Distance to Frontier, Selection, and Economic Growth*

Daron Acemoglu[†], Philippe Aghion[‡] and Fabrizio Zilibotti[§]

December 2, 2003

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 in return, 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, that 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 Numbers: O31, O33, O38, O40, L16.

Keywords: appropriate institutions, convergence, economic growth, innovation, imitation, political economy of growth, selection, technical change, technology adoption, traps.

*We thank seminar participants at Birkbeck College, Brown University, Canadian Institute of Advanced Research, DELTA, IIES, London School of Economics, MIT, NBER, SITE Stockholm, Universite de Toulouse, Universita Bocconi, University of Chicago and Abhijit Banerjee, Gary Becker, Mathias Dewatripont, and Byeongju Jeong for helpful comments, and Mauricio Prado for research assistance.

[†]Massachusetts Institute of Technology

[‡]Harvard University

[§]Institute for International Economic Studies, Stockholm University

"... 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 regards 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 (*Economic Backwardness in Historical Perspective,* p. 7)

1 Introduction

In his famous essay, *Economic Backwardness in Historical Perspective*, Gerschenkron 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, large firms and state intervention, facilitate such convergence. If this assessment is correct, the institutions/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 relatively 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 lower growth, and how the political influence of the beneficiaries of existing policies may prevent policy reform.

To understand the main mechanism in our model, imagine the following stylized economy with three key features: (i) firms (entrepreneurs) are either high skill or low skill (or high and low type); (ii) there are credit constraints restricting the amount of investment; and (iii) firms engage both in innovation and adoption of existing technologies from the world technology frontier. If a firm is successful and revealed to be high skill, it will continue to operate. If it is revealed to be low skill, it can be terminated and replaced by a new draw, which will on average have higher skills. However, because of the credit market imperfections, their retained earnings enable existing firms ("insiders") to undertake greater investments. Consequently, the decision to terminate unsuccessful firms 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" firms[1] are more important for innovation than for adoption of existing technologies: adoption and imitation are relatively more straightforward activities compared to innovation. This leads to a key implication of our model: retaining unsuccessful firms and 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 firms becomes less likely.

A likely equilibrium sequence is for an economy to start with *an investment-based strategy*, relying on existing firms 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 less successful firms and entrepreneurs are terminated.

Furthermore, as suggested by Gerschenkron, government intervention to encourage the investment-based strategy might be useful, because the investment-based strategy may fail to emerge even when it is 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 limiting the extent of competition, which increases 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, there is the *rent-shield effect*: the cash (rents) in the hands of insiders creates a shield protecting 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

[1] Our argument applies both to the selection of firms/entrepreneurs and selection of managers to run existing firms. In the model, for simplicity we focus on the selection of entrepreneurs.

switch to the innovation-based strategy reduces growth, because the economy is not making best use of innovation opportunities. But 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 get 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 and typically focuses on whether the world technological leadership is taken over by a newcomer.[2]

But this analysis poses another important question: why do governments not choose institutions/policies that favor the investment-based strategy when the country is at 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.[3] An interesting possibility is that societies may get trapped with "inappropriate institutions" and relatively backward technologies, precisely because earlier they adopted appropriate institutions for their circumstances at the time, but in the process also created a powerful constituency against change.

The rest of the paper is organized as follows. Section 2 presents some motivating

[2] The type of leapfrogging implied by our model may help explain why some of the 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.

[3] Both the Korean and the Japanese cases illustrate the political economy problems created by the investment-based strategy. The close links between government officials and the chaebol in the Korean case and the bureaucrats and the keiretsu in the Japanese case appear to have turned into major obstacles to progress. On the influence of Korean chaebol on policy, Kong (2002, p. 3) writes "...political—not economic—considerations dominated policymaking... [in Korea].... and ...corruption was far greater than the conventional wisdom allows". In fact, the patriarchs of Samsung, Daewoo and Jinro, the three major chaebol, were convicted in the late 1990s of major bribing of two former presidents. Significantly, their jail sentences were pardoned in 1997 (see Asiaweek, October 10, 1997).

evidence and discusses the related literature. Section 3 outlines the basic model. Section 4 characterizes the equilibrium. Section 5 discusses government policy and the possibility of political economy traps. Section 6 concludes. The Appendix contains details on the empirical evidence discussed above and theoretical extensions.

2 Motivating Evidence and Related Literature

The main assumption of our analysis is that R&D and innovation become more important as an economy approaches the world technology frontier. The data from the OECD sectoral database used by Griffith, Redding and Van Rennan (2003) are consistent with this prediction. Griffith, Redding and Van Rennan construct a measure of distance to frontier for each industry in each country and year, DTF_{ict}.[4] This measure is defined as the highest TFP in industry i at time t divided by TFP in industry i in country c at time t (so it takes higher values when a particular country is further behind the frontier in this industry). The same dataset also provides data on R&D intensity, RD_{ict} (R&D divided by sales). Table 1 reports the correlation between these two measures with or without controlling for year effects, country effects, industry effects, between the years 1974 and 1990. The first three columns to use a measure of DTF_{ict} without correcting for differences in skills and hours, while the last three columns use a measure that corrects for these differences (see Griffith, Redding and Van Rennan, 2003).

All columns show the same pattern; there is a statistically highly significant negative relationship between distance to frontier and R&D intensity. The relationship in column 1, which does not control for country industry or year effects, yields and coefficient of -0.018 with standard error 0.003.[5] The second column shows that a similar, though quantitatively smaller, relationship is present once we condition on a full set of year, country and industry effects. Therefore, industries that are farther from the frontier in a given country are less R&D intensive than industries closer to the frontier in the same country. The third column shows a very similar result once we condition on a full set of industry times country effects as well as the effects (i.e., δ_{ic} rather than δ_i and δ_c separately).

Columns 4-6 show similar results using the distance to frontier measure corrected

[4]We are grateful to Rachel Griffith for providing us with these data and for generous help with the empirical evidence presented here. See Griffith, Redding and Van Rennan (2003) for the construction of the distance to frontier measures, descriptive statistics and details.

[5]This estimate implies that a one standard deviation increase in distance to frontier, which is 0.33, translates into an effect of 0.006, which is about 15 percent of the standard error of R&D intensity (0.045).

for differences in skills and hours. In all cases, the relationship is highly statistically significant.

Overall, the evidence indicates that there is generally less R&D in industries farther from the frontier, which is consistent with the main premise of our model.

Though somewhat less directly related, the evidence in Acemoglu, Johnson and Mitton (2003) is also consistent with our approach. They find greater industry-level and aggregate concentration in poorer countries. These results are consistent with the notion that there is less selection and typically larger firms in countries farther from the frontier, which are typically does that are poorer.

In addition, our analysis implies that certain non-competitive policies may have limited costs, or even benefits, when countries are far from the world technology frontier, but 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 discussed in footnote 2, we are not aware of any systematic empirical investigation. While a detailed empirical analysis is beyond the scope of the current paper, a brief look at the data reveals some notable patterns consistent with this implication.

Figures 1a and 1b look at the relationship between growth and initial distance to frontier (GDP per capita relative to the U.S.) in the sample of non-OECD, non-socialist countries separately for those with high and low degree of "non-competitive" policies/barriers to entry (here we do this using the measure of number of procedures necessary for opening a new business, from Djankov, La Porta, Lopez-de-Silanes, and Shleifer, 2002). The figures show growth in per capita income between 1965 and 1995 plotted against distance to frontier in 1965, where we also control for a dummy for sub-Saharan African countries, which have much lower growth rates. While there is a strong negative relationship between growth and distance to frontier for countries with high barriers, the relationship is much weaker for countries with low barriers. In other words, high-barrier countries do relatively well when they are far from the frontier, but slow down significantly near the frontier, while low-barrier countries grow almost equally successfully near or far from the frontier.[6] This is consistent with the notion that barriers to entry

[6] In the regression of country growth rates between 1965 and 1995 on the sub-Saharan Africa dummy and distance to frontier in 1965 in the sample of low-barrier countries, the coefficient on distance to frontier is -0.021 (s.e.=0.030), which is insignificant at the 5 percent level (shown in Figure 1b). The same coefficient is -0.073 (s.e.=0.029) in the sample of high-barrier countries, which is significant at the 5 percent (shown in Figure 1a). 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. Details on the estimates shown in these figures, information on samples and robustness

are more harmful to growth closer to the frontier, though this cross-country relationship may be driven by other omitted cross-country differences.

Figures 1d and 1e show the same pattern when we look at growth in 5-year intervals and control for country fixed effects and time effects. These figures show that 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.[7]

Figures 2a-2d show that the same results hold when we look at the differential growth experiences of countries with different degrees of openness to international trade. Here we split the sample according to the predicted openness measure constructed by Frankel and Romer (1999), which exploits "exogenous" differences in openness from a standard "gravity equation" due to differences in population, land area, proximity and common borders to other countries, and whether a country is landlocked.[8]

Finally, Figures 3a-d turn to another implication of our model, that skills should matter more nearer to frontier. If this is the case, when we split the sample according to low and high human capital (using total years of schooling in 1965), we should see a more negative relationship between growth and distance to frontier for low-human capital than for high-human capital countries. This is the pattern we find in the data, though now the contrast is somewhat weaker in the cross-sectional regressions, and but still strong in the fixed effect regressions.[9] The evidence presented in Figures 1, 2 and 3 therefore suggests that cross-country growth patterns are broadly consistent with the basic implications of our approach, though this is only a first pass, and more detailed empirical analysis of these patterns is necessary in future work.

FIGURES 1, 2 AND 3

Our paper relates to a number of different literatures. First, the notion that skills are

checks are provided in Appendix A.

[7]With country fixed effects and time effects, the coefficient on the distance to frontier in the low-barrier sample is -0.036 (s.e.= 0.036), while in the high-barrier sample it is -0.105 (s.e.= 0.046). We also obtain similar results in fixed effect regressions when distance to frontier is instrumented by its past values in order to avoid biases resulting from the fact that distance to frontier is correlated with lags of the dependent variable. See Appendix A for details.

[8]In the cross-sectional regressions, the coefficient on the distance to frontier for the "closed" economies is -0.046 (s.e.= 0.012), while for the "open" economies, it is -0.020 (s.e.= 0.028). In the fixed effect regressions, the coefficient for closed economies is -0.194 (s.e.= 0.051), while for open economies, it is -0.81 (s.e.= 0.031). See Appendix A for details.

[9]In the cross-sectional regressions, the coefficient on the distance to frontier for low-education countries is -0.094 (s.e.= 0.047), while for high-education countries, it is -0.050 (s.e.= 0.026). In the fixed effect regressions, the coefficient for the low-education countries is -0.221 (s.e.= 0.065) and for the high-education countries, it is -0.048 (s.e.= 0.030).

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),[10] and to the emphasis in Galor and Tsiddon (1997) and Hassler and Rodriguez (2000) on the importance of ability and skill in times of economic change and turbulence. Second, our model is related to work on finance and growth, including Greenwood and Jovanovic (1990), King and Levine (1993), and Acemoglu and Zilibotti (1997). Third, our focus is related to work on technological convergence, including Barro and Sala-i-Martin (1997), Aghion and Howitt (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.

Perhaps more closely related are Tong and Xu (2000) and Rajan and Zingales (1999). Tong and Xu extend the model by Dewatripont and Maskin (1995) and 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, investigate how certain arrangements that are at first growth enhancing later reduce growth or even cause non-convergence traps. Rajan and Zingales argue that the same practices that were useful for the success of East Asian economies were also 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.

Another link is to the debate on the optimal degree of government intervention in less developed countries. Consistent with the Gerschenkron view, some economists, e.g., Stiglitz (1995) and Hausmann and Rodrik (2002), call for greater government intervention in less developed countries where market failures tend to be more severe than in more advanced economies. Countring this, several economists and political scientists emphasize the greater danger of government failures in less developed nations, where checks on governments are weaker (e.g., Shleifer and Vishny, 1999). Our model combines these two insights. We derive a reason for possible government intervention at the early stages of development, while also highlighting why such intervention can be counterproductive because of political economy considerations.

[10] Nelson and Phelps (1966) rank activities according to 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). They 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" (p. 69).

3 The model

3.1 Agents and production

The model economy is populated by overlapping generations of two-period lived risk-neutral agents, discounting 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 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 entrepreneurship (management). In particular, we assume that each worker has high skill (ability) in entrepreneurship with probability λ and low skill with probability $1-\lambda$.

There is a unique final good in the economy, also used as an input to produce intermediate inputs. We take this good as the numeraire. The final good is produced competitively from labor and a continuum 1 of intermediate goods as inputs with the aggregate production function:

$$y_t = \frac{1}{\alpha} L_t^{1-\alpha} [\int_0^1 (A_t(\nu))^{1-\alpha} x_t(\nu)^{\alpha} \; d\nu], \tag{1}$$

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

In each intermediate sector ν, one production site has access to the most productive technology, $A_t(\nu)$, so this "leading firm" will enjoy monopoly power. Each leading firm has access to a technology to transform one unit of the final good into one unit of intermediate good of productivity $A_t(\nu)$. A fringe of additional firms can "steal" this technology, and produce the same intermediate good, with the same productivity $A_t(\nu)$, 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 both technological factors and 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 for the incumbent to be forced to charge a limit price to prevent entry by the fringe. This limit price is equal to the marginal cost of the fringe:

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

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

$$\pi_t(\nu) = (p_t(\nu) - 1) x_t = \delta A_t(\nu) L_t, \tag{3}$$

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

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

$$A_t \equiv \int_0^1 A_t(\nu) d\nu. \tag{4}$$

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:

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

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

$$y_t^{net} = y_t - \int_0^1 x_t(\nu) d\nu = \zeta A_t L_t, \tag{6}$$

where $\zeta \equiv (\chi - \alpha) \chi^{-\frac{1}{1-\alpha}} / \alpha$ is monotonically decreasing in χ. Thus for given average technology A_t, both total output 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 only capture a fraction of the greater productivity in the final goods sector (or of the consumer surplus) created by their production.

3.2 Technological progress and productivity growth

Each leading firm (capitalist) requires an entrepreneur to operate. This leaves $L_t = N$ production workers (recall that the total size of worker population is $N + 1$). Firm productivity is determined by entrepreneurial skill and 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", and that the large project requires an additional investment, while the small one does not.[11] 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, which collect funds from consumers. Intermediation is without any costs and there is free entry into this activity. Moreover, since intermediation takes place within a period, there are no interest costs to be covered.

Entrepreneurial skills, which affect productivity growth, are initially unknown, and are revealed after an agent works as an entrepreneur for the first time. Entrepreneurs perform two important tasks: (1) They engage in *innovation*, and entrepreneurial skills are important for success in this activity. (2) They also *adopt* technologies from the frontier, and here skills play a less important role than in innovation. This assumption captures the notion that relatively backward economies can grow by adopting already well-established technologies, and 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, i.e.,

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

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

The productivity of intermediate good ν at time t is expressed as:

$$A_t(\nu) = s_t(\nu)\left(\eta \bar{A}_{t-1} + \gamma_t(\nu) A_{t-1}\right), \tag{8}$$

where $s_t(\nu) \in \{\sigma, 1\}$ denotes the size of the project, with $s_t(\nu) = \sigma < 1$ corresponding to a small project and $s_t(\nu) = 1$ corresponding to a large project. $\gamma_t(\nu)$ denotes the skill 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

[11] The NBER working paper version shows that the qualitative results are identical with a positive but smaller investment cost for a small project.

skills as captured by the term $\gamma_t(\nu)$. 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), we have the growth rate of aggregate technology as:

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

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 1, so that the country is close to the frontier, innovation matters relatively more, and growth is driven by the $\gamma_t(\nu)$ term. Consequently, as the country develops and approaches the world technology frontier, innovation and entrepreneurial selection become more important.

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

Finally, the cost of investment for the large project is:

$$k_t(\nu \mid s = 1) = \kappa\bar{A}_{t-1}, \tag{10}$$

while, as noted above, the small project, $s = \sigma$, requires no investment, i.e., $k_t(\nu \mid s = \sigma) = 0$. The assumption that investment cost is proportional to $\bar{A}_{t-1}$ ensures balanced growth.[12] 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 $\bar{A}_{t-1}$.

[12] Alternatively, investment costs of the form $k_t(\nu) = \kappa\bar{A}_{t-1}^{\rho}A_{t-1}^{1-\rho}$ 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 for the expressions when $\rho < 1$.

3.3 Contracts, incentive problems and credit constraints

Capitalists make contract offers to a subset of workers to become entrepreneurs, specifying the loan amount from intermediaries, and payments to entrepreneurs, as well as the level of investment. Investment costs are financed either through the retained earnings of entrepreneurs or by borrowing.[13] To simplify the discussion, we also assume that young capitalists (new firms) cannot hire old entrepreneurs (e.g., because old cohorts' 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 market constraints, restricting 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 firms) to shield them against the threat of entry by new entrepreneurs (firms).

To specify the incentive compatibility constraints more formally, define $\pi_t(\nu \mid s, e, z)$ as the ex post cash-flow generated by firm ν 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 $z \in \{L, H\}$, where Y denotes young, O denotes old, L stands for low skill and H for high skill. $\pi_t(\nu \mid s, e, z)$ is simply given by the expression in (3) with $A_t(\nu)$ 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:[14]

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

where $S_t(\nu \mid s, e, z)$ is the payment to an entrepreneur of age e and skill z, running a project of size s.

[13] Whether old capitalists inject their own funds or borrow from intermediaries is immaterial, since there is no cost of intermediation, and the incentive problems are on the side of entrepreneurs.

[14] This specification rules out long-term contracts where the payment to an old entrepreneur is conditioned on whether he has diverted funds in the first period or not. 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.

Let us define $\widehat{RE}_t(\nu \mid s, e, z) \leq k_t(\nu \mid s)$ as the retained earnings injected by an entrepreneur to finance part of the investment costs,[15] and $RE_t(\nu \mid s, e, z)$ as total retained earnings. Naturally, we have $0 \leq \widehat{RE}_t(\nu \mid s, e, z) \leq RE_t(\nu \mid s, e, z)$, and $\widehat{RE}_t(\nu \mid s, \cdot, z = Y) = 0$, since young entrepreneurs have no funds to contribute. 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 capitalists with a project of size s, entrepreneur of age e and skill z, and

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

as the profit-maximizing project size choice for capitalists when the entrepreneur has age e and skill z be, where E_t is the expectations operator at time t which 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 B, where we also show that for N 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 $S_t(\nu \mid s, e, z)$ such that the incentive compatibility constraints in (11) are satisfied.

Since the participation constraint is slack, there will be an excess supply of young agents willing to become entrepreneurs. Thus young entrepreneurs will be paid the lowest salary consistent with incentive compatibility, (11). The same also applies to old low-skill entrepreneurs (since 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). 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:

$$V_t^*(e = O, z = H) \leq \max \left\langle V_t^*(e = O, z = L); E_t V_t^*(e = Y, \cdot)\right\rangle. \tag{15}$$

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

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, offer a higher salary to attract an old high-skill entrepreneur, and 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 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 the 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 of the economy, a_t) 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 $R_t = 1$ when $E_t V_t^*(e = Y, \cdot) \geq \max \langle 0; V_t^*(e = O, z = L) \rangle$ and $R_t = 0$ when $E_t V_t^*(e = Y, \cdot) < V_t^*(e = O, z = L)$.

A dynamic equilibrium is obtained by piecing together static equilibria as defined in Definition 1 through the law of motion of aggregate productivity as given by (9). We provide the equilibrium law of motion in greater detail below.

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,[16] credit market imperfections and moral hazard can make the capitalists unwilling to undertake the investment. In particular, because of the incentive compatibility constraint (11), profits must be shared between

[16]Without moral hazard, the large project is profitable for all a_t if $(1 - \sigma)\,\delta N \eta > \kappa$.

the capitalist and the entrepreneur with the shares $(1-\mu)$ and μ. Since capitalists only appropriate a fraction $1-\mu$ of the returns, they will tend to underinvest.

Since young entrepreneurs have no wealth, capitalists hiring young entrepreneur must cover the entire investment cost, $\kappa\bar{A}_{t-1}$. This implies that, conditional on the investment, the expected value of a firm hiring a young entrepreneur is equal to:

$$E_t V_t(\nu \mid s=1, e=Y, \cdot) = [(1-\mu)\,\delta N((\eta+\lambda\gamma a_{t-1})) - \kappa]\,\bar{A}_{t-1}. \tag{17}$$

The capitalist compares (17) to the value of the firm conditional on no investment, i.e.:

$$E_t V_t(\nu \mid s=\sigma, e=Y, \cdot) = (1-\mu)\,\delta N\sigma\,(\eta+\lambda\gamma a_{t-1}))\bar{A}_{t-1}. \tag{18}$$

Which of the two expressions is higher depends on a_{t-1} and on parameters. While our general argument only requires the credit constraint to bind in some range of a_{t-1}, we simplify the exposition by focusing on economies where (18) is larger than (17) for *all* values of a_{t-1}. For this to be the case, we assume that (see Appendix C for the case where this assumption is relaxed):

$$\kappa > (1-\mu)\,(1-\sigma)\,(\eta+\lambda\gamma)\,N\delta. \tag{A1}$$

Although Assumption (A1) ensures that young entrepreneurs run small projects, the same may not be true for old entrepreneurs. The key difference is that while young entrepreneurs have no wealth, old entrepreneurs can use their retained earnings to finance part of the investment cost. In fact, in equilibrium firms would never employ low-skill old entrepreneurs and operate small projects. To see this, simply note that

$$\begin{aligned} V_t(\nu \mid s=\sigma, e=O, z=L) &= (1-\mu)\,\delta N\sigma\eta\bar{A}_{t-1} \\ &< (1-\mu)\,\delta N\sigma\,(\eta+\lambda\gamma a_{t-1}))\bar{A}_{t-1} = E_t V_t(\nu \mid s=\sigma, e=Y, \cdot)\,, \end{aligned}$$

where the first expression follows from the fact that the small project does not require any investment costs. Thus, since young entrepreneurs are on average more productive, firms always prefer to have young entrepreneurs operate small projects rather than to have these projects operated by an old low-skill entrepreneur. Therefore, we obtain:

Lemma 1 If an old low-skill entrepreneur is retained (i.e., $R_t = 1$), then he operates a large project (i.e., $s_t^*(e=O, z=L)=1$). Moreover, if Assumption (A1) holds, then, for all $a \in [0,1]$, young entrepreneurs operate small projects (i.e., $s_t^*(e=Y, z=\cdot)=\sigma$).

The Lemma establishes that if a low-skill entrepreneur is retained, he must be running a large project (note that this conclusion does not require Assumption (A1) to hold). When does the firm prefer to retain a low-skill old entrepreneur rather than hiring a young entrepreneur? The answer depends on the amount of retained earnings of old low-skill entrepreneurs as well as the cost of foregoing the opportunity cost of trying a new entrepreneur (the assumptions that participation constraints are slack even when entrepreneurs inject all their return earnings and that capitalists make the contract offers imply that old low-skill entrepreneurs inject all their earnings, i.e., $\widehat{RE}_t = RE_t$, see Appendix B). Retained earnings are equal to the capitalized first period entrepreneurial earnings. In the case of a low-skill entrepreneur born at $t-1$, this is given by:[17]

$$RE_t = \frac{1+r}{1+g}\sigma\mu\delta N\eta\bar{A}_{t-1}. \tag{19}$$

To simplify the discussion, we focus in the text on the case where $RE_t < \kappa\bar{A}_{t-1}$, so that old low-skill entrepreneurs can cover only part of the investment cost (for the analysis of the converse case, see Appendix C). More formally, this corresponds to:

$$\kappa > \frac{1+r}{1+g}\sigma\mu\eta N\delta. \tag{A2}$$

(Note that like Assumption (A1), (A2) requires κ to be sufficiently large).

We can now characterize the value of a firm that retains an old low-skill entrepreneur and operates the large project:

$$V_t\left(\nu \mid s=1, e=O, z=L\right) = \left[\left(1-\mu\right)\delta N\eta\bar{A}_{t-1} - \left(\kappa\bar{A}_{t-1} - RE_t\right)\right], \tag{20}$$

where RE_t is given by (19), and $\kappa\bar{A}_{t-1} - RE_t > 0$ by Assumption (A2).

The firm retains low-skill old entrepreneurs whenever

$$V_t\left(\nu \mid s=1, e=O, z=L\right) > E_t V_t\left(\nu \mid s=\sigma, e=Y, \cdot\right), \tag{21}$$

where the left hand side is given by (20), while the right hand side is given by (18). If condition (21) does not hold, low-skill entrepreneurs are terminated. Condition (21) defines a threshold level of the distance to frontier, $a_r\left(\mu,\delta\right)$, such that below this threshold, low-skill old entrepreneurs are retained ($R=1$), and above this threshold they are

[17] To obtain the expression for RE_t, note that the entrepreneur has low skill and given Assumption (A1), in his youth, he ran a small project thus receiving a wage of $\mu\delta\sigma N\eta\bar{A}_{t-2} = \mu\delta\sigma N\eta\bar{A}_{t-1}/\left(1+g\right)$. Using (7) and taking into account the interest payments at the rate r gives us the right-hand side of (19).

terminated ($R = 0$). Using (18) and (20), we obtain this threshold as:

$$a_r(\mu, \delta) \equiv \frac{\left((1-\mu)(1-\sigma) + \frac{1+r}{1+g}\mu\sigma\right)\eta - \frac{\kappa}{\delta N}}{(1-\mu)\sigma\lambda\gamma}. \tag{22}$$

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 pointed out above, 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, they obtain only a fraction of the returns, consequently they have a bias against retaining old entrepreneurs, i.e., $R = 1$, which is associated with greater investment expenditures.[18] A higher δ weakens the extent of this appropriability effect and enables firms, and hence capitalists, to capture more of the surplus, encouraging $R = 1$. Second, as shown by (19), 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 retention ($R = 1$) more likely.

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

$$\delta > \frac{\kappa}{\sigma\eta L}\frac{1+g}{1+r}, \tag{23}$$

the former effect dominates and a_r is increasing in μ, and more severe moral hazard/credit market problems encourage $R = 1$. In contrast, when (23) does not hold, these problems encourage the termination of low-skill entrepreneurs.

Finally, high-skill entrepreneurs are always retained, and it can also be shown that $a_r(\mu, \delta) > 0$ is sufficient to ensure that the firm operates large projects with old high-skill entrepreneurs for all $a \in [0, 1]$ (see Appendix D for a proof).

We now summarize the analysis of the static equilibrium as follows:

Proposition 1 Suppose that Assumptions (A1) and (A2) hold. Then, for given a_{t-1}, there exists a unique equilibrium such that (i) young entrepreneurs operate small projects

[18] Capitalists do not pay the full investment costs, since entrepreneurs also contribute their retained earnings. Nevertheless, Assumption (A2) ensures that capitalists pay a sufficiently large fraction of the costs and that there is a tendency for underinvestment.

$(s_t^*(e=Y,\cdot)=\sigma)$; (ii) old low-skill entrepreneur are retained $(R_t=1)$ and operate large projects $(s_t^*(e=O,z=L)=1)$ if $a_{t-1}<a_r(\mu,\delta)$ and are terminated $(R_t=0)$ if $a_{t-1}>a_r(\mu,\delta)$, where $a_r(\mu,\delta)$ is given by (22); (iii) old high-skill entrepreneurs are always retained and, as long as $a_r(\mu,\delta)>0$, they 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 fully characterize the dynamic equilibrium of the economy. Let us first characterize the law of motion of a_t conditional on the retention decision, R_t. Note that half of the firms are young and use (4) to write $A_t \equiv \int_0^1 A_t(\nu)\,d\nu = \left(A_t^Y + A_t^O\right)/2$, where A_t^Y is average productivity among young firms and A_t^O is average productivity among old firms. In addition, since all young firms hire young entrepreneurs who, from Lemma 1, choose $s=\sigma$ and a fraction λ of those are high skill, we have $A_t^Y=\sigma(\eta\bar{A}_{t-1}+\lambda\gamma A_{t-1})$.

Average productivity among old firms depends on whether we have $R=1$ or $R=0$. With $R=1$, all entrepreneurs are retained, so a fraction λ are high ability, and all old entrepreneurs choose $s=1$, so $A_t^O[R=1]=\eta\bar{A}_{t-1}+\lambda\gamma A_{t-1}$. If, on the other hand, $R=0$, only a fraction λ of the entrepreneurs, those revealed to be high skill, are retained, and the remaining $1-\lambda$ are replaced by young entrepreneurs. Thus, 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:

$$a_t = \begin{cases} \frac{1+\sigma}{2(1+g)}\left[\eta+\lambda\gamma a_{t-1}\right] & \text{if } R_t=1 \\ \frac{1}{2(1+g)}\left[\left(\lambda+\sigma+(1-\lambda)\sigma\right)\eta+\left(1+\sigma+(1-\lambda)\sigma\right)\lambda\gamma a_{t-1}\right] & \text{if } R_t=0 \end{cases}. \quad (24)$$

This equation, which is also depicted in Figure 4, 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, since $(1+\sigma)\lambda\gamma a_{t-1} < (1+\sigma+(1-\lambda)\sigma)\lambda\gamma a_{t-1}$. 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 if this comes at the expense of sacrificing entrepreneurial selection. This 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 where unsuccessful entrepreneurs are terminated and only successful entrepreneurs are retained.

FIGURE 4 HERE

The full equilibrium is then simply determined by combining this with the equilibrium law of motion, (24), which, using Proposition 1, can be written as:

$$a_t = \begin{cases} \frac{1+\sigma}{2(1+g)}\left(\eta + \lambda\gamma a_{t-1}\right) & \text{if } a_{t-1} \leq a_r(\mu,\delta) \\ \frac{1}{2(1+g)}\begin{pmatrix} (\lambda+\sigma+(1-\lambda)\sigma)\eta \\ +(1+\sigma+(1-\lambda)\sigma)\lambda\gamma a_{t-1} \end{pmatrix} & \text{if } a_{t-1} > a_r(\mu,\delta) \end{cases}. \tag{25}$$

FIGURE 5 HERE

Figure 5 depicts the equilibrium dynamics. As (25) 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, while when a_{t-1} exceeds $a_r(\mu,\delta)$, 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 in the world pursuing the innovation-based strategy. Equation (25) evaluated at $a = 1$ gives the world technology growth rate as:

$$g = \frac{1}{2}\left[(\lambda+\sigma+(1-\lambda)\sigma)\eta + (1+\sigma+(1-\lambda)\sigma)\lambda\gamma\right] - 1, \tag{26}$$

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:

$$(1-\sigma)\eta < \lambda\gamma. \tag{27}$$

Consequently, at $a = 1$, the $R = 0$ line intersects the 45 degree line and is above the $R = 1$ line. But then, as drawn in Figure 5, the $R = 1$ line must intersect the 45 degree line at some $a_{trap} < 1$. From (24), this threshold value can be calculated as:

$$a_{trap} = \frac{(1+\sigma)\eta}{2(1+g) - \lambda\gamma(1+\sigma)}. \tag{28}$$

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.

However, in practice, 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

$$a_{trap} < a_r(\mu, \delta),$$

which corresponds to the case depicted in Figure 5. In contrast, Figure 6 shows the case where $a_{trap} > a_r(\mu, \delta)$, so that the economy switches out of the investment-based strategy before a_{trap} is reached, and the non-convergence trap does not arise.

FIGURE 6 HERE

When is this condition likely to be satisfied? From (28), a_{trap} is an increasing function of $\lambda\gamma$, and is independent of $\kappa/\delta N$ and μ. Since $a_r(\mu, \delta)$ 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 (23) 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. Since 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 (23) 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 that Assumptions (A1) and (A2) hold, and the economy starts with distance to frontier 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 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

Imagine a social planner interested in maximizing the growth rate of the economy, and she has to take the equilibrium prices and project size choices of Lemma 1 as given. Will she choose an innovation-based strategy ($R = 0$) or an investment-based strategy ($R = 1$)?[19]

Inspection of (24) or of Figure 4 immediately shows that growth will be maximized when the economy reaches the highest level of a_t for a given a_{t-1}, or in other words, she should pursue a strategy of $R = 1$ whenever $a_{t-1} < \hat{a}$, and 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:

$$\hat{a} \equiv \frac{\eta(1-\sigma)}{\lambda\gamma\sigma}. \tag{29}$$

Condition (27) ensures that $\hat{a} < 1$. Therefore, similar to 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 the equilibrium.

How does $\hat{a}$ compare to the equilibrium threshold $a_r(\mu, \delta)$? The answer depends, among other things, on the degree of competition as measured by δ. The appropriability effect discussed above means that equilibrium behavior is biased against the investment-based strategy, creating a force towards $a_r(\mu, \delta) < \hat{a}$. However, countering this, there is what we might call 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 might increase or reduce the gap between the equilibrium and the growth-maximizing allocations depending on whether we start from

[19] We characterize the growth-maximizing strategy not for welfare comparisons, but to derive the implications of equilibrium behavior for aggregate growth rates. In Appendix E, 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.

a situation where $\hat{a} > a_r(\mu, \delta)$ or $\hat{a} < a_r(\mu, \delta)$. Given μ, there exists a unique level of competition δ, denoted by $\widehat{\delta}(\mu)$, such that $\hat{a} = a_r\left(\mu, \widehat{\delta}(\mu)\right)$, where

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

If the product market is less competitive than implied by this threshold, namely, if $\delta > \widehat{\delta}(\mu)$, then we have $\hat{a} < a_r(\mu, \delta)$, and the economy generates excess retention of low-skill entrepreneurs relative to the growth-maximizing allocation. In this case, which is the one shown in Figure 5, limiting competition (larger δ) would further increase the growth gap between the equilibrium and the growth-maximizing strategy. Conversely, if product market competition is high, namely if $\delta < \widehat{\delta}(\mu)$, then $\hat{a} > a_r(\mu, \delta)$ as shown in Figure 6, and 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 above 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, the discussion of non-convergence traps in the previous subsection also highlights that limiting product market competition may later become harmful to growth, and prevent convergence to the frontier. In particular, there exists a threshold competition level, $\delta^*(\mu)$, such that

$$a_r(\mu, \delta^*(\mu)) = a_{trap}, \tag{30}$$

where a_{trap} is given by (28). An economy with a sufficiently high level of competition, $\delta < \delta^*(\mu)$, will never fall into a non-convergence trap. Therefore, excessively high competition may cause a slowdown in the process of technological convergence at the earlier stages of development, but does not affect the long-run equilibrium. Low competition, on the other hand, may have detrimental effects in the long-run.

This analysis therefore 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, while economy 2 starts 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, since economy 2 still pursues the investment-based strategy despite the fact that growth is now maximized with the innovation-based strategy. Furthermore, if $a_{trap} < a_r(\mu, \delta_2)$, 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 certain rigid institutions, associated with the less competitive market structure supporting the investment-based strategy here, become more costly as an economy approaches the world technology frontier. It may also shed some light on why some economies, such as Brazil, Mexico or Peru, that initially grew relatively rapidly with highly protectionist policies, were then overtaken by economies with more competitive policies such as Hong Kong or Singapore.[20]

5 Policy and political economy traps

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, with 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 fail to 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.

The 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 if not reversed later. 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 (because they are useful only at the early stages of development). The bulk of the section is devoted to an analysis of how political economy considerations, in particular lobbying by groups benefiting from existing policies, might make it harder for the society to abandon these policies, thus turning appropriate institutions into "inappropriate institutions," and potentially generating non-convergence traps.

[20]Interestingly, 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 was reverted in the 1970s and 1980s.

5.1 Policy and appropriate institutions

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 threshold. A policy intervention that encourages greater investment will increase growth over the range $a \in (a_r(\mu,\delta), \hat{a})$.[21] 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 etc. Imagine the government subsidizes a fraction τ of the cost of investment. An analogous analysis to before gives the threshold for switching from the investment- to the innovation-based strategy as:

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

If τ is chosen appropriately, in particular if $\tau = \tilde{\tau}$ such that $\tilde{a}_r(\mu,\delta,\tilde{\tau}) = \hat{a}$, 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 exactly the level of investment made by firms. For this reason, we focus on another potential policy instrument that affects the equilibrium threshold $a_r(\mu,\delta)$, the extent of anti-competitive policies, such as entry barriers, merger policies etc.. Naturally, this discussion also applies to investment subsidies.

Anti-competitive policies are captured by the parameter χ in our model, and recall that δ is monotonically increasing in χ. Thus high values of χ or δ correspond to a less competitive environment. Starting from a situation where $a_r(\mu,\delta) < \hat{a}$, policies that restrict competition will 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 for encouraging faster growth and technological convergence.

The situation where the government chooses a less competitive environment in a backward economy in order to encourage long-term relationships, greater investment and faster technological convergence is reminiscent to Gerschenkron's analysis.[22] But

[21] The analysis in Appendix E 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 maximization rather than growth maximization.

[22] It is also reminiscent to the well-known "infant-industry" arguments calling for protection and government support for certain industries at the early stages of development.

our analysis also reveals that limiting in the degree of competition (and similarly investment subsidies) become harmful for economies closer to the world technology frontier. Appropriate institutions for early stages of development therefore become *inappropriate* for an economy close to the frontier. Thus an economy that adopts such institutions must later abandon them; otherwise, it will end up in a non-convergence trap.

A sequence of policies whereby certain interventions are first adopted and then abandoned raises important political economy considerations, however. Groups that benefit from anti-competitive policies will become richer while these policies are implemented, and will oppose a change in policy. To the extent that economic power buys political power, for example, 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", where groups enriched by these institutions successfully block reform, and the economy ends up in a non-convergence trap because, at earlier stages of development, it adopted appropriate institutions.

We now build a simple political economy model where special interest groups may capture politicians. Our basic political economy model is a 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 (see also Do, 2002, on this) and combined with our growth setup.

5.2 Political environment

Suppose that competition policy, χ, is determined in each period by a politician (or government) that cares about the current consumption, but is also sensitive to bribes—or to campaign contributions. For tractability, we adopt a very 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 (similar to the "myopic planner" discussed in Appendix E), and to B_t otherwise, 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} in order to ensure stationary policies in equilibrium, since bribes will be increasing in A.

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 check-and-balances that limit the ability of special interest groups to capture politicians. This formulation is similar to that in Grossman and Helpman (1994, 2001), but simpler since in their formulation, 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 is assumed to have perfect commitment to deliver the competition policy promised to an interest group in return for bribes.

Young agents have no wealth, so they cannot bribe politicians. To simplify the discussion, we also assume that only capitalists can organize as interest groups, so the only group with the capability to bribe politicians are old capitalists, and that the institutional choice facing the politician is between two policies, low and high competition, or between "competitive" and "anti-competitive" policies, i.e., $\chi_t \in \{\underline{\chi}, \bar{\chi}\}$ where $\underline{\chi} < \bar{\chi} \leq 1/\alpha$. We set, by analogy, $\delta_t \equiv (\chi_t - 1)\, \chi_t^{-\frac{1}{1-\alpha}} \in \{\underline{\delta}, \overline{\delta}\}$, which, recall, is the parameter in the profit function, (3). We will use χ and δ interchangeably to refer to the degree of competition. Finally, we also set $\zeta_t \equiv (\chi_t - \alpha)\, \chi_t^{-\frac{1}{1-\alpha}}/\alpha \in \{\underline{\zeta}, \overline{\zeta}\}$, which is the parameter in net output, (6), above. The assumption that χ is a discrete rather than a continuous choice variable is reasonable, since the ability of the politicians to fine-tune institutions is often limited (i.e., they can either impose entry barriers or not, etc.), and it approximates a situation where the main choice is whether or not to undertake some major reform.

5.3 Political equilibrium

As a benchmark, let us start with the case without bribes or an "honest" politician, i.e., $H \to \infty$. Such a politician will maximize total current consumption

$$C_t = \zeta N A_t - \int_0^1 k_t(\nu)\, d\nu,$$

that is, net output minus investment, at date t. Throughout the analysis, we maintain Assumptions (A1) and (A2), so Lemma 1 holds and Proposition 1 describes the static equilibrium.

Will an honest politician ever choose anti-competitive policies ($\chi = \bar{\chi}$ and $\delta = \overline{\delta}$)? It is straightforward to show that he will do so for $a \in (a_r(\mu, \underline{\delta}), a_{WM})$, where:[23]

$$a_{WM} \equiv \frac{\left(\bar{\zeta}(1+\sigma) - \underline{\zeta}(\lambda + \sigma(2-\lambda))\right)\eta - (1-\lambda)\kappa/\delta N}{\lambda\gamma\left(\underline{\zeta}(1+\sigma(2-\lambda)) - \bar{\zeta}(1+\sigma)\right)} < \hat{a}. \tag{31}$$

[23] a_{WM} is derived by equating consumption under (i) $R = 1$ and low competition, $\bar{\zeta}$, and (ii) $R = 0$ and high competition, $\underline{\zeta}$. See the working paper version for details. Note also that the set $(a_r(\mu, \underline{\delta}), a_{WM})$ could be empty.

Below $a_r(\mu, \underline{\delta})$, reducing competition does not affect retention decisions, since $R = 1$ anyway, and creates only static monopoly distortions. Above a_{WM}, inducing the investment-based strategy is not sufficiently beneficial. In the range $(a_r(\mu, \underline{\delta}), a_{WM})$, the benefits from inducing the investment-based strategy outweigh the static losses.

Next consider the competition policy set by a politician who responds to bribes (i.e., H finite). Clearly, capitalists always prefer low to high competition, as this increases their profits. Let $B_t^W \equiv B^W(a_{t-1})\bar{A}_{t-1}$ denote the maximum bribe that capitalists are *willing* to pay in order to induce anti-competitive policies, $\delta = \overline{\delta}$, rather than competitive policies, $\delta = \underline{\delta} < \overline{\delta}$. We also assume that agents cannot borrow to pay bribes, so the amount of bribes that they can pay will also be limited by their current income. This assumption introduces the link between economic power and political power (and through this channel, the possibility of history dependence): *richer agents can pay greater bribes and have a greater influence on policy.* Let $B_t^C \equiv B^C(\delta_{t-1}, a_{t-1})\bar{A}_{t-1}$ denote the maximum bribe that they *can* pay, where $\delta_{t-1} \in \{\underline{\delta}, \overline{\delta}\}$ is the level of competition at date $t-1$. It is equal to the profits generated by young firms in period $t-1$ that accrues to capitalists:

$$B^C(\delta_{t-1}, a_{t-1}) = \delta_{t-1}(1-\mu)\sigma(N\eta + \lambda\gamma a_{t-1}). \tag{32}$$

δ_{t-1} features in this equation, since the extent of competition in the previous period determines profits and the maximum bribe that capitalists can pay the politician.

Equilibrium bribes are therefore: $B(\delta_{t-1}, a_{t-1}) = \min\langle B^W(a_{t-1}), B^C(\delta_{t-1}, a_{t-1})\rangle$. We focus on economies where capitalists are credit constrained in the range of interest. Thus, from now on, we have:[24]

$$B(\delta_{t-1}, a_{t-1}) = B^C(\delta_{t-1}, a_{t-1}). \tag{33}$$

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.

As long as $a_{t-1} \notin [a_r(\mu, \underline{\delta}), a_{WM}]$, i.e., as long as the politician does not want to choose the anti-competitive policy, $\overline{\delta}$, for welfare-maximizing reasons, she will be induced to change the policy to $\overline{\delta}$ only when bribes are sufficient to cover the honesty cost, HA_{t-1}, or if and only if: $B^C(\delta_{t-1}, a_{t-1}) \geq Ha_{t-1}$. Using (32), we can rewrite this inequality as

$$\delta_{t-1}(1-\mu)\sigma N(\eta + \lambda\gamma a_{t-1}) \geq Ha_{t-1}. \tag{34}$$

Greater δ_{t-1} makes it more likely that (34) holds, since it corresponds to greater profits for capitalists, which they can use for bribing politicians. We define a_L and a_H as the

[24] See the NBER working paper version for the expression for B_t^W and more details on this point.

unique values of a_{t-1} such that (34) holds with equality for $\delta_{t-1} = \bar{\delta}$ and $\delta_{t-1} = \underline{\delta}$, respectively. Thus:

$$a_L \equiv \frac{\bar{\delta}(1-\mu)\sigma N \eta}{H - \lambda\gamma\bar{\delta}(1-\mu)\sigma N} > a_H \equiv \frac{\underline{\delta}(1-\mu)\sigma N \eta}{H - \lambda\gamma\underline{\delta}(1-\mu)\sigma N}. \tag{35}$$

Politicians will be bribed to maintain the anti-competitive policy, $\overline{\delta}$, as long as $a_{t-1} \leq a_L$. Similarly, they will be bribed to switch from competitive to the anti-competitive policies when $a_{t-1} \leq a_H$. That $a_L > a_H$ follows because capitalists make greater profits with low competition and have greater 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 cutoffs, a_L and a_H, are decreasing functions of H, because more honest politicians will be harder to convince to pursue the policy preferred by the capitalist lobby.

FIGURE 7 HERE

Figure 7 summarizes this pattern diagrammatically. When $a \in (a_r(\mu, \underline{\delta}), a_{WM})$, politicians choose anti-competitive policies without bribes, since this is the consumption-maximizing policy. Outside this range, anti-competitive policies will only be chosen when the lobby pays sufficient bribes to politicians. When when $a \leq a_H$, irrespective of current policy the capitalist lobby can pay enough bribes, and the politician chooses the anti-competitive policies. In contrast, when $a \geq a_L$, the politician cannot be bribed and the political equilibrium will involved high competition. Finally and perhaps most interestingly, if $a \in (a_H, a_L) \setminus (a_r(\mu, \underline{\delta}), a_{WM})$, the outcome is *history dependent.* If competition is initially low, capitalists enjoy greater monopoly profits and are sufficiently wealthy to successfully lobby to maintain the anti-competitive policies $(\overline{\chi}, \overline{\delta})$. If competition is initially high, capitalists make lower profits and do not have enough funds to buy politicians, so there is no effective lobbying activity, and equilibrium policies are competitive.

To discuss the possibility of political economy traps, now assume that

$$\underline{\delta} < \delta^*(\mu) < \overline{\delta}, \tag{36}$$

where recall that $\delta^*(\mu)$ is the threshold competition level such that $a_r(\mu, \delta^*(\mu)) = a_{trap}$ defined in (30). This assumption implies that $a_r(\mu, \underline{\delta}) < a_{trap} < a_r(\mu, \overline{\delta})$. So a non-convergence trap will arise when anti-competitive policies, $\delta = \overline{\delta}$, are being pursued, while with $\delta = \underline{\delta}$, the economy will switch out of the investment-based strategy before it reaches a_{trap}, and it will continue to converge to the world technology frontier. Whether the economy will get stuck in a non-convergence trap therefore depends on whether the political process leads to a switch from anti-competitive policies, $\overline{\delta}$, to more competitive policies, $\underline{\delta}$, before a_{trap} is reached. Also to simplify the discussion, in the rest of this section we assume that $(a_r(\mu, \underline{\delta}), a_{WM}) = \emptyset$, which removes the case where the politician chooses low competition without receiving bribes.

Next consider the evolution of an economy starting with initial level of technology $a_0 < a_L$, and with a low initial level of competition $\delta = \overline{\delta}$. Then, the politician will be bribed into maintaining low competition as long as a remains below a_L. If we also have

$$a_{trap} \leq a_L, \tag{37}$$

then the economy will reach a_{trap} with anti-competitive policies, $\delta = \overline{\delta}$. Assumption (36) implies that at this point it will be pursuing the investment-based strategy, and get stuck in a non-convergence trap. We think of this type of non-convergence trap as a "political economy trap", since the reason why the economy fails to switch from the investment-based to the innovation-based strategy is successful lobbying by the capitalists in favor of anti-competitive policies. In this case, if the economy ever reached distance to frontier $a = a_L$, it would switch to high competition and to an innovation-based strategy, and would eventually converge to the frontier. But this stage is never reached since convergence stops at $a = a_{trap} < a_L$.

In contrast, if (37) does not hold, then eventually, a will exceed a_L, and the capitalist lobby will no longer be able to capture the politician, and the economy will revert to the high competition policy, $\delta = \underline{\delta}$, switch to the innovation-based strategy, and converge to the frontier. Inspection of (37) shows that it is more likely to be satisfied when H is low, that is, when the political system is more corruptible. Therefore, in societies with weak political institutions, political economy traps are more likely and government intervention is more "risky" (potentially much more costly, especially for long-run growth).

Notice also that whether or not a political economy trap arises depends on initial conditions. For example, suppose that the economy starts with initial distance to frontier $a_0 > a_H$, but differently from before, with competitive policies, $\delta = \underline{\delta}$. Then, as Figure 6 shows, the capitalist lobby will not have enough funds to bribe the politician, because with $\delta = \underline{\delta}$ profits are low, and the economy will remain competitive, i.e., with $\delta = \underline{\delta}$. Assumption (36) then ensures that the economy switches to the innovation-based strategy before a_{trap} and converges to the frontier.

This discussion establishes:

Proposition 3 Suppose that Assumptions (A1), (A2), (27) and (36) hold. Suppose also that competition policy, $\delta \in \{\underline{\delta}, \overline{\delta}\}$, is decided by a sequence of politicians with honesty cost H, potentially receiving bribes from the capitalist lobby according to equation (33), and that $(a_r(\mu, \underline{\delta}), a_{WM}) = \emptyset$.

Then, there exists a threshold level a_L such that as long as $a < a_L$, starting with $\delta_0 = \overline{\delta}$ the politician will be bribed into maintaining a low level of competition, $\delta = \overline{\delta}$, and a threshold level $a_H < a_L$ such that as long as $a < a_H$ the politician will always be bribed into maintaining or switching to a low level of competition, $\delta = \overline{\delta}$. Both a_H and a_L are decreasing in H. The dynamic equilibrium takes the following form:

1. If (37) holds, and the economy starts with $a_0 < a_L$ and $\delta_0 = \overline{\delta}$ or with $a_0 < a_H$, then the equilibrium will feature bribes and the investment-based strategy throughout. It will grow until it reaches a_{trap}, and then it will be stuck in a non-convergence trap.

2. If (37) does not hold, and the economy starts with $a_0 < a_L$ and $\delta_0 = \overline{\delta}$ or with $a_0 < a_H$, then the equilibrium will start with bribes and the investment-based strategy. Bribes will stop and there will be a switch to the innovation-based strategy at a_L, before a_{trap} is reached, and the economy will converge to the frontier.

3. If the economy starts with $a_0 > a_H$ and $\delta = \underline{\delta}$, then there will be no bribes in equilibrium, the economy will switch out of the investment-based strategy and converge to the frontier.

This proposition therefore demonstrates the existence of multiple steady state equilibria, one with the political economy trap and no convergence to the frontier, and the other without the trap and convergence to the frontier. Which of these two long-run equilibria emerges depends on whether the economy starts with low or high competition and on its initial level of development. Since a_H and a_L are decreasing in H, the proposition also shows that political economy traps are more likely when there are few checks and balances on politicians.

6 Conclusion

In this paper, we proposed a model where the economic organization of technological leaders and followers differ systematically. In the model economy, equilibrium in technological leaders corresponds to an innovation-based strategy, with younger firms, greater churning and greater selection, while technological followers pursue an investment-based strategy, with greater investment and long-term relationships. We use to this model to evaluate the pros and cons of investment-based and innovation-based strategies.

In the model economy, firms engage both in copying and adoption technologies from the world frontier and in innovation activities. The selection of high-skill entrepreneurs and firms is more important for innovation than for adoption. Consequently, as the economy approaches the frontier, selection becomes more important. As a result, countries that are far from the technology frontier pursue an investment-based strategy, with 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, and 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 show 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.

Equally interesting, we find that retained earnings may shield insiders so much that some societies may never switch out of the investment-based strategy, and these societies never converge to the world technology frontier. The reason is that they fail to take advantage of the innovation opportunities that require selection. Consequently, policies encouraging investment-based strategies might also lead to non-convergence traps.

The growth-maximizing (and welfare-maximizing) policy sequence is therefore a set of policies encouraging investment and protecting insiders, such as anti-competitive policies, at the early stages of development, followed by more competitive policies. Such a sequence of policies creates obvious political economy problems. Beneficiaries of existing policies can bribe politicians to prevent policy reform. Moreover, these groups, in our model the capitalists, will be politically powerful precisely because they are benefiting from the less competitive policies in place. In this context, the model also sheds some light on the debate about whether government intervention should be more prevalent in less developed countries. The answer suggested by the model is that, abstracting from political economy considerations, there may be greater need for government intervention when the economy is relatively backward. But unless political institutions are sufficiently developed (or become developed in the process of economic growth) and impose effective constraints on politicians and elites, such government intervention may result in the capture of politicians by groups that benefit from the intervention, paving the way for political economy traps.

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 and the type of equilibrium financial and employment relationships, and the political economy implications of these differences, constitute a very interesting, and relatively underexplored, area for future research.

References

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

[2] Acemoglu, Daron, Simon Johnson and Todd Mitton (2003) "Institutions and Market Structure" MIT mimeo.

[3] Acemoglu, Daron and Fabrizio Zilibotti, (1997) "Was Prometheus Unbound by Chance? Risk Diversification and Growth." *Journal of Political Economy*, vol.105, pp.709-51.

[4] Aghion, Philippe, and Peter Howitt, (1992) "A Model of Growth Through Creative Destruction." *Econometrica*, vol. 60, pp. 323-51.

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

[6] Arellano, Manuel and Stephen Bond, (1991) "Some Tests of Specification for Panel Data: Monte Carol Evidence and an Application to Employment Equations" *Review of Economic Studies* Vol 58, pp 277-97.

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

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

[9] Dewatripont, Mathias and Eric Maskin, (1995) "Credit and Efficiency in Centralized and Decentralized Economies", *Review of Economic Studies,* Vol. 62, No. 4., pp. 541-555.

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

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

[12] Frankel, Jeffrey and David Romer, (1999) "Does Trade Cause Growth?" *American Economic Review*, vol. 89, pp. 379-399.

[13] Galor, Oded and Daniel Tsiddon, (1997) "Technology, Mobility, and Growth", *American Economic Review*, vol. 87, pp.363-382.

[14] Gerschenkron, Alexander, (1962) *Economic Backwardness in Historical Perspective*, Harvard University Press, Cambridge MA.

[15] Greenwood, Jeremy, and Boyan Jovanovic, (1990) "Financial Development, Growth, and the Distribution of Income", *Journal of Political Economy*, vol. 98, No. 5, Part 1. pp. 1076-1107.

[16] Griffith, Rachel, Stephen Redding and John Van Rennan (2003) "Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries" University College London, mimeo, forthcoming *Review of Economics and Statistics.*

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

[18] Grossman, Gene and Elhanan Helpman, (2001) *Special Interest Politics,* MIT Press, Cambridge MA and London UK

[19] Hassler, John and Jose V. Rodriguez Mora, (2000) "Intelligence, Social Mobility, and Growth", *American Economic Review* vol. 90: pp. 888-908

[20] Hausmann, Ricardo and Dani Rodrik, (2002) "Economic Development As Self-Discovery", Mimeo Harvard University.

[21] Howitt, Peter, (2000), "Endogenous Growth and Cross-Country Income Differences", *American Economic Review,* vol. 90, pp. 829-46.

[22] Howitt, Peter and David Mayer, (2002) "Technological Inovation, Implementation and Stagnation: A Schumpeterian Theory of Convergence Clubs", Mimeo Brown University

[23] King, Robert G. and Ross Levine, (1993) "Finance and Growth: Schumpeter Might Be Right" *Quarterly Journal of Economics*, vol. 108, No. 3. pp. 717-737.

[24] Kang, David, (2002) *Crony Capitalism: Corruption and Development and South Korea and the Philippines,* Cambridge, Cambridge University Press.

[25] Nelson, Richard R. and Edmund S. Phelps, (1966): "Investment in Humans, Technological Diffusion and Economic Growth", *American Economic Review* vol. 56: pp. 69-75.

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

[27] Shleifer, Andrei and Robert Vishny, (1999) *The Grabbing Hand : Government Pathologies and Their Cures ,* Harvard University Press, Cambridge MA

[28] Stiglitz, Joseph E., (1995), *Whither Socialism*?, MIT Press. Cambridge, MA

[29] Tong, Jian and Chenggang Xu, (2001) "Endogenous Financial Institutions, R&D Selection and Growth", Mimeo London School of Economics.

[30] Wooldridge, Jeffrey (2002) *Econometric Analysis of Cross-section and Panel Data*

7 Appendix A: Details on empirical work and data

In this appendix, we provide details on the empirical results briefly presented in the Introduction. The sample of countries includes all non-OECD (including those that joined the OECD in the 1990s, such as Korea and Mexico) and all non-socialist countries for which we have data. 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, La Porta, Lopez-de-Silanes, and Shleifer (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 procedures is smaller or equal to 10 and into the "high-barriers" group otherwise. This implies that 19 countries are classified as high barrier and 23 countries as low barrier.[25] The barrier measures are time invariant, and hence so is our classification.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, while 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. Distance to frontier is defined as the ratio of the country's GDP to the U.S. GDP at the beginning of the sample. For the cross-sectional regressions, per capita GDP growth rates are for 1965-95, and the initial data are for 1965. The output data are from the Summers-Heston data set, obtained from the National Bureau of Economic Research.

Appendix Table 1 reports a number of regressions using this sample. The first column reports the regressions shown in Figure 1a. We define two dummy variables, HB and LB. HB is equal to 1 for high-barrier countries and zero otherwise, while LB takes the value 1 for the low-barrier countries. We also control for a dummy for sub-Saharan Africa, since sub-Saharan African countries have experienced much slower growth than the rest of the world during this time period, and 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 Easterly and Levine, 1997, on the role of ethno-linguistic fragmentation in explaining low growth in Africa). Thus the estimating equation is:

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

[25] The median number of procedures is 11 and 4 countries have exactly 11 procedures (these are Egypt, Indonesia, Kenya and Uganda). The results are robust to classifying these countries into the low-barrier group.

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 U.S. 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 above. The coefficients of interest are $\alpha_{1,HB}$ 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. We are particularly interested in the difference between $\alpha_{1,HB}$ and $\alpha_{1,LB}$.

The coefficients shown in column 1 correspond to those in Figures 1a and 1b. As was visible in these figures, there is a stronger negative relationship between distance to frontier and growth for high-barrier countries. For example, $\alpha_{1,LB}$ is estimated as -0.021 (s.e.=0.030), while $\alpha_{1,HB}$ is -0.073 (s.e.=0.029). 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, since the implication of our approach is that at any point in time, there should be a stronger relationship between distance 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:

$$g_{i,t} = \alpha_{0,HB} \cdot LB_i + \alpha_{0,LB} \cdot LB_i + \alpha_{1,HB} \left(\frac{y_{i,t-1}}{y_{U,t-1}} \cdot HB_i \right) + \alpha_{1,LB} \left(\frac{y_{i,t-1}}{y_{U,t-1}} \cdot LB_i \right) + \alpha_2 SA_i + f_t + \varepsilon_{it},$$

$g_{i,t}$ 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 i at date $t-1$, the f_t's denote a full set of time effects, and we take the time intervals to be 5 years.[26] The results are reported in column 2, and show a similar pattern, with no relationship between growth and distance to frontier for low-barrier countries, and a strong negative relationship for the high-barrier countries. For example, $\alpha_{1,LB}$ is now estimated to be 0.009 (s.e.= 0.015), while $\alpha_{1,HB}$ is -0.061 (s.e.= 0.017). The difference between these two coefficients is now statistically significant at the 1 percent level. The next column adds controls for years of schooling (we use male use of schooling from the Barro-Lee data set), and the pattern is unchanged.

The patterns shown in Columns 1-3 could be driven by some other omitted county characteristic. A stronger test of the implication of our model would be to see whether high-barrier countries slow down more significantly as they approach the frontier than the low-barrier countries.

[26] 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.

To investigate this column 4 changes the specification to

$$g_{i,t} = \alpha_{0,HB} \cdot LB_i + \alpha_{0,LB} \cdot LB_i + \alpha_{1,HB} \left(\frac{y_{i,t-1}}{y_{U,t-1}} \cdot HB_i \right) + \alpha_{1,LB} \left(\frac{y_{i,t-1}}{y_{U,t-1}} \cdot LB_i \right) + d_i + f_t + \varepsilon_{it},$$

where the d_i's denotes a full set of country 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 -0.036 (s.e.= 0.036), while $\alpha_{1,HB}$ is estimated at -0.105 (s.e.= 0.046). But the difference between the coefficients is once again statistically insignificant.

Nevertheless, the results in columns 3 and 4 are difficult to interpret because of the standard bias in models with fixed effects and lagged dependent variables (see, for example, Wooldridge, 2002, chapter 10). Distance to frontier is correlated with the lags of the dependent variable, since $g_{i,t} \approx (y_{i,t} - y_{i,t-1}) / y_{i,t}$. This creates a bias in the estimation of the fixed effects, and therefore in the estimates of the α_1's. To deal with this problem, in columns 5 and 6, we report regressions where distance to frontier is instrumented by its one-period lags.[27] The results are similar to those reported in columns 6 and 7: the estimate of $\alpha_{1,LB}$ is -0.033 (s.e.= 0.048), while $\alpha_{1,HB}$ is estimated at -0.206 (s.e.= 0.071). The difference between the coefficients $\alpha_{1,HB}$ and $\alpha_{1,LB}$ is statistically significant at the 5 percent level without years of schooling and at the 1 percent level with years of schooling in the regression

In the bottom panel of Appendix Table 1, we report regressions that do not split the sample, but interact the barrier variable with distance to frontier. Thus, the general model we are estimating is:

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

where 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. The interaction term, β_2, which loosely corresponds to the difference between $\alpha_{1,HB}$ and $\alpha_{1,LB}$, is negative and statistically significant (-0.009, with s.e.=0.003). This implies that high-barrier countries slow down more closer to the frontier. Other results, with the exception of those in column 4, also show a significant interaction term.

Appendix Table 2 has an identical form to Appendix Table 1, but splits the sample according to the degree of openness to international trade. Rather than using the ratio of actual trade to GDP, which is a highly endogenous variable, we use the degree of openness predicted by a standard "gravity" equation as in Frankel and Romer (1999). The

[27] This instrumental-variables strategy leads to consistent estimates. Estimators that use additional moment restrictions can improve efficiency, though they may also have undesirable small sample properties. See Wooldridge (2002).

gravity equation estimates degree of openness as a function of differences in population, land area, proximity and common borders to other countries, and whether a country is landlocked. As in the previous case, we split the sample into two parts, "open" and "closed" countries.[28] The results of interest in Appendix Table 2 are similar: both in the cross-sectional and the fixed effect regressions, the estimates of $\alpha_{1,LO}$, the distance to frontier term for low-openness countries, are substantially larger (more negative) than $\alpha_{1,HO}$, the distance the frontier term for high-openness countries, and this difference is statistically significant in the fixed-effect specifications. In the bottom panel, we estimate the model without splitting the sample, but with the interaction specification. The results are again similar, with the exception of the fixed-effects specification with years of schooling controls, which leads to the opposite of the sign we expect, though now the standard error is very very large, and the interaction coefficient is highly insignificant.

Finally, Appendix Table 3 reports results with the sample split according to the average years of schooling for males over 25. Countries are classified as low-education if the years of education was below 2.47 in 1965. Our mechanism suggests that skills, and thus human capital and schooling, should be more useful near the frontier, so we should find more negative estimate of $\alpha_{1,LE}$, the coefficient of distance to frontier for low-education countries than $\alpha_{1,HE}$, the coefficient of distance to frontier for high-education countries. The point estimates in the cross-sectional regressions are consistent with this, but have large standard errors, so the picture here is not clear. But again the fixed-effect regressions show considerable support for our hypothesis.

8 Appendix B: The participation constraint

In this Appendix we prove that when N is sufficiently large, the incentive compatibility constraint, (11), is always more binding than the relevant participation constraints.

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

$$PC_t(\nu \mid s,e=O,z) \equiv S_t(\nu \mid s,e=O,z) - \widehat{RE}_t(\nu \mid s,e=O,z) - w_t \geq 0, \qquad (38)$$

which simply states that the payments minus retained earnings that are injected must be

[28] The open economies are Barbados, Belize, Benin, Bostwana, Burundi, Cape Verde, Comoros, Congo, Costa Rica, Cote d'Ivoire, Cyprus, Dominica, Dominican Republic, El Salvador, Gabon, Gambia, Ghana, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hong Kong, Israel, Jamaica, Jordan, Korea, Lesotho, Malaysia, Mauritania, Mauritius, Namibia, Nicaragua, Panama, Rwanda, Senegal, Seychelles, Sierra Leone, Singapore, Syria, Taiwan, Togo, Trinidad and Tobago, and Tunisia, while the close economies are Angola, Argentina, Bangladesh, Bolivia, Brazil, Burkina Faso, Cameroon, Central African Republic, Chile, Columbia, Democratic Republic of Congo, Ecuador, Egypt, Ethiopia, Fiji, India, Indonesia, Iran, Kenya, Madagascar, Malawi, Mexico, Morocco, Mozambique, Nepal, Nigeria, Pakistan, Papua New Guinea, Paraguay, Peru, South Africa, Sri Lanka, Tanzania, Thailand, Uganda, Uruguay, Venezuela, Zambia and Zimbabwe.

greater than the wage rate. We will ensure that these participation constraints hold even when all entrepreneurs inject all their retained earnings, i.e., when $\widehat{RE}_t(\nu \mid s, e = O, z) = RE_t(\nu \mid s, e = O, z)$.

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

$$PC_t(\nu \mid s, e = Y, z) \equiv S_t(\nu \mid s, e = Y, z) + \frac{1}{1+r} E_t \text{Rent}_{t+1} - w_t \geq 0, \tag{39}$$

where the expected future rent is given by

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

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 this type, and he will receive future rents when he has high skill or when he has low skill and the economy is in the investment-based regime, i.e., $R_t = 1$.

We prove the main result in three steps. First, we prove that $\exists N_Y < \infty$ such that, for $N \geq N_Y$, the participation constraint is slack for young entrepreneurs. Second, we show that, if the participation constraints holds for the old low-skill entrepreneurs, it holds, *a fortiori*, for the old high-skill entrepreneurs. Finally, we prove that $\exists N_{OL} < \infty$ such that, if $N \geq N_{OL}$, then the participation constraint is slack for old low-skill entrepreneurs. Therefore, if $N \geq \max\{N_Y, N_{OL}\}$, then the participation constraints both for young and old entrepreneurs our slack.

For the first step, note that since the young have no retained earnings, a sufficient condition for the participation constraint not to bind when the incentive constraint binds is that $\mu\pi_t(\nu \mid \sigma, Y, L) \geq w_t$. Using equation (3) for equilibrium profits, and the equilibrium wage equation (5), we can re-express his participation constraint (39) as:

$$\mu\delta\sigma N(\eta\bar{A}_{t-1} + \lambda\gamma A_{t-1}) + \frac{1}{1+r} E_t \text{Rent}_{t+1} \geq \left((1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}\right) A_t. \tag{40}$$

Since E_tRent> 0 and $A_t \leq (1+g)\bar{A}_{t-1}$, a sufficient condition for (42) to hold for all a's is

$$N \geq \frac{(1+g)(1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}}{\mu\delta\sigma\eta} \equiv N_Y \tag{41}$$

To establish the second step, suppose that $(1+g) \geq (1+r)\sigma$, which will be satisfied as long as the economy is dynamically efficient, i.e., $r \geq g$. Then for $s \in \{\sigma, 1\}$, we have:

$$\mu\pi_t(\nu \mid 1, O, H) - RE_t(1 \mid \sigma, O, H) \geq \mu\pi_t(\nu \mid 1, O, L) - RE_t(\nu \mid 1, O, L). \tag{42}$$

To see why, replace π_t and RE_t by their expressions to obtain:

$$\mu\delta N\left(\eta\bar{A}_{t-1}+\lambda\gamma A_{t-1}\right)-\frac{1+r}{1+g}\mu\sigma\delta N\left(\eta\bar{A}_{t-1}+(1+g)\,\lambda\gamma A_{t-2}\right)\geq\mu\delta N\eta\bar{A}_{t-1}-\frac{1+r}{1+g}\mu\sigma\delta N\eta\bar{A}_{t-1},$$

where the inequality follows from the observation that:

$$\mu\delta N\lambda\gamma\frac{A_{t-1}}{A_{t-2}}\geq(1+g)\,\mu\delta N\lambda\gamma\geq(1+r)\,\mu\sigma\delta N\lambda\gamma.$$

Combining (42) with (38) shows that if the participation constraint holds for old low-skill entrepreneurs, it holds, *a fortiori*, for old high-skill entrepreneurs.

Therefore, to establish that the participation constraints of old entrepreneurs is slack, we only need to show $\mu\pi_t\left(\nu\mid 1,O,L\right)-RE_t\left(\nu\mid 1,O,L\right)\geq w_t$. Substituting for π_t, w_t and RE_t, and dividing through by $\bar{A}_{t-1}$, we can re-express this condition as:

$$\mu\delta N\eta\bar{A}_{t-1}-\frac{1+r}{1+g}\mu\sigma\delta N\eta\bar{A}_{t-1}\geq(1-\alpha)\,\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}A_t,$$

Rearranging terms, and noting that $A_t\leq(1+g)\,\bar{A}_{t-1}$, proves that if

$$N\geq\frac{(1+g)\,(1-\alpha)\,\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}}{\left(1-\frac{1+r}{1+g}\sigma\right)\mu\delta\eta}\equiv N_{OL},\tag{43}$$

then the participation constraint is slack. Therefore, if $N\geq\max\left\{N_Y,N_{OL}\right\}$, then both participation constraints are slack, even when entrepreneurs inject all their retained earnings.

9 Appendix C: Equilibrium without Assumptions (A1) and (A2)

9.1 Relaxing Assumption (A1)

When Assumption (A1) does not hold, young entrepreneurs will also undertake large projects for some values of a_t. In particular, a comparison of (17) and (18) shows that $s\left(e=Y,\cdot\right)=1$ whenever

$$a_{t-1}\geq a_s\left(\mu,\delta\right)\equiv\frac{\kappa/\delta N-(1-\mu)(1-\sigma)\eta}{(1-\mu)(1-\sigma)\lambda\gamma}.$$

(Note that Assumption (A1) guaranteed that $a_s\left(\mu,\delta\right)>1$, so this case never arose in the text).

Now suppose that $s\left(e=Y,\cdot\right)=1$, then it is straightforward to see that old low-skill entrepreneurs will be retained when: $(1-\mu)\delta N\eta+\frac{1+r}{1+g}\mu\delta\eta N>(1-\mu)\delta N(\eta+\lambda\gamma a_{t-1})$, or when

$$a_{t-1}<\tilde{a}_r\left(\mu\right)\equiv\frac{1+r}{1+g}\frac{\mu}{1-\mu}\frac{\eta}{\lambda\gamma}.\tag{44}$$

We can also see that whenever $a_s(\mu,\delta) > 0$, we have $a_r(\mu,\delta) > \tilde{a}_r(\mu)$, where $a_r(\mu,\delta)$ is the retention threshold given by (22) in the text, which applies when old entrepreneurs run large projects and young entrepreneurs run small projects.

Consequently, the analysis in the text applies whenever $a_s(\mu,\delta) > a_r(\mu,\delta)$, but now after a_t reaches $a_s(\mu,\delta)$, young entrepreneurs also run large projects. In contrast, when $a_s(\mu,\delta) < a_r(\mu,\delta)$, the retention threshold is $\tilde{a}_r(\mu)$. This is because in this case, we must also have $a_s(\mu,\delta) < \tilde{a}_r(\mu) < a_r(\mu,\delta)$, so by the time $\tilde{a}_r(\mu)$ arrives young entrepreneurs are already running large projects.[29] Therefore, the equilibrium sequence is as follows:

- If $a_s(\mu,\delta) > a_r(\mu,\delta)$, then we have: $s(e=Y,\cdot) = \sigma$ and $R = 1$ when $a_{t-1} \leq a_r(\mu,\delta)$, $s(e=Y,\cdot) = \sigma$ and $R = 0$ when $a_r(\mu,\delta) < a_{t-1} \leq a_s(\mu,\delta)$, and $s(e=Y,\cdot) = 1$ and $R = 0$ when $a_s(\mu,\delta) < a_{t-1}$.

- If $a_s(\mu,\delta) < a_r(\mu,\delta)$, then we have: $s(e=Y,\cdot) = \sigma$ and $R = 1$ when $a_{t-1} \leq a_s(\mu,\delta)$, $s(e=Y,\cdot) = 1$ and $R = 1$ when $a_s(\mu,\delta) < a_{t-1} \leq \tilde{a}_r(\mu)$, and $s(e=Y,\cdot) = 1$ and $R = 0$ when $\tilde{a}_r(\mu) < a_{t-1}$.

Turning to the growth-maximizing threshold, it is clear that terminating old low-skill entrepreneurs is always beneficial whenever both young and old entrepreneurs are running large projects. This implies that the growth-maximizing retention threshold becomes: $\hat{a}_1 = \min\langle \hat{a}, a_s(\mu,\delta)\rangle$, where $\hat{a}$ given by (29) is the growth-maximizing threshold given in the text when young entrepreneurs run small projects and old entrepreneurs run large projects.

9.2 Relaxing Assumption (A2)

Suppose that, contrary to (A2), we have $RE \equiv \frac{1+r}{1+g}\sigma\mu\delta N\eta \geq \kappa$. In this case, in equation (20) we have $\kappa\bar{A}_{t-1} \leq RE_t$, so $V_t(\nu \mid s=1, e=O, z=L) = (1-\mu)\delta N\eta\bar{A}_{t-1}$, and old low-skill entrepreneurs are retained whenever

$$a_{t-1} \leq \hat{a} \equiv \frac{(1-\sigma)\eta}{\sigma\lambda\gamma},$$

so in this case, the equilibrium and growth-maximizing thresholds coincide.

10 Appendix D: Proof of $s(e=O, z=H) = 1$ for all $a \in [0,1]$

In this appendix, we prove that, if $a_r(\mu,\delta) > 0$, then old high-skill entrepreneurs will operate large projects for all values of a_{t-1}. Define

$$\begin{aligned}\Delta_L &\equiv V_t(\nu \mid s=1, e=O, z=L) - V_t(\nu \mid s=\sigma, e=O, z=L),\\ \Delta_H &\equiv V_t(\nu \mid s=1, e=O, z=H) - V_t(\nu \mid s=\sigma, e=O, z=H).\end{aligned}$$

[29] To see this notice that $\tilde{a}_r(\mu) = (1-\sigma)a_s(\mu,\delta) + \sigma a_r(\mu,\delta)$, i.e., since $\sigma \in (0,1)$, it is a convex combination of $a_s(\mu,\delta)$ and $a_r(\mu,\delta)$. Then $a_s(\mu,\delta) < a_r(\mu,\delta)$ immediately ensures that $a_s(\mu,\delta) < \tilde{a}_r(\mu) < a_r(\mu,\delta)$.

We first prove that

$$a_r(\mu,\delta) > 0 \Rightarrow \Delta_L > 0 \tag{45}$$

for all values of a_{t-1}. To see why, first note that

$$\Delta_L = \left((1-\mu)(1-\sigma)\delta N\eta + \frac{1+r}{1+g}\sigma\mu\delta N\eta - \kappa\right)\bar{A}_{t-1}$$

is independent of a_{t-1} (or of A_{t-1}). Next, note that $a_r(\mu,\delta) > 0$ implies that, for all $a_{t-1} < a_r(\mu,\delta)$,

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

which in turn implies that $\Delta_L > 0$.

Next, we prove that

$$\Delta_L > 0 \Rightarrow \Delta_H > 0 \tag{46}$$

for all A_{t-1} and A_{t-2}. Suppose not. Then, given $\Delta_L > 0$, there should exist A_{t-1} and A_{t-2} such that: $\Delta_H = (1-\mu)(1-\sigma)\delta N\left(\eta\bar{A}_{t-1} + \gamma A_{t-1}\right) + \max\left\{RE_t^H - \kappa\bar{A}_{t-1}, 0\right\} < 0$, where $RE_t^H = \frac{1+r}{1+g}\sigma\mu\delta N\left(\eta\bar{A}_{t-1} + (1+g)\gamma A_{t-2}\right) \geq RE_t$ denotes the retained earnings of a high-skill entrepreneur. Moreover, if $\Delta_L > 0$ and $\Delta_H < 0$, then, $\Delta_L - \Delta_H > 0$. However,

$$\Delta_L - \Delta_H = -(1-\mu)(1-\sigma)\delta N\gamma A_{t-1} - \max\left\{RE_t^H - \kappa\bar{A}_{t-1}, 0\right\} + \left(RE_t - \kappa\bar{A}_{t-1}\right) < 0$$

which is a contradiction. This proves (46). Finally, (45) and (46) establish our initial claim.

11 Appendix E: 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 $\beta \equiv 1/(1+r)$, i.e., 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 $\int_0^1 k_t(\nu)\,d\nu = \frac{\kappa}{2}\bar{A}_{t-1}$ 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 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 will retain 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, i.e.,

$$a_{mfb} \equiv \frac{\eta(1-\sigma) - \kappa/\zeta N}{\sigma\lambda\gamma}. \tag{47}$$

This can be compared with the growth-maximizing policy. Since 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 also cares 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, while generating first-order gains in productivity, A_t, and in the present discounted value of future consumption. Thus, the non-myopic planner will choose a threshold, $a_{fb} > a_{mfb}$. 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, since 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 switches to an innovation-based strategy too late, since it would feature $a_r(\mu, \delta) > \hat{a}$. We can also verify that an economy with sufficiently small μ switches to an innovation-based strategy ($R_t = 0$) too soon relative to the welfare-maximizing allocation, i.e., $a_r(\mu, \delta) < a_{fb}$. 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 (22)), except that here ζ replaces δ in (22). 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 $a_{fb} > a_r(\mu, \delta)$.