

and suppose that $\delta \in (\delta_L, \delta_H)$. Then, for all $a \in [0, 1]$, the following statements hold:

1. *Young entrepreneurs operate small projects (i.e., $s_t^*(e = Y) = \sigma$).*
2. *If an old low-skill entrepreneur is retained (i.e., $R_t = 1$), then he operates a large project (i.e., $s_t^*(e = O, z = L) = 1$) and contributes to his entire retained earnings,*

$$RE_t = \frac{1 + r}{1 + g}\sigma\mu\delta N\eta\bar{A}_{t-1}, \tag{17}$$

   to finance the project.
3. *All high-skill entrepreneurs are always retained and operate large projects, that is, $s_t^*(e = O, z = H) = 1$.*

This lemma is proved in Appendix B, where we also show that $\delta_L < \delta_H$ for all $\phi < \bar{\phi}$, so that the interval (δ_L, δ_H) is nonempty. The assumption that ϕ is sufficiently low ensures that capital market imperfections, rather than productivity differences, are the main determinant of the firm's decision on project size.

Lemma 1 establishes that, for a non empty range of competition policies $\delta \in (\delta_L, \delta_H)$, if a low-skill entrepreneur is retained, then he must run a large project. In the rest of the paper, we restrict our attention to this range.[12] Note also that the expression for RE_t follows immediately, because the entrepreneur in question is low-skill and operated a small project in his youth.

When does the firm prefer to retain a low-skill old entrepreneur rather than hire a young entrepreneur? Consider the value of a firm that retains an old low-skill entrepreneur and operates the large project:

$$V_t(\nu \mid s = 1, e = O, z = L) = \left[(1 - \mu)\delta N \eta \bar{A}_{t-1} - \max(\kappa \bar{A}_{t-1} - RE_t, 0)\right],$$
(18)

where RE_t is given by (17).

To simplify the discussion, we wish to ensure that $\kappa \bar{A}_{t-1} > RE_t$, so that the cost of the large project is greater than retained earnings (see the NBER Working paper version for the analysis when this assumption is relaxed and so the entrepreneur finances the entire cost of the project). The following assumption is sufficient to ensure that this will be the case for all $\delta \in (\delta_L, \delta_H)$.

ASSUMPTION 1.   We assume that $\delta_H < \kappa/[\sigma \mu \eta N(1 + r)/(1 + g)]$.

Under this assumption, the term $\max[\kappa \bar{A}_{t-1} - RE_t, 0]$ in the right-hand side of (18) simplifies to $\kappa \bar{A}_{t-1}$ for all $\delta \in (\delta_L, \delta_H)$.

In contrast, the value of the firm that hires a young entrepreneur and runs a small project is

$$E_t V_t(\nu \mid s = \sigma, e = Y) = (1 - \mu)\delta N \sigma (\eta + \lambda \gamma a_{t-1})\bar{A}_{t-1} - \phi \kappa \bar{A}_{t-1}.$$ (19)

The firm retains low-skill old entrepreneurs whenever

$$V_t(\nu \mid s = 1, e = O, z = L) > E_t V_t(\nu \mid s = \sigma, e = Y),$$ (20)

where the left-hand side is given by (18) and the right-hand side by (19). If condition (20) does not hold, then low-skill entrepreneurs are terminated. Condition (20) defines a threshold level $a_r(\mu, \delta)$ of the distance to frontier such that low-skill old entrepreneurs are retained ($R = 1$) below this threshold but are terminated ($R = 0$) above this threshold.

---

12.   The analysis can easily be extended to the case where $\delta \notin (\delta_L, \delta_H)$. If $\delta < \delta_L$, old low-skill entrepreneurs, if retained, are assigned to small projects that they partially or totally finance through retained earnings. In this case, there is no trade-off between selection and investment size. On the other hand, if $\delta > \delta_H$, young entrepreneurs operate large projects, and once again there is no trade-off.

Using (18) and (19), we obtain this threshold as

$$a_r(\mu, \delta) \equiv \frac{\left[(1-\mu)(1-\sigma) + \frac{1+r}{1+g}\mu\sigma\right]\eta - \frac{\kappa(1-\phi)}{\delta N}}{(1-\mu)\sigma\lambda\gamma}. \tag{21}$$

The threshold $a_r(\mu, \delta)$ is increasing in δ: When product markets are less competitive (higher δ), the switch to $R = 0$ occurs later. This comparative static reflects two forces. The first is the appropriability effect, which (as already pointed out) implies that firms do not capture the entire surplus created by technological progress. Capitalists bear the costs of investment but, because of the appropriability effect, obtain only a fraction of the returns. Consequently, they have a bias against retaining old entrepreneurs, which is associated with greater investment expenditures. A higher δ weakens the extent of this appropriability effect and enables firms, and hence capitalists, to capture more of the surplus and so encouraging the retention of old entrepreneurs. Second, as shown by (17), a higher δ implies greater profits and greater retained earnings for old entrepreneurs, which they can use to shield themselves against competition from young entrepreneurs, making their own retention more likely.

However, the effect of incentive problems μ, on $a_r(\mu, \delta)$ is ambiguous. On the one hand, a higher μ increases the earnings retained by entrepreneurs and raises these insiders' shields against competition from newcomers, encouraging $R = 1$. On the other hand, a higher μ reduces the profit differential between hiring a young versus an old low-skill entrepreneur. If

$$\delta > \frac{(1-\phi)\kappa}{\sigma\eta L}\frac{1+g}{1+r}, \tag{22}$$

then the former effect dominates and a_r is increasing in μ, so that more severe moral hazard problems encourage $R = 1$. In contrast, when (22) does not hold, these problems encourage the termination of low-skill entrepreneurs.

The static equilibrium is summarized in the following proposition.

PROPOSITION 1. *Suppose Assumption 1 holds and $\delta \in (\delta_L, \delta_H)$, and let $a_r(\mu, \delta)$ be defined by (21). Then, for given a_{t-1}, there exists a unique equilibrium such that (i) young entrepreneurs operate small projects ($s_t^*(e = Y) = \sigma$); (ii) old low-skill entrepreneurs are retained ($R_t = 1$) and operate large projects ($s_t^*(e = O, z = L) = 1$) when $a_{t-1} < a_r(\mu, \delta)$ but are terminated ($R_t = 0$) when $a_{t-1} > a_r(\mu, \delta)$; and (iii) old high-skill entrepreneurs are always retained and operate large projects ($s_t^*(e = O, z = H) = 1$) for all $a_{t-1} \in [0, 1]$. The threshold $a_r(\mu, \delta)$ is increasing in δ.*

4.3. Dynamic Equilibrium

We now characterize the dynamic equilibrium of the economy. Let us first determine the law of motion of a_t conditional on the retention decision R_t. We observe that half the firms are young and then use (4) to write

$$A_t \equiv \int_0^1 A_t(v)dv = (A_t^Y + A_t^O)/2,$$

where A_t^Y is average productivity among young firms and A_t^O is average productivity among old firms. In addition, because all young firms hire young entrepreneurs (who, by Lemma 1, choose $s = \sigma$), and because a fraction λ of those are high skill, it follows that $A_t^Y = \sigma(\eta \bar{A}_{t-1} + \lambda \gamma A_{t-1})$.

Average productivity among old firms depends on whether we have $R = 1$ or $R = 0$. With $R = 1$, all entrepreneurs are retained; a fraction λ are high ability, and all old entrepreneurs choose $s = 1$, so

$$A_t^O[R = 1] = \eta \bar{A}_{t-1} + \lambda \gamma A_{t-1}.$$

On the other hand, if $R = 0$ then only a fraction λ of the entrepreneurs (those revealed to be of high skill) are retained, and the remaining $1 - \lambda$ are replaced by young entrepreneurs. Hence, in this case

$$A_t^O[R = 0] = \lambda(\eta \bar{A}_{t-1} + \gamma A_{t-1}) + (1 - \lambda)\sigma(\eta \bar{A}_{t-1} + \lambda \gamma A_{t-1}).$$

Combining the definitions for A_t^Y, A_t^O and a_t (from (16)) and using the fact that $\bar{A}_t$ grows at the rate g, we obtain

$$a_t = \begin{cases} \frac{1+\sigma}{2(1+g)}[\eta + \lambda \gamma a_{t-1}] & \text{if } R_t = 1, \\ \frac{1}{2(1+g)}[(\lambda + \sigma + (1-\lambda)\sigma)\eta + (1 + \sigma + (1-\lambda)\sigma)\lambda \gamma a_{t-1}] & \text{if } R_t = 0. \end{cases}$$

$$(23)$$

This equation, which is also depicted in Figure 2, shows that the economy with $R_t = 1$ achieves greater growth (higher level of a_t for given a_{t-1}) through the imitation/adoption channel, as captured by the fact that $(1 + \sigma)\eta > (\lambda + \sigma + (1 - \lambda)\sigma)\eta$. However, it also achieves lower growth through the innovation channel, because

$$(1 + \sigma)\lambda \gamma a_{t-1} < (1 + \sigma + (1 - \lambda)\sigma)\lambda \gamma a_{t-1}.$$

In light of this observation, we can think of an equilibrium with $R_t = 1$ as corresponding to an *investment-based strategy*, where firms undertake greater investments even at the expense of sacrificing entrepreneurial selection. This
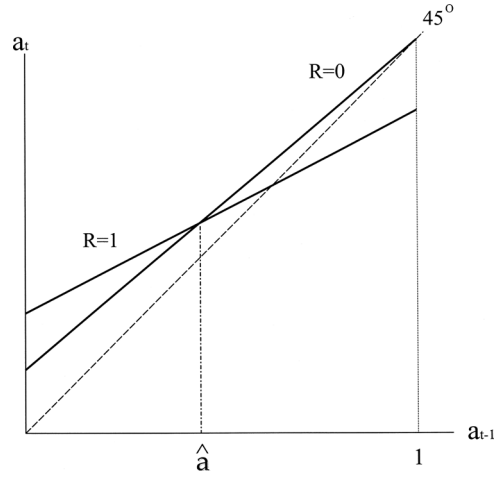
FIGURE 2. Technological convergence under investment-based strategy ($R = 1$) and innovation-based strategy ($R = 0$).

strategy involves longer-term relationships (entrepreneurs are never terminated) and the protection of older entrepreneurs from the competition of younger ones. In contrast, with $R_t = 0$ we can think of the economy as pursuing an *innovation-based strategy*, where there is greater selection of entrepreneurs and where the emphasis is on maximizing innovation at the expense of investment. Consequently, the innovation-based strategy results in a more "competitive" environment in which unsuccessful entrepreneurs are terminated and only successful entrepreneurs are retained.

The full equilibrium is then simply determined by combining this with the equilibrium law of motion, (23), which, by Proposition 1, can be written as

$$
a_t = \begin{cases} \frac{1+\sigma}{2(1+g)}(\eta + \lambda\gamma a_{t-1}) & \text{if } a_{t-1} \leq a_r(\mu, \delta) \\[2mm] \frac{1}{2(1+g)}\begin{bmatrix}(\lambda + \sigma + (1-\lambda)\sigma)\eta \\ +(1+\sigma + (1-\lambda)\sigma)\lambda\gamma a_{t-1}\end{bmatrix} & \text{if } a_{t-1} > a_r(\mu, \delta) \end{cases} . \quad (24)
$$

Figure 3 depicts the equilibrium dynamics. As (24) shows, equilibrium dynamics are given by a piecewise linear first-order difference equation. When $a_{t-1} \leq a_r(\mu, \delta)$, the economy pursues the investment-based strategy, but if a_{t-1} exceeds $a_r(\mu, \delta)$, then the economy switches to the steeper line, which corresponds to the innovation-based strategy.

The figure shows the possibility of a *non convergence* trap, where an economy stops converging to the frontier. To elaborate on this further, let us first characterize the world growth rate. It is plausible to assume that the growth rate of the technology frontier is determined endogenously by the most advanced economy
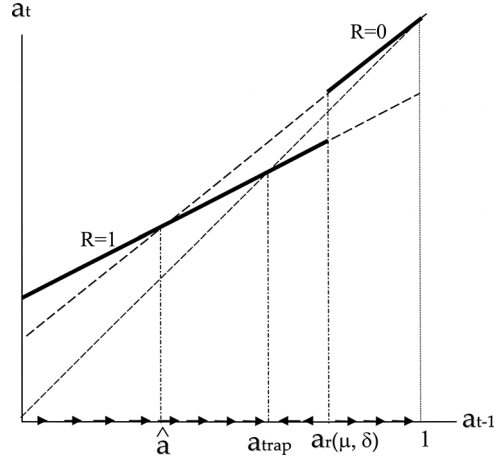
FIGURE 3. Equilibrium dynamics with non-convergence trap.

in the world pursuing the innovation-based strategy. Then, equation (24) evaluated at $a = 1$ gives the world technology growth rate as

$$g = \frac{1}{2}[(\lambda + \sigma + (1 - \lambda)\sigma)\eta + (1 + \sigma + (1 - \lambda)\sigma)\lambda\gamma] - 1, \qquad (25)$$

which we assume to be positive. In addition, for the innovation-based strategy to generate higher growth than the investment-based strategy at the frontier, $a = 1$, we need

$$\frac{(1 - \sigma)\eta}{\sigma} < \lambda\gamma. \qquad (26)$$

Consequently, at $a = 1$, the $R = 0$ line intersects the 45° line and is above the $R = 1$ line. But then, as drawn in Figure 3, the $R = 1$ line must intersect the 45° line at some $a_{trap} < 1$. From (23), this threshold value can be calculated as

$$a_{trap} = \frac{(1 + \sigma)\eta}{2(1 + g) - \lambda\gamma(1 + \sigma)}. \qquad (27)$$

If the economy is pursuing the investment-based strategy when it reaches $a = a_{trap}$, then it will stay there forever. In other words, it will have fallen into a non convergence trap.

In practice, however, an economy may switch out of the investment-based strategy before a_{trap} is reached. Therefore, the necessary and sufficient condition for an equilibrium non convergence trap is
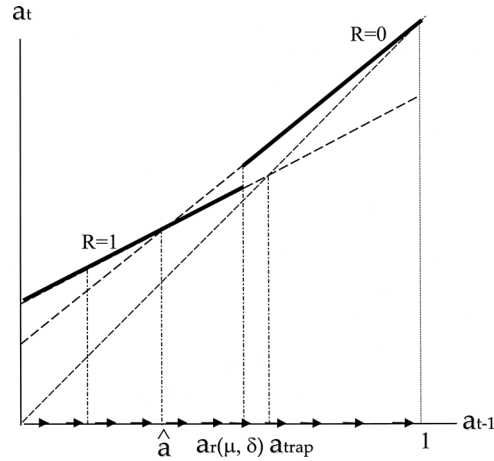
$$a_{trap} < a_r(\mu, \delta),$$

FIGURE 4. Equilibrium dynamics without non-convergence trap.

which corresponds to the case depicted in Figure 3. In contrast, Figure 4 shows the case where $a_{trap} > a_r(\mu, \delta)$; here the economy switches out of the investment-based strategy before a_{trap} is reached, and the non convergence trap does not arise.

When is this condition likely to be satisfied? From (27), a_{trap} is an increasing function of $\lambda\gamma$ and is independent of $\kappa/\delta N$ and μ. Because $a_r(\mu, \delta)$ is a decreasing function of $\kappa/\delta N$ and of $\lambda\gamma$, smaller values of $\kappa/\delta N$ and $\lambda\gamma$ make it more likely that $a_{trap} < a_r(\mu, \delta)$. Furthermore, if condition (22) holds, then traps are more likely in economies with severe incentive problems/credit market imperfections. These comparative statics are intuitive. First, smaller values of κ and greater values of δN make the retention of low-skill entrepreneurs more likely. Because a trap can only arise due to excess retention, a greater $\kappa/\delta N$ reduces the likelihood of traps. Second, large values of $\lambda\gamma$ increase the opportunity cost of employing low-skill entrepreneurs and make it less likely that a trap can emerge due to lack of selection. Finally, when condition (22) holds, more severe credit market imperfections (incentive problems) favor insiders by raising retained earnings and increase the likelihood of a non convergence trap.

The next proposition summarizes the equilibrium dynamics.

PROPOSITION 2.   *Suppose Assumption 1 holds and $\delta \in (\delta_L, \delta_H)$. Let $a_r(\mu, \delta)$ and a_{trap} be defined by (21) and (27), and denote the initial distance to frontier by a_0. Then the unique dynamic equilibrium is as follows:*

1. *If $a_0 < a_r(\mu, \delta)$ and $a_{trap} \geq a_r(\mu, \delta)$, then the economy starts with the investment-based strategy, switches to the innovation-based strategy at $a = a_r(\mu, \delta)$, and converges to the world technology frontier, $a = 1$.*

2. *If $a_0 < a_r(\mu, \delta)$ and $a_{trap} < a_r(\mu, \delta)$, then the economy starts with the investment-based strategy and converges towards the world technology frontier until it reaches $a = a_{trap} < 1$, where both convergence and the growth of a_t stop.*

3. *If $a_r(\mu, \delta) \leq a_0$, then the economy starts with the innovation-based strategy and converges to the world technology frontier, $a = 1$.*

4.4. Growth-maximizing Strategies and a Theory of Leapfrogging

In this section, we analyze the growth implications of different development strategies. We first characterize the growth-maximizing strategy. Clearly, growth maximization is not the correct criterion for welfare comparisons, because it ignores the cost of investments. Nevertheless, it is the most appropriate way to derive implications of the theory that are comparable with the evidence presented in Section 2. In Appendix C, we characterize the welfare-maximizing strategies and show that the comparison of those to the equilibrium is very similar to the comparison of the growth-maximizing strategy to the equilibrium.

Inspection of (23) or of Figure 2 immediately shows that growth will be maximized when the economy reaches the highest level of a_t for a given a_{t-1}. This is attained by pursuing the strategy of $R = 1$ whenever $a_{t-1} < \hat{a}$ and pursuing the innovation-based strategy, $R = 0$, whenever $a_{t-1} > \hat{a}$, where $\hat{a}$ is given by the intersection of the $R = 0$ and $R = 1$ lines in Figure 4 or by

$$\hat{a} \equiv \frac{\eta(1 - \sigma)}{\lambda \gamma \sigma}. \tag{28}$$

Condition (26) ensures that $\hat{a} < 1$. Therefore, similar to the case of equilibrium behavior, the growth-maximizing sequence also starts with the investment-based strategy and then switches to an innovation-based strategy. But the switch does not necessarily occur at the same point as in the equilibrium case.

How does $\hat{a}$ compare to the equilibrium threshold $a_r(\mu, \delta)$? The answer depends on (among other things) the degree of competition as measured by δ. The appropriability effect discussed previously means that equilibrium behavior is biased against the investment-based strategy, creating a force toward $a_r(\mu, \delta) < \hat{a}$. However, countering this is the *rent-shield* effect: Retained earnings are used to finance part of the investment costs, creating a transfer to the capitalists and shielding old entrepreneurs from the competition of young entrepreneurs. In other words, while the appropriability effect creates a bias (relative to the growth-maximizing allocation) against the investment-based strategy, retained earnings (rents) of the insiders protect them from competition and create a bias in favor of the investment-based strategy.

Which effect dominates is ambiguous. A greater δ increases $a_r(\mu, \delta)$ relative to $\hat{a}$ (which does not depend on δ), but this could either increase or decrease the gap between the equilibrium and the growth-maximizing allocations depending on whether we start from a situation where $\hat{a} > a_r(\mu, \delta)$ or $\hat{a} < a_r(\mu, \delta)$, respectively. Given μ, there exists a unique level of competition δ, denoted by $\widehat{\delta}(\mu)$, such that $\hat{a} = a_r(\mu, \widehat{\delta}(\mu))$, where[13]

$$\widehat{\delta}(\mu) = \frac{\kappa}{N} \frac{1 - \phi}{\frac{1+r}{1+g}\mu\sigma\eta}.$$

If $\delta > \widehat{\delta}(\mu)$ that is, product market competition is low, then $\hat{a} < a_r(\mu, \delta)$ and there is an excessive retention of low-skill entrepreneurs relative to the growth-maximizing allocation. In this case, which is the one shown in Figure 3, limiting competition (i.e., increase δ) would further increase the growth gap between the equilibrium and the growth-maximizing strategy. Conversely, if $\delta < \widehat{\delta}(\mu)$, namely, product market competition is high, then $\hat{a} > a_r(\mu, \delta)$ as shown in Figure 4. In this case, the economy switches out of the investment-based strategy too quickly, and limiting competition would increase growth in the range where $a_t \in (a_r(\mu, \delta), \hat{a})$.

One implication of the foregoing discussion is that less competitive environments may foster growth at early stages of development (far from the technology frontier). For example, starting with an economy featuring $\hat{a} > a_r(\mu, \delta)$ and $a_{t-1} \in (a_r(\mu, \delta), \hat{a})$, an increase in δ (a reduction in competition) may induce the investment-based strategy in this range and secure more rapid growth. However, our previous discussion of non convergence traps also underscored that limiting product market competition may later become harmful to growth, preventing convergence to the frontier. In particular, there exists a threshold competition level $\delta^*(\mu)$, defined by

$$a_r[\mu, \delta^*(\mu)] = a_{trap}, \tag{29}$$

such that an economy with $\delta < \delta^*(\mu)$ will never fall into a non convergence trap. Therefore, competitive markets may slow down technological convergence at the earlier stages of development, but this does not affect the long-run equilibrium. Low competition, on the other hand, may have detrimental effects in the long run.

Our analysis thus leads to a new theory of "leapfrogging." Imagine two economies that start with the same distance to frontier, a_{t-1}, but differ in terms of their competitive policies, δ_1 and δ_2, with $a_r(\mu, \delta_1) < a_{t-1} < \hat{a} < a_r(\mu, \delta_2)$. Given this configuration, economy 1 will pursue the innovation-based strategy,

---

13. Note that $\widehat{\delta}(\mu) > \delta_L$. Moreover, if $(1 - \mu)(1 - \sigma)(\eta + \lambda\gamma) > (1 + r)\mu\sigma\eta/(1 + g)$, then $\widehat{\delta}(\mu) < \delta_H$, so the restriction $\delta \in (\delta_L, \delta_H)$ in Lemma 1 does not preclude the possibility that changes in δ can either augment or reduce growth.

while economy 2 will start with the investment-based strategy and initially grows faster than economy 1. However, once these economies pass beyond $\hat{a}$, economy 1 starts growing more rapidly, because economy 2 still pursues the investment-based strategy (despite the fact that growth is now maximized via the innovation-based strategy). Furthermore, if $a_{trap} < a_r(\mu, \delta_2)$, then economy 2 will get stuck in a non convergence trap before it can switch to the innovation-based strategy, and will be leapfrogged by economy 1, which avoids the non convergence trap and converges to the frontier. This result further illustrates the claim made in the Introduction that those rigid institutions associated with the less-competitive market structures supporting the investment-based strategy become more costly as an economy approaches the world technology frontier. It may also shed some light on why some economies (e.g., Brazil, Mexico, Peru) that initially grew relatively rapidly with highly protectionist policies were then overtaken by economies (e.g., Hong Kong, Singapore) adopting more competitive policies.[14]

5. Policy and Political Economy Traps

5.1. Policy and Appropriate Institutions

The analysis so far has established a number of results. First, the dynamic equilibrium typically starts with the investment-based regime, which features high investment and long-term relationships. As the economy approaches the world technology frontier, this is followed by a switch to an innovation-based regime featuring lower investment, younger firms, and more selection. Second, if there is no switch to the innovation-based regime, the economy will get stuck in a non convergence trap and will not converge to the frontier. Finally, for some parameter values far from the world technology frontier, the growth rate can be increased by inducing the economy to stay longer in the investment-based regime.

   This last observation raises the possibility of useful policy interventions along the lines suggested by Gerschenkron: Governments in relatively backward economies can intervene to increase investment and to induce faster adoption of existing technologies. However, the second observation points out that this type of intervention may have long-run costs unless it is abandoned at later stages of development. In this section, we start with a brief discussion of possible policies to foster growth, which can be interpreted as corresponding to "appropriate" institutions for countries at different stages of development (in the sense that they are useful only at specific stages of development).

   Consider an equilibrium allocation with $a_r(\mu, \delta) < \hat{a}$, where the economy switches out of the investment-based strategy before the growth-maximizing

---

14. Before 1967, the growth of GDP per worker was indeed slower in Singapore (2.6% per year) than in both Mexico (3.9%) and Peru (5.3%). This ranking reversed in the 1970s and 1980s.

threshold. A policy intervention that encourages greater investment will increase growth over the range $a \in (a_r(\mu, \delta), \hat{a})$.[15] A number of different policies can be used for this purpose. Probably the most straightforward is an investment subsidy, which might take the form of direct subsidies or preferential loans at low interest rates. Suppose the government subsidizes a fraction τ of the cost of investment. Analyzing as before, we analogously derive the threshold for switching from the investment- to the innovation-based strategy:

$$\tilde{a}_r(\mu, \delta, \tau) \equiv \frac{\left[(1 - \mu)(1 - \sigma) + \frac{1+r}{1+g}\mu\sigma\right]\eta - \frac{(1-\phi)\kappa(1-\tau)}{\delta N}}{(1 - \mu)\sigma\lambda\gamma}.$$

If τ is chosen appropriately—in particular, if $\tau = \tilde{\tau}$ such that $\tilde{a}_r(\mu, \delta, \tilde{\tau}) = \hat{a}$— then the economy can be induced to switch out of the investment-based strategy exactly at $\hat{a}$ (or at some other desired threshold, if the government is pursuing a different objective).

Investment subsidies are difficult to implement, however, especially in relatively backward economies where tax revenues are scarce. Furthermore, it may be difficult for the government to observe and monitor the exact level of investment undertaken by firms. For this reason, we focus on another potential policy instrument that affects the equilibrium threshold $a_r(\mu, \delta)$: the extent of such anticompetitive policies as entry barriers, merger policies, and so forth. Naturally, this discussion also applies to investment subsidies.

Anticompetitive policies are captured by the parameter χ in our model. We recall that δ is monotonically increasing in χ, so high values of χ or δ correspond to a less competitive environment. Starting from a situation where $a_r(\mu, \delta) < \hat{a}$, anticompetitive policies close the gap between the equilibrium threshold and the growth-maximizing threshold. Although restricting competition creates static losses (recall equation (6)), in the absence of feasible tax/subsidy policies, this may be the best option available to encourage faster growth and technological convergence.

When the government chooses a less competitive environment in a backward economy in order to encourage long-term relationships, greater investment, and faster technological convergence the situation is reminiscent of Gerschenkron's analysis. It is also related to the well-known "infant industry" arguments calling for protection and government support for certain industries at early stages of their development. But our analysis also reveals that anticompetitive policies (and similarly investment subsidies) become harmful for economies closer to the world technology frontier. Institutions that are appropriate for early stages of

---

15. The analysis in Appendix C also shows that, with μ or δ sufficiently small, we can also have $a_r(\mu, \delta)$ less than the threshold at which a welfare-maximizing social planner would choose to switch from the investment- to the innovation-based strategy, so this discussion could be carried out in terms of policies to encourage welfare (rather than growth) maximization.

development therefore become inappropriate for an economy close to the frontier. An economy that adopts such institutions must later abandon them; otherwise, it will end up in a non convergence trap.

Yet a sequence of policies whereby certain interventions are first adopted and then abandoned does raise important political economy considerations. Groups that benefit from anticompetitive policies will become richer under these policies and so will oppose a change in policy. To the extent that economic power buys political power, say, via lobbying, these groups can be quite influential in opposing such changes. Therefore, the introduction of appropriate institutions to foster growth also raises the possibility of political economy traps: Groups enriched by these institutions successfully block reform, and the economy ends up in a nonconvergence trap because it adopted but could not abandon institutions appropriate for an early stage of development. This is the subject of the following section.

5.2. *Political Economy and Traps*

We now describe a simple political economy example where special interest groups may capture politicians and lead the economy into a "political economy trap." Our example is a much-simplified version of the special interest group model of Grossman and Helpman (1994, 2001), extended to include a link between economic power and political influence (on this, see also Do 2002).[16]

Suppose that competition policy χ is determined in each period by a politician (or government) that cares about the agents' welfare but is also sensitive to bribes or to campaign contributions. For tractability, we adopt a simple setup: politicians at time t can be bribed to affect policies at time $t + 1$. The politician's pay-off is equal to hA_{t-1} (where $h > 0$) if she behaves honestly and chooses the policy that maximizes current consumption. Else, his pay-off is B_t, where B denotes a monetary bribe the politician might receive in order to pursue a different strategy. The utility of pursuing the right policy is assumed to be linearly increasing in A_{t-1}, which ensures stationary policies in equilibrium (because bribes will be increasing in A).[17]

We assume that agents cannot borrow to pay bribes and that only capitalists can organize lobbies. Moreover, young agents have no wealth, so they cannot

---

16.   We limit our analysis to a particular case where the equilibrium sequence illustrates the possibility of political economy traps. Details and the more comprehensive analysis are in the NBER working paper version of this paper.

17.   In this formulation, the honesty parameter H can be interpreted as a measure of the aggregate welfare concerns of politicians, or more interestingly, as the quality of the system of checks and balances that limits the ability of special interest groups to capture politicians. This formulation is similar to that in Grossman and Helpman (1994, 2001) but simpler because in their model, the utility that the politician gets from adopting various policies is a continuous function of the distance from the ideal policy. As in their setup, the politician can commit to deliver the competition policy promised to an interest group in return for bribes.

bribe politicians. Hence the only group with the capability to bribe politicians is old capitalists. We also assume that the institutional choice facing the politician is binary: low versus high competition.

More formally, the politician chooses $\chi_t \in \{\underline{\chi}, \bar{\chi}\}$ where $\underline{\chi} < \bar{\chi} \leq 1/\alpha$. By analogy, we set $\delta_t \equiv (\chi_t - 1)\chi_t^{-1/(1-\alpha)} \in \{\underline{\delta}, \bar{\delta}\}$, which, we recall, is the parameter in the profit function, (3). The assumption that χ is a discrete rather than a continuous choice variable is for simplicity. Because we want to emphasize the possibility of traps induced by political economy factors, we choose $\underline{\delta}$ and $\bar{\delta}$ such that $\delta^*(\mu) \in (\underline{\delta}, \bar{\delta})$: hence an economy choosing $\bar{\delta}$ forever converges to a non convergence trap, whereas an economy choosing forever $\underline{\delta}$ converges to the frontier.

Clearly, capitalists always prefer low to high competition, because this increases their profits. Under some conditions, old capitalists will be willing to pay the entire first-period revenue to increase their monopoly power. We focus on parameter values where these conditions hold, so that capitalists are credit constrained.[18] This implies that they will be able to bribe the politicians as long as the maximum bribe they can pay which is given by their first-period earnings exceeds the cost of buying politics. Capitalists' first-period earnings, on the other hand, depend on δ_{t-1}, the competition policy at $t-1$. Therefore, capitalists will successfully bribe politicians when

$$\delta_{t-1}(1 - \mu)\sigma N(\eta + \lambda\gamma a_{t-1}) \geq h a_{t-1}, \tag{30}$$

where the left-hand side is the first-period earnings of capitalists and thus the maximum bribes they can pay.

It is evident that when δ_{t-1} is greater, (30) is more likely to hold; in less competitive markets, capitalists have higher profits that they can use for bribing politicians. Let us define a_L and a_H as the unique values of a_{t-1} such that (30) holds with equality for $\delta_{t-1} = \bar{\delta}$ and $\delta_{t-1} = \underline{\delta}$, respectively. Thus

$$a_L \equiv \frac{\bar{\delta}(1 - \mu)\sigma N\eta}{h - \lambda\gamma\bar{\delta}(1 - \mu)\sigma N} > a_H \equiv \frac{\underline{\delta}(1 - \mu)\sigma N\eta}{h - \lambda\gamma\underline{\delta}(1 - \mu)\sigma N}. \tag{31}$$

This discussion immediately implies that politicians will be bribed to maintain the anticompetitive policy $\bar{\delta}$ as long as $a_{t-1} \leq a_L$, which ensures that the left-hand side of (30) is greater than the right-hand inside. Similarly, politicians will be bribed to switch from a procompetitive policy $\delta_{t-1} = \underline{\delta}$ to an anticompetitive policy $\bar{\delta}$ when $a_{t-1} \leq a_H$. That $a_L > a_H$ follows because capitalists

---

18. This is in the spirit of capturing the notion that economic and political power are related. If capitalists were not credit constrained this link would be absent in our model. See the NBER Working paper for the conditions to ensure that capitalists are credit constrained. There, we also discuss the case in which politicians may choose anticompetitive policies even without bribes (because this would be the welfare-maximizing policy). Here, we focus on the case where such range is empty and without bribes, the politician would always choose high competition.

make greater profits with low competition and have more funds to bribe politicians. This formalizes the idea that, once capitalists become economically more powerful, they also become politically more influential and consequently more likely to secure the policy that they prefer. Note that both the cutoffs a_L and a_H are decreasing functions of h, because politicians who are more honest will be harder to bribe into pursuing the policy preferred by the capitalist lobby.

Finally, in order to focus on an interesting case, let us further assume that $a_{trap} \in (a_H, a_L)$. In this case, the equilibrium sequence has the following features:

1. If $a_0 < a_H$ then the economy starts at a point where the capitalists are sufficiently rich that, even when $\delta_{t-1} = \underline{\delta}$, they will have enough to bribe politicians successfully. Consequently, as long as $a < a_{trap}$ there will be growth with anticompetitive policies, but the economy will eventually converge to the non convergence trap, a_{trap}. We refer to this situation as a political economy trap, because politicians, once in it, always receive bribes and adopt policies that prevent a switch from the investment-based to the innovation-based strategy.

2. If $a_0 > a_L$, then capitalists have insufficient funds to bribe politicians even when $\delta_{t-1} = \bar{\delta}$. Hence, the policy is procompetitive and the economy switches to an innovation-based strategy and therefore converges to the world technology frontier, $a = 1$.

3. Finally, if $a_0 \in (a_H, a_L)$, then the long-run outcome is *history dependent*. If competition is initially low, $\delta_{t-1} = \bar{\delta}$, capitalists enjoy greater monopoly profits and are sufficiently wealthy to lobby successfully to maintain the anticompetitive policies ($\delta_{t-1} = \bar{\delta}$) in place; the equilibrium is identical to case 1 and eventually ends up in a non convergence trap. However, if competition is initially high ($\delta_{t-1} = \underline{\delta}$), capitalists make lower profits and do not have enough purchasing power to buy politicians. Consequently, in this case, there is no effective lobbying activity, procompetitive policies are never abandoned, and the economy converges to $a = 1$.

In conclusion, when our theory is augmented with a stylized political process describing the possible capture of politicians by lobbies, it predicts multiple steady state political equilibria. One of these eventually leads to procompetitive policies and an innovation-based equilibrium, ensuring convergence to the world technology frontier; the other leads to a political economy trap, with investment-based strategy throughout and no convergence to the frontier.

Two other points are noteworthy. First, because a_H and a_L are decreasing in h, the analysis demonstrates that political economy traps are more likely when there are few checks and balances on politicians. Second, it also suggests that potentially well-meaning government policy directed at improving the short-run allocation of resources (e.g., by encouraging the investment-based strategy) may have unintended adverse long-run consequences.

6. Conclusion

In this paper we have proposed a growth model where firms engage in copying and adopting technologies from the world frontier and also in innovation activities. The closer an economy is to the world technology frontier, the higher the relative importance of innovation relative to imitation as a source of productivity growth. Because the selection of high-skill entrepreneurs and firms is more important for innovation than for adoption, firms in countries that are far from the technology frontier pursue an investment-based strategy, which features long-term relationships, high average size and age of firms, large investments, but little selection. Closer to the technology frontier, there is less room for copying and adoption of well-established technologies; consequently, there is an equilibrium switch to an innovation-based strategy with short-term relationships, younger firms, less investment, and better selection of entrepreneurs.

We showed that economies may switch out of the investment-based strategy too soon or too late. A standard appropriability effect, resulting from the fact that firms do not internalize the greater consumer surplus they create by investing more, implies that the switch may occur too soon. In contrast, the presence of retained earnings that incumbent entrepreneurs can use to shield themselves from competition makes the investment-based strategy persist for too long. When the switch is too soon, government intervention in the form of policies limiting product market competition or providing subsidies to investment may be useful because it encourages the investment-based strategy. Nevertheless, anticompetitive policies can also lead to a non convergence trap whereby the economy never switches out of the investment-based strategy and fails to converge to the world technology frontier.

Even though much of the emphasis in this paper is on cross-country comparisons, the same reasoning also extends to cross-industry comparisons. In particular, our analysis suggests that the organization of firms and of production should be different in industries that are closer to the world technology frontier. More generally, cross-industry differences in the internal organization of the firm constitute an interesting and relatively underexplored area for future research.

Appendix A The Participation Constraint

In this appendix we prove that if $\sigma < (1+r)^{-1}$ (sufficient condition) and N is sufficiently large, then the incentive compatibility constraint (11) is always more binding than the relevant participation constraints.

Let us denote $IC_t(v \mid s, e, z) \equiv S_t(v \mid s, e, z) - \mu\pi_t(v \mid s, e, z) \geq 0$, where $S_t(v \mid s, e, z)$ is the payment to the entrepreneur and the inequality corresponds to the incentive compatibility condition. The participation constraint for an old entrepreneur is

$$PC_t(v \mid s, e = O, z) \equiv S_t(v \mid s, e = O, z) - \widehat{RE}_t(v \mid s, e = O, z) - w_t \geq 0,$$
$$(A.1)$$

which simply states that the payments minus retained earnings that are injected must be greater than the wage rate. We will ensure that these participation constraints hold even when all entrepreneurs inject all their retained earnings, that is, when $\widehat{RE}_t(v \mid s, e = O, z) = RE_t(v \mid s, e = O, z)$.

The participation constraint for a young entrepreneur is slightly more involved, because he anticipates potential rents if he remains an entrepreneur in the future. We can write this constraint as

$$PC_t(v \mid s, e = Y) \equiv S_t(v \mid s, e = Y) + \frac{1}{1+r} E_t \text{Rent}_{t+1} - w_t \geq 0; \quad (A.2)$$

here the expected future rent is given by

$$E_t \text{Rent}_{t+1} = \lambda\, PC_{t+1}(v \mid s, e = O, z = H)$$
$$+ (1 - \lambda)\, R_t\, PC_{t+1}(v \mid s, e = O, z = L),$$

which uses the fact that future rents correspond to the future participation constraints being slack. This expression also takes into account that the entrepreneur is uncertain about his own type and will receive future rents when he has high skill or when he has low skill and the economy is in the investment-based regime $R_t = 1$.

We prove the main result in two steps. First, we prove that there exists an $N_Y < \infty$ such that, for $N \geq N_Y$, the participation constraint is slack for young entrepreneurs. Second, we prove that there exists an $N_O < \infty$ such that, for $N \geq N_O$, the participation constraint is slack for both low- and high-skill old entrepreneurs. Therefore, if $N \geq \max\{N_Y, N_{OL}, N_{OY}\}$, then the participation constraints both for young and old entrepreneurs are slack.

For the first step, observe that because the young have no retained earnings, a sufficient condition for the participation constraint not to bind when the incentive constraint binds is that $\mu \pi_t(v \mid \sigma, Y, L) \geq w_t$. Using equation (3) for equilibrium profits along with the equilibrium wage equation (5), we can re-express this participation constraint (A.2) as

$$\mu \delta \sigma N(\eta \bar{A}_{t-1} + \lambda \gamma A_{t-1}) + \frac{1}{1+r} E_t \text{Rent}_{t+1} \geq \left((1-\alpha)\alpha^{-1}\chi^{-\alpha/(1-\alpha)}\right) A_t.$$
$$(A.3)$$

Because $E_t \text{Rent} > 0$ and $A_t \leq (1+g)\bar{A}_{t-1}$, a sufficient condition for (A.3) to hold for all a is

$$N \geq \frac{(1+g)(1-\alpha)\alpha^{-1}\chi^{-\alpha/(1-\alpha)}}{\mu \delta \sigma \eta} \equiv N_Y. \quad (A.4)$$

To establish the second step, we note that the participation constraints of old low-skill entrepreneurs is slack if and only if $\mu\pi_t(\nu \mid 1, O, L) - RE_t(\nu \mid 1, O, L) \geq w_t$. Substituting for π_t, w_t and RE_t, we can re-express this condition as

$$\mu\delta N \left(\eta\bar{A}_{t-1} - (1+r)\sigma\eta\bar{A}_{t-2}\right) \geq (1-\alpha)\alpha^{-1}\chi^{-\alpha/(1-\alpha)}A_t. \quad \text{(A.5)}$$

Similarly, the participation constraints of old high-skill entrepreneurs is slack if and only if $\mu\pi_t(\nu \mid 1, O, H) - RE_t(\nu \mid 1, O, H) \geq w_t$. Substituting for π_t, w_t, and RE_t, we can re-express this condition as

$$\mu\delta N\big[\left(\eta\bar{A}_{t-1} + \lambda\gamma A_{t-1}\right) - (1+r)\sigma\left(\eta\bar{A}_{t-2} + \lambda\gamma A_{t-2}\right)\big]$$
$$\geq (1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}A_t. \quad \text{(A.6)}$$

First, note that a sufficient condition for the LHS of both inequalities (A.5) and (A.6) to be positive is that $\sigma < 1/(1+r)$ (because $\bar{A}_{t-1} > \bar{A}_{t-2}$ and $A_{t-1} > A_{t-2}$, obviously $\eta\bar{A}_{t-1} > \eta\bar{A}_{t-2}$ and $\eta\bar{A}_{t-1} + \lambda\gamma A_{t-1} > \eta\bar{A}_{t-2} + \lambda\gamma A_{t-2}$). Second, as long as $\sigma < 1/(1+r)$ there exists an $N_O < \infty$ such that, if $N \geq N_O$, then both (A.5) and (A.6) hold, that is, the participation constraint is slack for both the low-skill and high-skill old entrepreneurs.

Therefore, if $N \geq \max\{N_Y, N_O\}$ then both participation constraints are slack, even when entrepreneurs inject all their retained earnings.

Appendix B Proof of Lemma 1

Let

$$\bar{\phi} \equiv \frac{\frac{1+r}{1+g}\mu\sigma\eta - (1-\mu)(1-\sigma)\lambda\gamma}{\frac{1+r}{1+g}\mu\sigma\eta + (1-\mu)(1-\sigma)\eta}. \quad \text{(B.1)}$$

It is immediate to verify that if $\phi < \bar{\phi}$ then $\delta_H > \delta_N$ (in particular, note that δ_H is decreasing in ϕ and that $\delta_H = \delta_N$ if and only if $\phi = \bar{\phi}$). We first show that if $\delta < \delta_H$ then, for all a, young entrepreneurs are assigned to small projects. More formally, $\delta < \delta_H$, implies that, for all a,

$$E_t V_t(\nu \mid s = \sigma, e = Y) = \big[(1-\mu)\sigma\delta N(\eta + \lambda\gamma a_{t-1}) - \phi\kappa\big]\bar{A}_{t-1}$$
$$\geq \big[(1-\mu)\delta N(\eta + \lambda\gamma a_{t-1}) - \kappa\big]\bar{A}_{t-1} = E_t V_t(\nu \mid s = 1, e = Y).$$

Because the right-hand side increases in a_{t-1} faster than the left-hand side, it suffices to show this inequality for $a_{t-1} = 1$. Letting

$$\delta \leq \frac{\kappa}{N}\frac{1-\phi}{(1-\mu)(1-\sigma)(\eta + \lambda\gamma)} \equiv \delta_H$$

ensures that this is so.

Next we establish that the retained earnings of an old low-skill entrepreneur at t are as given by (17). Recall first that we have assumed participation constraints to be slack even when entrepreneurs inject all their retained earnings. This, plus the fact that capitalists make the contract offers, imply that old low-skill entrepreneurs inject all their earnings, namely, $\widehat{RE}_t = RE_t$ (see Appendix A). Retained earnings are therefore equal to the capitalized first-period entrepreneurial earnings. We have shown above that all young entrepreneurs run small projects, so a low-skill entrepreneur born at $t-1$ will have a level of retained earnings equal to $\mu \delta \sigma N \eta \bar{A}_{t-2}$. Equation (7) implies that $\bar{A}_{t-1} = \bar{A}_{t-2}/(1+g)$, and adding the interest payments at the rate r establishes that the retained earnings of an old low-skill entrepreneur at time t are as given by (17).

Now we show that, if $\delta > \delta_L$, then old entrepreneurs operate large projects. We prove this to be true for a low-skill entrepreneur. Then, *a fortiori*, it must be true for a high-skill entrepreneur. An old low-skill entrepreneur operates a large project if and only if

$$
\begin{aligned}
V_t(\nu \mid s=1, e=O, z=L) \\
= \left[(1-\mu)\delta N\eta - \max\left\{ \kappa - \frac{1+r}{1+g}\mu\delta N\sigma\eta, 0 \right\} \right] \bar{A}_{t-1} \\
\geq V_t(\nu \mid s=1, e=O, z=L) \\
= \left[(1-\mu)\sigma\delta N\eta - \max\left\{ \kappa - \frac{1+r}{1+g}\mu\delta N\sigma\eta, 0 \right\} \right] \bar{A}_{t-1}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
(1-\mu)(1-\sigma)\delta N\eta \geq &\ \max\left\{ \kappa - \frac{1+r}{1+g}\mu\delta N\sigma\eta, 0 \right\} \\
&- \max\left\{ \phi\kappa - \frac{1+r}{1+g}\mu\delta N\sigma\eta, 0 \right\}. \qquad \text{(B.2)}
\end{aligned}
$$

We first show that, if $\phi < \bar{\phi}$ and $\delta \geq \delta_L$, then $\phi\kappa \leq \mu\delta N\sigma\eta(1+r)/(1+g)$. In order to derive a contradiction, we suppose that $\phi\kappa > \mu\delta N\sigma\eta(1+r)/(1+g)$ for some $\delta \geq \delta_L$. Let $\delta = \delta_L$. Then

$$
\phi\kappa > \frac{1+r}{1+g}\mu\delta_N N\sigma\eta = \frac{\frac{1+r}{1+g}\mu\sigma\eta}{(1-\mu)(1-\sigma)\eta + \frac{1+r}{1+g}\mu\sigma\eta}\kappa > \bar{\phi}\kappa,
$$

contradicting the assumption that $\phi < \bar{\phi}$. Hence, we have established that $\phi\kappa \leq \mu\delta_L N\sigma\eta(1+r)/(1+g)$, which immediately implies that $\phi\kappa \leq \mu\delta N\sigma\eta(1+r)/(1+g)$ for all $\delta \geq \delta_L$. Thus, (B.2) can be rewritten as

$$
(1-\mu)(1-\sigma)\delta N\eta \geq \max\left\{ \kappa - \frac{1+r}{1+g}\mu\delta N\sigma\eta, 0 \right\}.
$$

Because the left-hand side increases faster in δ than the right-hand side, ensuring that this inequality holds for $\delta = \delta_L$ is sufficient. Evaluating it at $\delta = \delta_L$ yields

$$\frac{(1-\mu)(1-\sigma)\eta}{(1-\mu)(1-\sigma)\eta + \frac{1+r}{1+g}\mu\sigma\eta}\kappa \geq \max\left\{\frac{(1-\mu)(1-\sigma)\eta}{(1-\mu)(1-\sigma)\eta + \frac{1+r}{1+g}\mu\sigma\eta}\kappa, 0\right\},$$

which is obviously true. Hence, for all $\delta > \delta_L$ we have

$$V_t(\nu \mid s = 1, e = O, z = L) \geq V_t(\nu \mid s = \sigma, e = O, z = L),$$

and

$$V_t(\nu \mid s = 1, e = O, z = H) \geq V_t(\nu \mid s = \sigma, e = O, z = H),$$

which concludes the proof of the lemma.

Appendix C Welfare Analysis

In this appendix we compare the equilibrium with the retention policy that maximizes social welfare. Consider a planner who maximizes the present discounted value of the consumption stream with a discount factor $\beta \equiv 1/(1 + r)$; that is, she maximizes $C_t + \sum_{j=1}^{\infty} \beta^t C_{t+j}$, where $C_t = \zeta N A_t - \int_0^1 k_t(\nu)d\nu$ is equal to net output minus investment at date t, with $\int_0^1 k_t(\nu)d\nu = \kappa \bar{A}_{t-1}/2$ if $R_t = 1$ and equal to 0 if $R_t = 0$. As before, we start with an allocation where prices $p_t(\nu)$ satisfy (2) and the wage rate w_t is given by (5), and we assume that Lemma 1 holds. The planner takes all decentralized decisions, including those regarding project size, as given as in Section 4, and only chooses R.

A useful benchmark is the choice of a "myopic planner" who puts no weight on future generations, (i.e., $\beta = 0$). The myopic planner chooses the retention policy at t so as to maximize total consumption at t, and he retains old low-skill entrepreneurs if and only if $a_{t-1} < a_{mfb}$, where the threshold a_{mfb} is such that $R_t = 0$ and $R_t = 1$ yield the same consumption, that is

$$a_{mfb} \equiv \frac{\eta(1-\sigma) - (1-\phi)\kappa/\zeta N}{\sigma\lambda\gamma}. \tag{C.1}$$

This threshold can be compared with the growth-maximizing threshold. Because the planner takes into account the cost of innovation, which is ignored by the growth-maximizing strategy, the myopic planner sets $a_{mfb} < \hat{a}$.

Now consider a non-myopic planner who cares also about future consumption (i.e., she has $\beta > 0$). She will realize that, by increasing the retention threshold on a_{mfb}, she can increase future consumption at the expense of current consumption. For any positive β and in particular for $\beta = 1/(1 + r)$, a small increase

of the threshold starting at a_{mfb} involves no first-order loss in current consumption, but it generates first-order gains in productivity (A_t) and in the present discounted value of future consumption. The non-myopic planner will therefore choose a threshold of $a_{fb} > a_{mfb}$. Moreover, we can see that a_{fb} cannot exceed the growth-maximizing threshold $\hat{a}$. Any candidate threshold larger than $\hat{a}$ (say, $\tilde{a} > \hat{a}$), can be improved upon, because any threshold in the range $(\tilde{a}, \hat{a}]$ increases both current and future consumption relative to $\tilde{a}$. Thus, the optimal threshold cannot be to the right of $\hat{a}$. In summary, we have

$$a_{mfb} < a_{fb} < \hat{a}.$$

Therefore, an economy with sufficiently high μ and δN will switch to an innovation-based strategy too late, because it would feature $a_r(\mu, \delta) > \hat{a}$. We can also verify that an economy with sufficiently small μ will switch to an innovation-based strategy ($R_t = 0$) too soon relative to the welfare-maximizing allocation; that is, $a_r(\mu, \delta) < a_{fb}$. To see this, note that for $\mu \to 0$, the expression of a_{mfb} is identical to the expression of $a_r(\mu, \delta)$ (see equation (21)), except that here ζ replaces δ in (21). However, because of the appropriability effect, we have $\zeta > \delta$. By continuity this implies that, for μ sufficiently small, $a_{mfb} > a_r(\mu, \delta)$ and thus, *a fortiori*, $a_{fb} > a_r(\mu, \delta)$.

References

Acemoglu, Daron, Philippe Aghion, and Fabrizio Zilibotti (2002). "Distance to Frontier, Selection, and Economic Growth." NBER working paper no. W9066.

Acemoglu, Daron, Simon Johnson, and James Robinson (2001). "Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, 91, 1369–1401.

Acemoglu, Daron, and Fabrizio Zilibotti (1997). "Was Prometheus Unbound by Chance? Risk Diversification and Growth." *Journal of Political Economy*, 105, 709–751.

Aghion, Philippe, and Peter Howitt (1992). "A Model of Growth Through Creative Destruction." *Econometrica*, 60, 323–351.

Aghion, Philippe, and Peter Howitt (1998). *Endogenous Growth Theory.* MIT Press.

Barro, Robert, and Xavier Sala-i-Martin (1997). "Technological Diffusion, Convergence and Growth." *Journal of Economic Growth*, 2, 1–26.

Barro, Robert J., and Jong Wha Lee (1996). "International Measures of Schooling Years and Schooling Quality." *American Economic Review*, 86, 218–223.

Brezis, Elise, Paul Krugman, and Dani Tsiddon (1994). "Leapfrogging International Competition: A Theory of Cycles and National Technological Leadership." *American Economic Review*, 83, 1211–1219.

Dewatripont, Mathias, and Eric Maskin (1995). "Credit and Efficiency in Centralized and Decentralized Economies." *Review of Economic Studies,* 62, 541–555.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer (2002). "The Regulation of Entry." *Quarterly Journal of Economics*, 117, 1–38.

Do, Quy-Toan (2002). "Institutions, Institutional Change and the Distribution of Wealth." Working paper, MIT.

Easterly, William, and Ross Levine (1997). "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics*, 112, 1203–1250.

Galor, Oded, and Daniel Tsiddon (1997). "Technology, Mobility, and Growth." *American Economic Review*, 87, 363–382.

Gerschenkron, Alexander (1962). *Economic Backwardness in Historical Perspective*. Harvard University Press.

Greenwood, Jeremy, and Boyan Jovanovic (1990). "Financial Development, Growth, and the Distribution of Income." *Journal of Political Economy*, 98, 1076–1107.

Griffith, Rachel, Stephen Redding, and John Van Reenen (2004). "Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries." *Review of Economics and Statistics*, 86, 883–895.

Grossman, Gene, and Elhanan Helpman (1994). "Protection for Sale." *American Economic Review*, 84, 833–850.

Grossman, Gene, and Elhanan Helpman (2001). *Special Interest Politics*. MIT Press.

Hassler, John, and Jose Rodriguez Mora (2000). "Intelligence, Social Mobility, and Growth." *American Economic Review* 90, 888–908.

Howitt, Peter (2000). "Endogenous Growth and Cross-Country Income Differences." *American Economic Review,* 90, 829–846.

Howitt, Peter, and David Mayer (2002). "Technological Innovation, Implementation and Stagnation: A Schumpeterian Theory of Convergence Clubs." NBER working paper no. W9104.

King, Robert G, and Ross Levine (1993). "Finance and Growth: Schumpeter Might Be Right." *Quarterly Journal of Economics*, 108, 717–737.

Nelson, Richard, and Edmund Phelps (1966). "Investment in Humans, Technological Diffusion and Economic Growth." *American Economic Review,* 56, 69–75.

Rajan, Raghu, and Luigi Zingales (1999). "Which Capitalism? Lessons from the East Asian Crisis." *Journal of Applied Corporate Finance*, 11, 40–48.

Tong, Jian, and Chenggang Xu (2004). "Financial Institutions and the Wealth of Nations: Tales of Development." Discussion paper STICERD no. TE/04/469, London School of Economics.

Wooldridge, Jeffrey (2002). *Econometric Analysis of Cross-section and Panel Data*. MIT Press.

Zeira, Joseph (1998). "Workers, Machines and Economic Growth." *Quarterly Journal of Economics,* 113, 1091–1117.