NBER WORKING PAPER SERIES

DISTANCE TO FRONTIER, SELECTION, AND ECONOMIC GROWTH

Daron Acemoglu
Philippe Aghion
Fabrizio Zilibotti

Working Paper 9066
http://www.nber.org/papers/w9066

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2002

We thank seminar participants at Birkbeck College, Canadian Institute of Advanced Research, DELTA, MIT, SITE Stockholm, Universite de Toulouse, Universita Bocconi, and Mathias Dewatripont, Byeongju Jeong and Abhijit Banerjee for helpful comments, and Mauricio Prado for research assistance. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Distance to Frontier, Selection, and Economic Growth
Daron Acemoglu, Philippe Aghion and Fabrizio Zilibotti
NBER Working Paper No. 9066
July 2002
JEL No. O31, O33, O38, O40, L16

ABSTRACT

We analyze an economy where managers engage both in the adaptation of technologies from the world frontier and in innovation activities. The selection of high-skill managers is more important for innovation activities. As the economy approaches the technology frontier, selection becomes more important. As a result, countires at early stages of development pursue an investment-based strategy, with long term relationships, high average size and age of firms, large average investments, but little selection. Closer to the world technology frontier, there is a switch to innovation-based strategy with short-term relationships, younger firms, less investment and better selection of managers. We show that relatively backward economies may switch out of the investment-based strategy too soon, so certain economic institutions and policies, such as limits on product market competition or investment subsidies, that encourage the investment-based strategy may be beneficial. However, societies that cannot switch out of the investment-based strategy fail to converge to the world technology frontier. Non-convergence traps are more likely when policies and institutions are endogenized, enabling beneficiaries of existing policies to bribe politicians to maintain these policies.

Daron Acemoglu
Department of Economics
MIT, E52-380B
50 Memorial Drive
Cambridge, MA 02142-1347
and NBER
daron@mit.edu

Philippe Aghion
Department of Economics
Harvard University
Littauer Center 309
Cambridge, MA 02138
and NBER
p_aghion@harvard.edu

Fabrizio Zilibotti
University College London
Gower Street
London WC1E 6BT
Tel: 00 44 20 7679 5852
f.zilibotti@ucl.ac.uk

> "... 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 (1962, p. 7)

1 Introduction

In his famous essay, *Economic Backwardness in Historical Perspective*, Gerschenkron argued that relatively backward economies, such as Germany, France, Belgium and Russia during the nineteenth century, could rapidly catch up to more advanced economies by introducing "appropriate" economic institutions to encourage investment and technology adoption. He emphasized the role of long-term relationships between firms and banks, of large firms and of state intervention. Underlying this view is the notion that relatively backward economies can grow rapidly by investing in, and adopting, already existing technologies, or by pursuing what we call an *investment-based growth strategy*. If this assessment is correct, the institutions that are appropriate to such nations should encourage investment and technology adoption, even if this comes at the expense of various market rigidities and a relatively less competitive environment.

Although there are numerous cases in Africa, the Caribbean, Central America and South Asia where state involvement in the economy has been disastrous (e.g., Acemoglu, Johnson and Robinson, 2001), various pieces of evidence are consistent with the notion that rapid investment-based growth is possible with, or even sometimes encouraged by, relatively rigid institutions and considerable government involvement. These include the experiences of a number of European countries during the nineteenth century discussed by Gerschenkron (1962), the correlation between tariff rates and economic growth in the nineteenth century among countries following the technological leader, Britain (O'Rourke, 2000, but see also Irwin, 2002, for a different interpretation); the correlation between protection of high-skill and high-tech industries and economic growth in the postwar period (Nunn and Trefler, 2002); and the relatively rapid growth of economies pursuing import-substitution and infant-industry protection policies, such as Brazil, Mexico, Peru and Turkey, until the mid-1970s.

Perhaps the two most well-known cases of rapid investment-based growth are the post-war experiences of Japan and South Korea. In Japan, the Ministry of International Trade and Industry (MITI) played a crucial role by regulating foreign currency allocations, import licenses, and the extent of competition, by directing industrial activity and by encouraging investment by the *keiretsu*, the large groupings of industrial firms and banks (e.g., Johnson, 1982, Evans, 1995, Hoshi and Kashyap, 2002). In the Korean case, the large family-run conglomerates, the *chaebol* appear to have played an important role, especially in generating large investments and rapid technological development. The chaebol, similar to the keiretsu in Japan, received strong government support in the form of subsidized loans, anti-union legislation and preferential treatment that sheltered them from both internal and external competition. An additional important feature of both the chaebol and the keiretsu was their low managerial turnover, emphasis on long-term relationships and generally rigid structures (e.g., Wade, 1990, Vogel, 1991, Evans, 1995).

At the other extreme, we can think of the process of *innovation-based growth*, where the selection of successful managers and firms, as well as a variety of other innovation-type activities, are more important. Many view the current U.S. economy, with its market-based financing, important role for venture capital and high rate of business failures, and also the relatively competitive British economy of the nineteenth century as approximating this type of innovation-based growth. Interestingly, both the performance of economies pursuing the two different types of strategies and the views of the economics profession on the merits of the two strategies have varied significantly over time. While, until the mid-1970s or even the 1980s, a number of economies pursuing the investment-based strategy were successful and many economists were enthusiastic about this strategy, today the pendulum appears to have swung the other way.

This paper analyzes the equilibrium (and socially beneficial) choice between these two strategies. We discuss when each arises in equilibrium, their relative efficiencies for different stages of development, the possibility of traps and non-convergence resulting from the choice of an incorrect strategy, and the political economy of the two strategies.

The choice between the investment-based and innovation-based strategies involves a trade-off between investment and experience, on the one hand, and *selection*, on the other. Economies can often maximize investment by channeling money to existing firms, and making use of the experiences of established firms and managers. This is particularly the case in the presence of incentive problems, which are partly relaxed for existing firms and managers because of their retained earnings, thus increasing their investment capacity relative to newcomers. But the investment-based strategy also shelters less successful firms and managers from competition, and as a result, it involves less selection of successful firms and managers, worse matches between agents and economic activities, and less innovative activity by new entrants.

Our analysis builds on the notion that this trade-off between investment-based and innovation-based strategies changes over the course of development, especially as an economy approaches the world technology frontier. Relatively backward economies can grow with an investment-based strategy. In contrast, nearer the frontier, growth has to rely relatively more on innovation activities, thus selecting the right entrepreneurs, and the right matches between managers and economic activities, becomes more important.

Incorporating this trade-off into a standard endogenous technical change model, we show that economies will first tend to pursue an investment-based strategy with longer-term relationships, greater investment and less selection, and at some point may switch to an innovation-based strategy with greater selection, shorter-term relationships and younger firms. Moreover, we show that the equilibrium switch out of an investment-based strategy may occur sooner or later than the growth-maximizing or welfare-maximizing policies. Because monopolists do not appropriate the full returns from greater investments, the switch tends to occur too soon; and because retained earnings create an advantage for existing firms (insiders), the switch tends to occur too late.

As a result, similar to Gerschenkron's argument, in relatively backward countries, there may be room for government intervention, by direct subsidies, cheap loans and anti-competitive policies, to encourage the investment-based strategy. However, different from Gerschenkron's emphasis, an economy may fail to converge to the world technology frontier, precisely because it does not switch (or it switches too late) out of an investment-based to an innovation-based strategy. The reason is that after a certain stage of development, innovation activities are necessary to ensure further growth and convergence, and these activities are limited in the investment-based regime. Thus state intervention may have short-term benefits but also considerable long-term costs. Our model therefore implies that appropriate institutions and policies depend on the stage of development, thus justifying, and qualifying, both the views that emphasize certain beneficial effects from state intervention as well as those pointing out the inefficiencies resulting from these interventions.

The potential costs of policies encouraging investment-based growth become more substantial once we endogenize the political economy of government intervention. These policies enrich existing capitalists, who prefer the investment-based equilibrium to the innovation-based equilibrium. When economic power buys political power, it becomes difficult to reverse these policies that have an economic and politically powerful constituency.[1] An interesting implication of this political-economy analysis is that under certain circumstances societies may get trapped with "inappropriate institutions" and relatively backward technology, precisely because earlier they adopted appropriate institutions for their circumstances at the time, but in the process also creating a powerful constituency against change.[2]

An immediate implication of our analysis is a new theory of "leapfrogging". Economies that adopt policies encouraging the investment-based strategy may initially grow faster than others, but then get stuck in a non-convergence trap and taken over by the initial laggards. This is a very different view of leapfrogging from the standard one (e.g., Brezis, Krugman and Tsiddon, 1994), where leapfrogging emerges because of comparative advantage and learning-by-doing, and where the focus is on whether the world technological leadership is taken over by a newcomer. 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 anti-competitive and import substitution policies until the mid-1970s or early 1980s, stagnated and were taken over by other economies with relatively more competitive policies, such as Hong Kong or Singapore.[3]

[1] The Korean case illustrates the influence of economically powerful groups on politicians. 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).

[2] Both the Korean and the Japanese cases illustrate the dangers of the investment-based strategy, and the political economy problems created by such a strategy. The close links between government officials and the chaebol in the Korean case and bureaucrats and the keiretsu in the Japanese case, which appear to have been important for the early success of these economies, later became obstacles to progress, especially after the Asian crisis for Korea and after the mid-1980s for Japan. Following the crisis in Korea, a number of the chaebol went bankrupt, while others were split, or like Daewoo, were forced into restructuring. Interestingly, there seems to have been much less reform in this dimension in Japan.

[3] Brazil and Peru started with GDP per worker levels equal to, respectively, 18% and 22% of the U.S. in 1950. By the mid-70s, after growing at an average 3.5-4% per year, they had reached a GDP level equal to 35% of the U.S., and from there on stayed at this level, or declined relative to the U.S.. Mexico converged steadily from 33% of the U.S. level in 1950 to 63% in 1981, with a 3.5% average annual growth rate, but declined thereafter. In contrast, Hong Kong and Singapore, which started at 17% and 20% of the U.S. GDP per worker in 1960 (earlier data are not available), surpassed Brazil and Peru during the 1970s and Mexico during the 1980s. In 2000, their GDP per capita was, respectively, 70% and 73% relative to the U.S. (all numbers are PPP-adjusted, and from the Penn World Tables, or from World Factbook, 2001). It is also noteworthy that in the 63-country sample of Trefler and Nunn (2002), Hong Kong has the lowest tariff rates in their sample of 63 countries, Singapore has a relatively

At the heart of our analysis is the trade-off between investment-based and innovation-based strategies, which is founded on three ingredients: (1) Experienced managers and incumbent firms can undertake larger investments, and everything else equal, achieve greater technological improvements and productivity growth. (2) Managers copy and adopt well-established technologies from the world technology frontier, and managerial skill is not crucial for this type of copying and adoption activities. (3) Managers also undertake innovations or adapt technologies to local conditions, and managerial skill is essential in these tasks. This last point makes the *selection* of high-skill managers important for productivity growth.

All three ingredients are reasonable. That experienced managers and firms are more productive and can undertake larger investments is plausible, and consistent with evidence on firm-level learning-by-doing and investment patterns.[4] Moreover, this feature is introduced as an assumption only to simplify the basic model. We later show that in the presence of moral hazard, experienced managers will naturally invest more, because their earnings from previous periods relax the credit constraints implied by moral hazard. Similarly, that managers engage both in copying and imitation-type activities as well as innovation-type activities is natural and well accepted. The crucial ingredient here is that selection of high-skill managers is more important for innovation-type activities than for imitation. Although we do not have direct evidence on this point, given the nature of innovation activities, we regard this as a plausible assumption, and we also note that it is consistent with historical accounts. For example, Rosenberg (1982) emphasizes the speed of technology transfer and imitation in the presence of the right conditions, and concludes: "... the transfer of industrial technology to less developed countries is inevitable." (p. 270), while innovation and technological breakthroughs in advanced economies often require continuous efforts by various successful firms and the talents of many exceptional individuals (e.g., pp. 141-192).

Ingredients 1 and 2 above imply that managerial selection becomes more important when an economy is close to the world technology frontier. Ingredients 2 and 3, on

low rate, while Brazil and Mexico are two of the 5 countries with the highest tariffs (they do not have data on Peru).

[4] The innovation literature places considerable emphasis on the ability of incumbent and experienced firms to take advantage of incremental improvements (e.g., Arrow, 1974, Abernathy, 1980, Freeman, 1982, Nelson and Winter, 1982). Empirical work by, among others, Hirsch (1952), Lieberman (1984) and Bahk and Gort (1993), and more recently, by Irwin and Klenow (1994) and Benkard (2000) for the U.S. and Ohashi (2002) for Japan, document this type of learning-by-doing effects. Also, the relationship between firm age and firm size (or investment) is empirically well documented, e.g., Dunne, Roberts and Samuelson (1989).

the other hand, generate the key trade-off: an economy can either rely on selection (e.g., by terminating less successful managers) to generate more innovation, or sacrifice selection for experience, and take advantage of larger investments. All three ingredients together imply that in relatively backward economies, selection is less important, so an investment-based strategy exploiting experience is preferable. Closer to the frontier, the society needs selection, and therefore, it is more likely to adopt an innovation-based strategy. All the results in this paper build on the trade-off introduced by these three ingredients, and on the feature that closer to the world technology frontier, the selection of high-skill managers and successful firms becomes more important.

Our paper relates to a number of different literatures. First, the notion that managerial skill is more important for innovation than copying is reminiscent to the emphasis in Galor and Tsiddon (1997) and Hassler and Rodriguez (2001) on skill in times of economic change and turbulence. In a related contribution, Tong and Xu (2000) extend the model by Dewatripont and Maskin (1995) and compare "multi-financier" and "single-financier" contractual relationships as a function of the stage of development; the main idea of their paper is that while single-financier relationships tend to dominate at early stages of development when countries incur high sunk costs of R&D, multi-financier relationships tend to dominate at later stages of development when selecting good R&D projects becomes more important. But this model of financial contracting and growth does not deal with dynamic convergence aspects, and does not develop the contrast between innovation-based and investment-based growth strategies.

Second, our paper is related to the literature on the relationship between growth and contracting, including Acemoglu and Zilibotti (1999), Martimort and Verdier (2001) and Francois and Roberts (2001), as well as more generally, to the literature on growth and finance, including the papers by Greenwood and Jovanovic (1990), King and Levine (1993), La Porta et al (1997), Acemoglu and Zilibotti (1997), Rajan and Zingales (1998), Carlin and Mayer (2002) and Tadesse (2002). For example, Acemoglu and Zilibotti (1999) develop a model where informational problems become less severe as an economy develops, and derive implications from this for the organization of firms and markets. Martimort and Verdier (2001) and Francois and Roberts (2001) show how a high rate of creative destruction may discourage long-term relationships within firms.

Third, our model also relates to work on technological convergence and growth, in particular, to the papers by Barro and Sala-i-Martin (1997), Howitt (2000) and Howitt and Mayer (2002), which extend the growth framework of Aghion and Howitt (1992) to a multi-countries setup. Howitt and Mayer (2002), for example, analyze convergence clubs, prolonged stagnations, and twin-peak convergence patterns. But they do not provide an explicit treatment of institutions and contractual relations and they do not emphasize the trade-off between innovation-based and investment-based growth strategies.

Finally, our political economy section builds on the lobbying models by Grossman and Helpman (1997, 2001). While we simplify these models considerably in many dimensions, by introducing credit constraints on lobbies we also add a link between current economic power and political power. In this respect, our analysis is also related to Do (2002) who analyzes a lobbying model with credit-constrained agents, where income distribution affects policy.

Perhaps the most interesting link is between our approach and the existing debate on the optimal degree of government intervention in less developed countries. A number of authors, including Stiglitz (1995), call for government intervention in situations where externalities and market failures are rampant. Since less developed countries approximate these situations of widespread market failure, this reasoning, just like Gerschenkron's (1962) message, recommends greater government intervention in these countries. A recent paper by Hausmann and Rodrik (2001) pushes this line further and argues that most of the growth related activities in less developed countries create externalities because of potential imitation by others and learning-by-doing, and suggest that successful less developed countries have to rely on government intervention and subsidies, as was this case in South Korea and Taiwan. They write: "the world's most successful economies during the last few decades prospered doing things that are more commonly associated with failure," and propose similar "infant-industry-type" intervention for other countries. The same point of view is developed by many political scientists, including those working in the literature on "State Autonomy", for example, Johnson (1982), and more nuanced versions of these, such as the thesis of "Embedded Autonomy" by Evans (1995). These arguments are criticized by several economists and political scientists, however, because they ignore the potential for government failure. For example, Shleifer and Vishny (2000), argue that governments are often captured by interest groups or by politicians themselves. This suggests that in less developed countries, where checks on governments are weaker, the case for government intervention should be weaker as well. 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.

The rest of the paper is organized as follows. Section 2 outlines the basic model. Section 3 characterizes the equilibrium. Section 4 analyzes the implications of the equilibrium allocations for growth and convergence (or non-convergence) patterns, and compares the equilibrium allocation to growth-maximizing and welfare-maximizing policies. Section 5 shows how the results of the simpler model of Section 2 can be obtained in a more micro-founded model with a choice over project size. Section 6 discusses how government policy may be useful in creating "appropriate institutions" for convergence, but also how such policies may be captured by groups that are their main beneficiaries, creating political economy traps. Section 7 concludes.

2 The model

2.1 Agents and production

The model economy is populated by a continuum of overlapping generations of two-period lived agents. The population is constant. Each generation consists of a mass $1/2$ of "capitalists" with property rights on "production sites", but no managerial skill, and a mass $(L+1)/2$ of workers who are born without any financial asset but are endowed with managerial skills. Property rights are transmitted within dynasties.[5] Each worker is endowed with one unit of labor per unit of time, which she supplies inelastically without disutility. All agents are risk neutral and discount the future at the rate r.

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 production function:

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

where $A_t(\nu)$ is the 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, and $\alpha \in [0,1]$.

In each intermediate sector ν, one production site at each date has access to the most productive technology, $A_t(\nu)$, and so, this "leading firm" enjoys monopoly power. Each leading firm employs a manager, for production as well as for innovation, and incurs a setup cost, which is described in detail below. It then 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 also imitate this monopolist, and produce the same intermediate good, with the same productivity $A_t(\nu)$, but without using the production site or a manager. They correspondingly face greater costs of production, and need χ

[5] Alternatively, we could introduce a market for production sites, where capitalists at the end of their lives would sell their sites to younger agents. This would not change any of the results in the paper.

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). We will think of the parameter χ as capturing both technological factors and government regulation regarding competitive policy. A higher χ corresponds to a less competitive market, with the upper bound, $\chi = 1/\alpha$, corresponding to the situation of unconstrained monopoly. The fact that $\chi > 1$ implies that imitators are less productive than the incumbent producer in any intermediate good sector, 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 competition from imitators (e.g., Grossman and Helpman, 1991). This limit price is equal to the marginal cost of imitators:

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

so as to deter entry from the competitive fringe.

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

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

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 profit for monopolists.

Using (1), we also have that output is:

$$y_t = \alpha^{-1} \chi^{-\frac{\alpha}{1-\alpha}} A_t L. \tag{4}$$

where

$$A_t \equiv \int_0^1 A_t(\nu) \, d\nu. \tag{5}$$

is the average level of technology in this society. The market clearing wage level is, in turn, given by:

$$w_t = (1-\alpha) \alpha^{-1} \chi^{-\frac{\alpha}{1-\alpha}} A_t. \tag{6}$$

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

$$y_t^{net} = y_t - \int_0^1 x_t(\nu) \, d\nu = \zeta A_t L, \tag{7}$$

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, (7), and profits, (3), are identical except that the output has the term ζ instead of $\delta < \zeta$. This reflects an *appropriability effect*: the monopolists only capture a fraction of the greater productivity in the final goods sector (or of the consumer surplus) created by their production and productivity.

2.2 Technological progress and productivity growth

In every period and in each intermediate good sector, the leading firm can improve over the existing technology. Recall that leading firms are owned by capitalists, and as a result, a fraction 1/2 of these firms are young and a fraction 1/2 are mature (old). Managers, and managerial skills, are crucial for improvements in technology. Each manager selected to run a production site must make an investment of a fixed amount. These costs can be financed either through retained earnings, or by borrowing from a set of competitive intermediaries ("funds"), who collect earnings from other agents and lend them to managers. We assume that these intermediaries function without any costs. Then, returns are realized and shared between managers, intermediaries and capitalists according to the contractual arrangements between the three parties that are described below.

Managerial skills are firm-specific, and are revealed after the manager undertakes production and innovation activities in the first period of his relationship with the firm. We assume that a manager is high skill with probability λ and low skill with probability $1 - \lambda$. The assumption that managerial skills are firm-specific is made for simplicity, and does not affect any of the major results, but is also in line with the interpretation that what matters for successful innovation is the match between a particular manager and the activity he is engaged in.

We now make three important assumptions on the process of technological progress:

1. Experienced managers run larger projects and are, all else equal, more productive. Correspondingly, they create larger technological improvements.[6] This assumption captures the notion that, everything else equal, it is beneficial to have agents who have already acted as managers to continue in these tasks.

2. Managers *adopt* technologies from the frontier. Skills play a minor role in man-

[6] For now, this is an assumption. In Section 5, we show that the presence of moral hazard creates a natural tendency for experienced managers to run larger projects, since their retained earnings from the first period of activity help relax the credit constraints introduced by the moral hazard problems.

agers' success in technology adoption. This assumption captures the notion that relatively backward economies can grow by adopting already well-established technologies, and the adoption of these technologies is often relatively straightforward.

3. Managers also engage in *innovation* or *adaptation* of existing technologies to their local conditions. Managerial skills, and the match between the manager and the activity he is undertaking, matter for success in this activity. This assumption builds in the notion that managerial ability and skills, and therefore the selection of high-skill managers, are important for technological improvements.

First, let us denote the growth rate of the world technology frontier, $\overline{A}_t$, by g, i.e.,

$$\overline{A}_t = (1+g)^t \overline{A}_0. \tag{8}$$

We return to the determination of this growth rate below. All countries have a state of technology, A_t, defined by (5), less than the frontier technology, i.e., $A_t \leq \overline{A}_t$. We formulate the three above assumptions as follows: the productivity of intermediate good ν at time t is given by

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

where $s_t(\nu)$ is a term that depends on the experience of the manager; $\gamma_t(\nu)$ denotes the skill of the manager running this firm, and η is a positive constant. This equation states that, irrespective of the skill of the manager, all intermediate goods benefit from the state of world technology in the previous period, $\bar{A}_{t-1}$, by copying or adopting existing technologies. They also "innovate" over the existing body of local knowledge, A_{t-1}, and success in innovation depends on skill. Experience, $s_t(\nu)$, affects the productivity of both adoption and innovation activities.

Rearranging equation (9), we obtain a simpler equation

$$\frac{A_t(\nu)}{A_{t-1}} = s_t(\nu)\left(\eta \frac{\bar{A}_{t-1}}{A_{t-1}} + \gamma_t(\nu)\right). \tag{10}$$

Equation (10) shows the importance of *distance to technology 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 improvements in technology come from adoption of already well-established technologies. When $\bar{A}_{t-1}/A_{t-1}$ is small and the country is close to the frontier, innovations matter more. Thus as the country develops and approaches the technology frontier, innovation and adaptation of less well-established technologies, managerial skills and the quality of matches between managers and activities, and hence overall *managerial selection*, become more important.

For simplicity, we assume that the innovation component is equal to 0, i.e., $\gamma_t(\nu) = 0$, when the manager is low skill, and denote the productivity of a high-skill manager by $\gamma > 0$, i.e., $\gamma_t(\nu) = \gamma$ when manager in sector ν is high skill. Recall that a proportion λ of managers within each cohort are high skill. To guarantee a decreasing speed of convergence to the world technology frontier, we assume throughout that $\lambda\gamma < 1$.

The term $s_t(\nu)$ in (10) specifies the importance of experience in technological improvements. We assume that $s_t(\nu) = \sigma < 1$ in all cases, except when the firm rehires the same manager from the previous period, in which case the manager can make use of his firm-specific experience and $s_t(\nu) = 1$. This assumption implies that managers with firm-specific experience are more productive.

Finally, $k_t(\nu)$ denotes the investment that the manager in sector ν must make at time t in order to undertake the project, and we assume that $k_t(\nu) = \phi\kappa\bar{A}_{t-1}$ where $\phi < 1$, in all cases, except when the firm is employing the same manager as it did in the previous period, in which case $k_t(\nu) = \kappa\bar{A}_{t-1}$. There are two assumptions embedded in these expressions:

1. Costs grow with the level of world technology, $\bar{A}_{t-1}$. Intuitively, an important component of managerial activity is to undertake imitation and adaptation of technologies from the world frontier. As this frontier advances, managers need to incur greater costs to keep up with, and make use of, these technologies, hence investment costs increase with $\bar{A}_{t-1}$. This assumption ensures balanced growth.[7]

2. Experience enables managers to run larger projects. We can think of the greater productivity of experienced managers, $s_t(\nu) = 1$ instead of $s_t(\nu) = \sigma$ above, resulting, in part, from the ability of these managers to run larger projects. In Section 5, we show that the feature that experienced managers run larger projects does not need to be imposed as an assumption; once we endogenize the size of projects, experienced managers run larger projects than young managers because

[7] 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 $\rho = 1$, since it simplifies some of the expressions, without affecting any of our major results. We give the relevant key expressions for the case of $\rho < 1$ in the Appendix.

Note that for all cases where $\rho > 0$, an improvement in the world technology frontier, $\bar{A}_{t-1}$, increases both the returns and the costs of innovation. The parameter restriction $\sigma\delta\eta L > \phi\kappa$, which we impose below, is sufficient to ensure that the benefits always outweigh the costs.

their retained earnings relax the credit constraints imposed by moral hazard. Nevertheless, to simplify the exposition, we simply assume this feature here, and give the details of the model with project size choice in Section 5.

The setup described above introduces the key trade-off in this paper; that between experience and selection. Everything else equal, more experienced managers invest more, and generate more innovation and higher profits. However, some of the more experienced managers will have been revealed to be low skill, and low-skill managers are naturally less productive. So a society might either choose to have greater selection (a better allocation of managers to activities), and younger managers, but less managerial experience and investment; or less selection and older firms, but greater experience and investment. The trade-off between experience and selection will vary over the process of development because the importance of innovation vs. adoption of well-established technologies changes with distance the frontier, as captured by equation (10).

The state of local knowledge in the economy is summarized by the average of the productivity in various intermediate product sectors. To specify the law of motion of A_t note three things: (1) half of the firms will be young and the other half old; (2) average productivity among young firms is simply $A_{Yt} = \sigma\left(\eta\bar{A}_{t-1} + \lambda\gamma A_{t-1}\right)$, since they will hire young managers, a fraction λ of these will be high skill, with productivity $A_t(\nu) = \sigma\left(\eta\bar{A}_{t-1} + \gamma A_{t-1}\right)$, and the remainder will be low skill, with productivity $A_t(\nu) = \sigma\eta\bar{A}_{t-1}$ (recall equation (9)); (3) clearly, all managers revealed to be high skill will be retained, and average productivity among mature firms will depend on their decision whether to refinance low-skill managers.

Next, denote the decision to refinance a low-skill experienced manager by $R_t \in \{0,1\}$, with $R_t = 1$ corresponding to refinancing. More generally, $R_t = 1$ stands for all organizational decisions that make use of the skills of experienced managers or the expertise of established firms, especially to achieve greater investments, even if this comes at the expense of sacrificing managerial selection. Below we think of economies where $R_t = 1$ as pursuing *an investment-based strategy,* since these economies manage to invest more by making use of experienced managers. In contrast, economies with $R_t = 0$ are pursuing *an innovation-based strategy* where the emphasis is on maximizing innovation at the expense of investment. Using this notation, average productivity among mature firms is:

$$A_{Mt} = \begin{cases} \eta\bar{A}_{t-1} + \lambda\gamma A_{t-1} & \text{if } R_t = 1 \\ (\lambda + (1-\lambda)\sigma)\,\eta\bar{A}_{t-1} + (1 + (1-\lambda)\sigma)\,(\lambda\gamma A_{t-1}) & \text{if } R_t = 0. \end{cases}$$

The first line has exactly the same reasoning as for the average productivity of young firms. The second line follows from the fact that a fraction λ of the managers were revealed to be high skill, are retained, and have productivity $A_t(\nu) = \eta \bar{A}_{t-1} + \gamma A_{t-1}$, and the remaining $1 - \lambda$ of managerial posts are filled with young managers, who have average productivity $\sigma\left(\eta \bar{A}_{t-1} + \lambda\gamma A_{t-1}\right)$. Combining the productivity of young and mature managers, we have that

$$A_t = \begin{cases} \frac{1+\sigma}{2}\left(\eta \bar{A}_{t-1} + \lambda\gamma A_{t-1}\right) & \text{if } R_t = 1 \\ \frac{1}{2}\left(\left(\lambda + \sigma + (1-\lambda)\sigma\right)\eta \bar{A}_{t-1} + \left(1 + \sigma + (1-\lambda)\sigma\right)\lambda\gamma A_{t-1}\right) & \text{if } R_t = 0. \end{cases} \tag{11}$$

At this point, it is also is useful to introduce the notation of a_t to denote (the inverse of) the distance to frontier, i.e.,

$$a_t \equiv \frac{A_t}{\bar{A}_t}. \tag{12}$$

This variable is the key state variable in our analysis below. Using this definition and equations (8), (10), and (12), we can rewrite (11) as:

$$a_t = \begin{cases} \frac{1+\sigma}{2(1+g)}\left(\eta + \lambda\gamma a_{t-1}\right) & \text{if } R_t = 1 \\ \frac{1}{2(1+g)}\left(\left(\lambda + \sigma + (1-\lambda)\sigma\right)\eta + \left(1 + \sigma + (1-\lambda)\sigma\right)\lambda\gamma a_{t-1}\right) & \text{if } R_t = 0 \end{cases}, \tag{13}$$

which is the key dynamic equation in this economy.

2.3 Incentive problems

The final element of the environment in this economy is the incentive problems faced by firms. Managers engaged in innovative activities, or even simply entrusted with managing firms, are difficult to monitor. This creates a standard moral hazard problem, often resulting in rents for managers, or at the very least, in lower profits for firms. We formulate this problem in the simplest possible way, and assume that after output, innovations and profits are realized, the manager can appropriate a fraction μ of the returns for his own use, and will never be prosecuted. We think of the parameter μ as a measure of the importance of incentive problems, or equivalently, a measure of credit market imperfections resulting from these incentive problems. The key role of moral hazard in our model is to create current and future rents for managers, so that they can use their current rents to shield themselves from competition, and obtain the future rents.

Below we analyze both the case of no moral hazard, i.e., $\mu = 0$ (or μ small so that the incentive compatibility constraints are slack) and the economy with moral hazard where $\mu > 0$.

3 Equilibrium analysis

In this section, we characterize the equilibrium behavior of the economy described in Section 2. We start by specifying the contractual relations between capitalists (firms), intermediaries and managers, next define an equilibrium, and then characterize the equilibrium allocations with and without moral hazard. In the next section, we compare these equilibrium thresholds to the growth-maximizing and welfare-maximizing allocations.

3.1 Financial intermediation, contracts and equilibrium

Production requires a production site (owned by a capitalist), a manager, and financing to pay for the set-up cost of the project. Production sites are a scarce factor in this economy, since they allow the use of a superior technology. So the capitalists who own them will appropriate rents subject to satisfying the individual rationality and/or incentive constraints of intermediaries and managers.

We assume that capitalists make contractual offers to a subset of workers to become managers and to intermediaries. Investments are financed either through the retained earnings of managers, or through borrowing from intermediaries (recall that young capitalists and managers have no wealth to finance projects).[8] There is free entry into financial intermediation, and no cost of financial intermediation.

Omitting time subscripts when this causes no confusion, we use I_e^s, P_e^s, W_e^s and V_e^s to denote, respectively, the investment by intermediaries, the contractual payments to intermediaries, to managers and to firms, conditional on the skill level, s, and the experience, e, of the manager. In particular, $e = M$ denotes the case in which a mature firm employs the same manager as in the previous period (the case in which the firm-specific skills of the manager can be used), $e = Y$ denotes all other cases, and we use I_e to denote the investment in the case of an inexperienced manager of unknown skill. The sum of the payments to the three agents involved in each relation cannot exceed the total profits of the firm:

$$P_e^s + W_e^s + V_e^s \leq \pi_e^s, \tag{14}$$

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

where π_e^s is a profit level of the leading firm of age e with a manager of age-experience $e \in \{Y, M\}$, with skill s. The offer of the capitalist is a vector $(I_e^s, P_e^s, W_e^s, V_e^s)$.

Free entry into intermediation implies that intermediaries make zero (expected) profits. In addition, we assume that financial intermediation takes place within a period, so that there are no interest costs to be covered. Thus, intermediaries must receive expected payments equal to the investments they make:

$$\begin{aligned} E\left(P_{Y,t}^s\right) &= I_{Y,t}^s = k_{Y,t} = \phi\kappa\bar{A}_{t-1}, \\ E\left(P_{M,t}^s\right) &= I_{M,t}^s = k_{O,t} = \kappa\bar{A}_{t-1} - \widehat{RE}_{t-1}^s, \end{aligned} \tag{15}$$

where E is the expectations operator, $k_{e,t}$ is required investment in a firm of age e, manager of age e, at time t, and $\widehat{RE}_{t-1}^s$ denotes the fraction of costs financed by an old manager through retained earnings. By assumption, young firms have to run smaller projects, and will be managed by young agents who have no wealth, hence they have to borrow the full cost of the project. Mature firms may be run by old managers, who may have some retained earnings. The amount of retained earnings (invested in the firm) is decided by the manager.

In the presence of moral hazard, the payment to a manager must satisfy the following incentive constraint:

$$IC_e^s = W_e^s - \mu\pi_e^s \geq 0. \tag{16}$$

In writing the incentive compatibility constant in this way, we are ruling out long-term contracts where the payment to an experienced manager is conditioned on whether he has stolen or not in the first period of his management. We assume that there is no commitment technology for such long-term contracts, and even when a manager has stolen in the first period, if it is still profitable for the capitalist to employ him, he will do so—he cannot commit not to doing so ex post. This justifies the use of (16) as the incentive compatibility constraint.

In their contract offers, capitalists have to satisfy not only the incentive compatibility but also the individual rationality constraints of managers. These constraints ensure that managers prefer the contract that they are offered to working for the market wage, w_t. For example, for experienced managers, we need:

$$IR_{M,t}^s = W_{M,t}^s - w_t - \widehat{RE}_{t-1}^s \geq 0.$$

Both types of managers have to be paid at least the market wage. In addition, if they decide to inject any of their own retained earnings from the previous period, they have to be compensated for these retained earnings as well. Throughout, whenever there is moral hazard, i.e., unless we set $\mu = 0$, we assume that the individual rationality constraints of all managers are slack as long as their incentive compatibility constraints are satisfied.[9] This amounts to assuming that μ is large enough, so that to satisfy the incentive compatibility constraint, capitalists are already paying a sufficient amount to managers. Therefore, typically incentive compatibility constraints will bind and individual rationality constraints will be slack.

As noted above, high-skill experienced managers are always retained. Low-skill experienced managers may be retained (refinanced) or terminated. Recall that the variable $R_t \in \{0,1\}$ denotes the retention decision of mature firms for low-skill managers. We will have $R_t = 1$ when $V_M^L > E(V_Y^s)$, and capitalists will retain low-skill experienced managers, and make them an offer that satisfies their incentive compatibility constraint.

Alternatively, mature firms may prefer $R_t = 0$, i.e., to hire young managers, instead of low-skill experienced managers, and run smaller projects, which will be the case when $V_M^L < E(V_Y^s)$. In this case, the old low-skill managers are fired and become workers in the second period of their lives, and capitalists make incentive-compatible offers to some randomly-selected young workers to become managers. In addition, we also have to make sure that $E(V_Y^s) \geq 0$ so that capitalists prefer to hire managers.[10]

We can now formally define an equilibrium in this economy.

Definition 1: (Static Equilibrium) Given a_t, an equilibrium is a set of intermediate good prices, $p_t(\nu)$, that satisfy (2), a wage rate, w_t, given by (6), profit levels given by (3), and a vector of investments by intermediaries, and payments to intermediaries, managers and capitalists, $\{I_e^s, P_e^s, W_e^s, V_e^s\}$ and a continuation decision with low-skill managers, R_t, such that the feasibility equation, (14), and the free entry equation for intermediaries, (15), are satisfied, the incentive compatibility and individual rationality constraints for young and low-skill old for managers

[9] The individual rationality constraint for a young manager working in a young firm is more complicated, since such an manager might receive rents in the second period of his life. To capture this, we can write the individual rationality constraint as

$$IR_{Y,t} = E\left(W_{Y,t}^s\right) - w_t + \frac{1}{1+r} E\left(IR_{M,t+1}^s\right) \geq 0,$$

where the final term is the expectation of rents in the second period, if any.

[10] Otherwise, there would be no equilibrium innovation activity, and production is undertaken by the non-innovating fringe. Note that even if $E(V_Y^s) < 0$, innovation might be profitable when expected revenues over both periods of the firm's life are taken into account. However, we assume that capitalists cannot enter into long-term relationships with financial intermediaries, thus $E(V_Y^s) \geq 0$ is necessary.

hold, and we have $E\left(V_Y^s\right) \geq 0$, and $R_t = 1$ when $V_M^L > E\left(V_Y^s\right)$, and $R_t = 0$ when $V_M^L < E\left(V_Y^s\right)$.

Definition 2: (Dynamic Equilibrium) A dynamic equilibrium is a sequence of static equilibria such that the law of motion of the state of the economy is given by (13) above.

Next, we characterize the decentralized ("laissez-faire") equilibrium of the economy as defined in the previous section. Throughout, the emphasis will be on whether the economy pursues an investment-based strategy or an innovation-based strategy, i.e., whether $R_t = 1$ or $R_t = 0$, and how this decision varies with the state of the economy/distance to frontier, a_t.

3.2 The case of no moral hazard

We start with the case of no moral hazard ($\mu = 0$). The individual rationality constraints of both young managers and experienced managers have to hold, and they will all have lifetime income $w_t + w_{t+1}/\left(1+r\right)$, where w is the market wage given by (6).

As a result, mature firms will refinance low-skill managers, i.e., $R_t = 1$, if and only if

$$V_{M,t}^L \geq E\left(V_{Y,t}^s\right) \Leftrightarrow \pi_{M,t}^L - w_t - \kappa \bar{A}_{t-1} \geq E\left(\pi_{Y,t}^s\right) - w_t - \phi\kappa \bar{A}_{t-1},$$

where $\kappa \bar{A}_{t-1}$ is the level of investment with an old manager, while $\phi\kappa\bar{A}_{t-1}$ is the smaller level of investment when the mature firm hires an inexperienced young manager. Using (3)-(10) as well as the definition of a_t from (12), and simplifying terms, we can rewrite this inequality as:

$$\delta L \eta - \kappa > \delta\sigma L\left(\eta + \lambda\gamma a_{t-1}\right) - \phi\kappa.$$

Therefore, the equilibrium without moral hazard will have a threshold property. It will feature an investment-based strategy (refinancing of unsuccessful managers) if and only if

$$a_{t-1} < a_r\left(\mu = 0, \delta\right) \equiv \frac{\left(1-\sigma\right)\eta - \left(1-\phi\right)\kappa/\delta L}{\sigma\lambda\gamma}, \tag{17}$$

where $a_r\left(\mu = 0, \delta\right)$ is the threshold of the distance the frontier such that mature firms are indifferent between $R_t = 1$ and $R_t = 0$.

In addition, we have to check that innovation is profitable, $E\left(V_Y^s\right) \geq 0$. In the economy with no moral hazard, this requires $\delta L\sigma\left(\eta + \lambda\gamma a_{t-1}\right) - \phi\kappa - w_t \geq 0$, where the equilibrium wage is given by (6). The following is a sufficient condition for $E\left(V_Y^s\right) \geq 0$:

$$\sigma\delta L\eta \geq \phi\kappa + \left(1-\alpha\right)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}, \tag{18}$$

which we impose as an assumption. Summarizing this discussion:

Proposition 1 In the economy without moral hazard, the equilibrium has $R_t = 1$ and an investment-based strategy for all $a_{t-1} < a_r(\mu = 0, \delta)$, and $R_t = 0$ and an innovation-based strategy for all $a_{t-1} > a_r(\mu = 0, \delta)$ where $a_r(\mu = 0, \delta)$ is given by (17).

Countries farther away from the world technology frontier follow an investment-based strategy, and are characterized by long-term relationships between firms and entrepreneurs, larger and older firms, more investment and less selection. As an economy approaches the world technology frontier and passes the threshold $a_r(\mu = 0, \delta)$, it switches to an innovation-based strategy with shorter relationships, younger firms, and more selection. The intuition for this result is that the selection of managers becomes more important when the economy is closer to the world technology frontier, because there remains less room for technological improvements simply based on copying and adoption.

The threshold $a_r(\mu = 0, \delta)$ is increasing in δ, so that the switch to the innovation-based phase is delayed in less competitive economies. The reason is that refinancing existing manager corresponds to "greater investment", since the amount of investment is κ as opposed to $\phi\kappa < \kappa$ with a new manager. While the capitalist pays the full cost of the investment, because of the standard appropriability effect, part of the returns go to consumers in the form of higher real wages and consumer surplus. This discourages the strategy with a greater investment. A higher δ enables the capitalist to capture more of the surplus, encouraging refinancing. The same reasoning will apply in the economy with moral hazard in the next subsection.

Notice that this equilibrium already has a flavor of some of the issues raised by Gerschenkron. When the economy is relatively backward, there will be a very different set of (equilibrium) arrangements compared to an economy close to the technology frontier. However, Gerschenkron's emphasis was on policies that relatively backward economies ought to pursue, which is a topic we will revisit in Section 6.

Finally, we note that the equilibrium allocation characterized in Proposition 1 also applies to a range of positive μ's, for $\mu \leq \underline{\mu}$, such that the incentive compatibility constraints of both young and old low-skill managers are not binding for all $a \in [0, 1]$. See the Appendix for the characterization of $\underline{\mu}$.

3.3 The case with moral hazard

In this section, we characterize the equilibrium allocation for levels of μ sufficiently high so that the individual rationality constraints of both young and old low-skill managers are slack, and their incentive compatibility constraints are binding. Typically, this case requires μ to be greater than some threshold $\bar{\mu}$.[11]

As before, the interesting question is whether low-skill managers are refinanced. If refinanced, managers can extract rents, so they are willing to use their first-period retained earnings to cover part of the investment cost. The retained earnings of a manager who was revealed to be of low skill at the end of period $t-1$ are given by:

$$RE_{t-1} = (1+r)\,\mu\pi^L_{M,t-1} = (1+r)\,\mu\delta L\sigma\eta\bar{A}_{t-2} = \frac{1+r}{1+g}\mu\delta L\sigma\bar{A}_{t-1}, \tag{19}$$

where recall that r denotes the interest rate. The amount on the right-hand side of (19) is the total retained earnings of the manager, which may not be equal to the amount that the manager wants to, or can, inject to finance the project. Throughout the paper, we assume that

$$\frac{1+r}{1+g}\mu\delta L\sigma\eta < \kappa, \tag{20}$$

so that retained earnings are less than the cost of the project for experienced managers, and hence they can inject all of their retained earnings to finance part of the costs of the project (thus $\widehat{RE}_{t-1} = RE_{t-1}$).

Since the incentive compatibility constraint of a low-skill manager is binding, the manager receives a fraction μ of the profits. Thus, the value of a mature firm retaining a low-skill manager at date t can then be expressed as:

$$\begin{aligned} V^L_{M,t} &= (1-\mu)\,\pi^L_{M,t} - E\left(P_{M,t}\right) \\ &= \left((1-\mu)\,\delta L\eta - \kappa + \frac{1+r}{1+g}\mu\delta L\sigma\eta\right)\bar{A}_{t-1}, \end{aligned} \tag{21}$$

where the last line makes use of the fact that $E\left(P_{M,t}\right) = \kappa\bar{A}_{t-1} - \widehat{RE}_{t-1}$ as specified in equation (15), and substituting for $\widehat{RE}_{t-1} = RE_{t-1}$ from (19).

In contrast, the expected value of a mature firm employing a young manager is equal to:

$$\begin{aligned} E\left(V^s_{Y,t}\right) &= (1-\mu)\,E\left(\pi^s_{Y,t}\right) - \phi\kappa\bar{A}_{t-1} \\ &= \left((1-\mu)\,\delta L\sigma\left(\eta + \lambda\gamma a_{t-1}\right) - \phi\kappa\right)\bar{A}_{t-1}. \end{aligned} \tag{22}$$

[11] It is straightforward to check that as we consider large economies where population, L, and investment costs, κ, are large, all individual rationality constraints will be slack, while incentive constraints will bind (this is because the wage rate, w_t, does not depend on the population, L, while profits do). In other words, in this case both $\underline{\mu}$ and $\bar{\mu}$ will become arbitrarily small. Here we focus on such large economies where incentive constraints are binding and individual rationality constraints are slack.

This expression incorporates the fact that the manager will be paid a fraction μ of the profits she generates irrespective of her revealed skill.

Mature firms will continue their relationship with experienced low-skill managers, i.e., $R_t = 1$, if and only if $V_{M,t}^L \geq E\left(V_{Y,t}^s\right)$ or, equivalently, whenever $a_{t-1} \leq a_r(\mu, \delta)$ where:

$$a_r(\mu, \delta) \equiv \frac{\left((1-\mu)(1-\sigma) + \frac{1+r}{1+g}\mu\sigma\right)\eta - \kappa(1-\phi)/\delta L}{(1-\mu)\sigma\lambda\gamma}, \tag{23}$$

is the threshold for the distance to frontier such that mature firms are indifferent between $R_t = 1$ and $R_t = 0$. The sequence of the economy moving from an investment-based strategy to an innovation-based strategy is the same as in the case of no moral hazard. It is also noteworthy that the threshold $a_r(\mu, \delta)$ in (23) limits to the no-moral hazard threshold $a_r(\mu = 0, \delta)$ in (17) as incentive problems disappear, i.e., as $\mu \to 0$.

Next observe that $a_r(\mu, \delta)$ is increasing with δ as was the case with $a_r(\mu = 0, \delta)$, and when product markets are less competitive (higher δ), the switch to an innovation-based strategy occurs later. This comparative static now reflects two effects. First, for the same reasons as discussed in the previous subsection, the standard appropriability effect discourages the investment-based strategy, and an increase in δ reduces the extent of the appropriability effect. Thus, a higher δ enables the capitalists to capture more of the surplus, encouraging the investment-based strategy. Second, a higher δ implies greater profits and greater retained earnings for old unsuccessful managers, which they can use to "shield" themselves against competition from young managers, making refinancing and the investment-based strategy more likely.

The impact of the extent of incentive problems/credit market imperfections, μ, on a_r is ambiguous, however. On the one hand, a higher μ increases the earnings retained by young managers, thereby shielding these insiders from outside competition, and encouraging refinancing. On the other hand, a higher μ reduces the profit differential between hiring a young and an old low-skill manager. If

$$\delta > \frac{(1-\phi)\kappa}{\sigma\eta L}\frac{1+g}{1+r}, \tag{24}$$

then, the former effect dominates and a_r is increasing in μ, and more severe moral hazard/credit market problems encourage the investment-based strategy. In contrast, when (24) does not hold, these problems encourage the termination of low-skill managers.

To complete the characterization of equilibrium, we finally have to check that innovation is profitable, i.e., $E\left(V_Y^s\right) \geq 0$, or $((1-\mu)\delta L\sigma(\eta + \lambda\gamma a_{t-1}) - \phi\kappa) \geq 0$. This requires that

$$a_{t-1} \geq a_{ng}(\mu, \delta) \equiv \frac{1}{\lambda\gamma}\left(\frac{\phi}{\sigma(1-\mu)}\frac{\kappa}{\delta L} - \eta\right). \tag{25}$$

Assumption (18), which ensured $E(V_Y^s) \geq 0$ with no moral hazard, is not sufficient to guarantee that $a_{ng}(\mu, \delta) < 0$ for $\mu > 0$. In particular, $a_{ng}(\mu, \delta) > 0$ and stagnation traps are possible when μ is large, i.e., moral hazard is important, and δ is small, i.e., the market is highly competitive. The former effect reflects the fact that with greater moral hazard, capitalists make lower profits. The latter is due to the appropriability effect: some degree of monopoly power is necessary for capitalists to make sufficient profits from innovation.

We summarize our discussion in this subsection with the following:

Proposition 2 In the economy with moral hazard, there is a no-innovation equilibrium for all $a_{t-1} < a_{ng}(\mu, \delta)$ where $a_{ng}(\mu, \delta)$ is given by (25). For all $a_{t-1} \geq a_{ng}(\mu, \delta)$*,* the equilibrium has $R_t = 1$ and an investment-based strategy for all $a_{t-1} < a_r(\mu, \delta)$, and $R_t = 0$ and an innovation-based strategy for all $a_{t-1} > a_r(\mu, \delta)$ where $a_r(\mu, \delta)$ is given by (23). Also the threshold of distance to frontier, $a_r(\mu, \delta)$*,* is increasing in δ, so the switch to an innovation-based strategy occurs later when the economy is less competitive.

4 Growth, convergence patterns and traps

In this section, we first contrast the growth path of a laissez-faire economy with that of an economy in which refinancing decisions maximize growth. Then, we analyze the possibility of *non-convergence traps* where productivity and output per capita never converge to the world technology frontier. In the last subsection, we compare the equilibrium allocation to the welfare-maximizing allocation, which differs from the growth-maximizing one, since greater growth may come at the cost of greater investments and lower current consumption.

4.1 Growth-maximizing strategies

Consider an allocation where prices $p_t(\nu)$ satisfy (2), the wage rate, w_t, is given by (6), high-skill mature managers are refinanced, exactly as in an equilibrium allocation. However, suppose that the decision to refinance low-skill old managers, R_t, is made to maximize the growth rate of the economy. What is this growth-maximizing refinancing decision, $R_t^{\max}$? The answer is straightforward: simply compare the two branches of equation (13) corresponding to $R_t = 1$ and $R_t = 0$, and pick whichever is greater. This immediately implies that the growth-maximizing decision will be $R_t^{\max} = 1$ if:

$$a_{t-1} < \widehat{a} \equiv \frac{\eta\,(1-\sigma)}{\lambda\gamma\sigma}. \tag{26}$$

Just like the equilibrium, the growth-maximizing strategy also has a threshold property: the investment-based strategy is pursued until the economy reaches a certain distance to frontier, $\widehat{a}$, and from then on, the innovation-based strategy maximizes growth.

Since near the world technology frontier, growth, by definition, must come from innovation, not imitation, in the neighborhood of the frontier, innovation-based strategy must yield faster growth than the investment-based strategy. Thus $\widehat{a} < 1$. This reasoning implies the following parameter restriction:

Condition 1: $\lambda\gamma\sigma > \eta\,(1-\sigma)$,

which we impose throughout.[12]

If $\hat{a} < a_r$ the laissez-faire economy generates excess refinancing, or spends too much time with an investment-based strategy: there is a range of states, $a_{t-1} \in [\hat{a}, a_r]$, where all managers are refinanced in equilibrium whereas, from a growth maximization standpoint, it would be better not to refinance managers who were revealed to be of low skill. In contrast, when $\hat{a} > a_r$, the laissez-faire economy generates insufficient refinancing: there is a range of states, $a_{t-1} \in [a_r, \hat{a}]$, where only high-skill managers are refinanced in equilibrium whereas, from a growth maximization standpoint, it would be better to refinance all managers, even the low-skill ones.

Figure 1 HERE

Whether the growth-maximizing cut-off $\hat{a}$ is larger or smaller than the laissez-faire no-refinancing threshold, a_r, depends on the level of competition (δ) and on the degree of capital market imperfection (μ). First, we have that $a_r\,(\mu = 0, \delta) < \hat{a}$: the economy with no moral hazard switches to an innovation-based strategy too quickly. To see this, simply observe that

$$\hat{a} = \frac{\eta\,(1-\sigma)}{\lambda\gamma\sigma} > \frac{\eta\,(1-\sigma) - (1-\phi)\,\kappa/\delta L}{\lambda\gamma\sigma} = a_r\,(\mu = 0, \delta)\,.$$

[12] We refer to this parameter restriction and the one in the next subsection as "conditions", since they are parameter configurations implied by the logic of the model.

The same result carries over to economies with sufficiently small μ's. Intuitively, the investment-based strategy involves greater investments, and as discussed above, because of the appropriability effect, capitalists are biased against greater investments. This makes the equilibrium switch to the innovation-based strategy with smaller investments come too soon relative the growth-maximizing allocation.

While the economy without any moral hazard always switches to an innovation-based strategy too soon, i.e., $a_r(\mu = 0, \delta) < \hat{a}$, the economy with moral hazard might have $a_r(\mu > 0, \delta) > \hat{a}$. This is because moral hazard generates a high salary for managers, and they can use these as retained earnings to "shield" themselves from competition and continue in their role as managers. In other words, a key role of moral hazard in our model is to generate rents and thus protect low-skill insiders. As a result, the economy with moral hazard may remain in an investment-based strategy with little managerial selection for an excessively long duration. The possibility of staying too long in this regime will play an important role in our discussion below.

In addition, we can see that the degree of competition also affects the comparison between the equilibrium and the growth-maximizing allocations. Recall that a less competitive environment, i.e., a lower δ, encourages the investment-based strategy (cfr., equation (23)), while the growth-maximizing allocation does not depend on δ (cfr., equation (26)). Greater competition may increase or reduce the gap between the equilibrium and the growth-maximizing allocations, however, depending on whether we start from a situation where $\hat{a} > a_r(\mu, \delta)$ or $\hat{a} < a_r(\mu, \delta)$. More specifically, given μ, there exists a unique level of competition δ, denoted by $\widehat{\delta}(\mu)$, such that $\hat{a} = a_r(\mu, \delta)$, where simply comparing equations (23) and (26), we have:

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

If product market competition is lower than this threshold, namely, if $\delta > \widehat{\delta}(\mu)$ (see upper panel in Figure 1), then $\hat{a} < a_r$, and the laissez-faire economy generates excess refinancing relative to the growth-maximizing allocation.[13] In this case, greater competition increases the gap, and worsens inefficiency. Conversely, if product market competition is high, namely if $\delta < \widehat{\delta}(\mu)$ (see lower panel in Figure 1), then $\hat{a} > a_r$ and the economy switches to an innovation-based strategy too quickly, and now lower competition reduces the gap between the equilibrium and the growth-maximizing allocations.

We summarize this discussion with the following:

[13] That there is a non-empty set of parameter values where $\hat{a} < a_r$ can be seen by considering large values of for μ and δL, and comparing (23) and (26)

Proposition 3 The growth-maximizing policy is to choose $R_t = 1$ and an investment-based strategy for all $a_{t-1} < \hat{a}$, and $R_t = 0$ and an innovation-based strategy for all $a_{t-1} > \hat{a}$ where $\hat{a}$ is given by (26).

The laissez-faire economy with sufficiently small μ switches to an innovation-based strategy ($R_t = 0$) too soon relative to the growth-maximizing allocation, i.e., $a_r (\mu = 0, \delta) < \hat{a}$. An economy with sufficiently high μ and δL, on the other hand, switches to an innovation-based strategy too late.

4.2 Growth, convergence patterns, traps and leapfrogging

We now discuss how an economy that fails to switch to an innovation-based strategy might not converge to the world technology frontier. We already saw that, for combinations of large μ and small δ's, there exists a non-empty range of economies with $a < a_{ng} (\mu, \delta)$, which will stagnate because moral hazard makes innovation unprofitable for capitalists. These economies will diverge from the frontier (i.e., stagnate while the frontier is advancing). In this subsection, we will see that there is also a more interesting non-convergence trap, where certain economies grow, but fail to converge to the frontier because they persistently pursue an investment-based strategy.

As was the case for Condition 1, the logic of the model, in particular the fact that world growth must come from innovation, dictates that in the neighborhood of the world technology frontie, an economy pursuing the innovation-based strategy must grow approximately at the same rate as the frontier. This implies another condition:

Condition 2: $1 + g = [(\lambda + \sigma + (1 - \lambda) \sigma) \eta + (1 + \sigma + (1 - \lambda) \sigma) \lambda\gamma] /2.$

This condition can be thought of as endogenizing the growth rate of the world technology frontier, g, as resulting from the most advanced economy pursuing an innovation-based strategy.

Figure 2 depicts the relationships between a_t and a_{t-1} under refinancing ($R = 1$ curve) and no-refinancing ($R = 0$ curve) of old low-skill managers. The fact that both curves are linear simply follows from the expressions in equation (13). The two lines intersect at $a_{t-1} = \widehat{a}$, defined by equation (26), since by construction at this point they generate the same amount of growth (that $\widehat{a} < 1$ is guaranteed by Condition 1). When a is greater than $\hat{a}$, refinancing reduces growth, and for a less than $\hat{a}$, it increases growth, as discussed in the previous section. Furthermore, Condition 2 implies that the $R = 0$ curve hits the 45 degree line at $a = 1$. This immediately implies that the $R = 1$ curve is below the 45 degree at $a = 1$. Therefore, an economy that always pursues an investment-based strategy (i.e., refinances low-skill managers) will *not* converge to the frontier. This result is implied by the structure of our model.

Since the $R = 1$ curve starts above the 45 degree line, and ends below it at $a = 1$, it must intersect it at some

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

This is the value of the distance to frontier at which an economy pursuing an investment-based strategy will get trapped: it is a fixed point of equation (13) for $R_t = 1$.

The existence of this point, a_{trap}, does not imply that there will be traps in equilibrium, since the economy may switch to an innovation-based strategy before a_{trap}. Therefore, the necessary and sufficient condition for an equilibrium trap is $a_{trap} < a_r$. When is this condition likely to be satisfied?

>From (27), a_{trap} is an increasing function of $\lambda\gamma$, and is independent of $\kappa/\delta L$ and μ. Also, recall that a_r is a decreasing function of $\kappa/\delta L$ and of $\lambda\gamma$, and, if condition (24) holds, it is an increasing function of μ, see equation (23). These observations imply that smaller values of $\kappa/\delta L$ and $\lambda\gamma$ make it more likely that $a_{trap} < a_r$. Furthermore, if condition (24) 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 δL make it easier for low-skill managers to get refinanced. Since a trap can only arise due to excess refinancing, a greater $\kappa/\delta L$ reduces the possibility for traps. Second, large values of $\lambda\gamma$ increase the opportunity cost of refinancing low-skill managers, and make it less likely that a trap can emerge due to lack of selection. Finally, when condition (24) holds, more severe credit market imperfections (incentive problems) favor insiders by raising retained earnings, and via this channel, they increase the probability of a trap due to excess refinancing.

An implication of this discussion is that less competitive environments may foster growth at early stages of development (farther away from the technology frontier), but will later become harmful to growth, and prevent convergence to the frontier. In particular, there exists a cut-off competition level, $\delta^*(\mu)$, such that

$$a_r\left(\mu, \delta^*\left(\mu\right)\right) = a_{trap.} \tag{28}$$

An economy with a sufficiently high level of competition, $\delta < \delta^*(\mu)$, will never fall into a non-convergence trap. Too 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.[14] Low competition, instead, has detrimental effects in the long-run.

Another related implication is that when $a_r < \hat{a}$, an economy with more competitive product markets will initially grow slower than a less competitive economy, but later "leapfrog" it, when the less competitive economy becomes stuck in a non-convergence trap. This result may shed some light on why some economies, such as Brazil, Mexico or Peru, that grew relatively rapidly with highly protectionist policies were then taken over by economies with more competitive policies such as Hong Kong or Singapore.[15]

Finally, it is useful to determine the basin of attraction of a_{trap} when the economy is pursuing an investment-based strategy. Clearly, this includes all $a \in [a_{ng}, a_{trap}]$. Moreover, for $a \in (a_{trap}, 1]$, the growth rate of an economy pursuing investment-based strategy is less than g, so the basin of attraction of a_{trap} when the economy is pursuing an investment-based strategy is the entire set $a \in [a_{ng}, 1]$ (see Figure 1). Thus we have:

Proposition 4 Economies with $a < a_{ng}(\mu, \delta)$ stagnate. Economies with $a \geq a_{ng}(\mu, \delta)$ that always pursue the innovation-based strategy converge to the frontier, $a = 1$. Economies with $a \geq a_{ng}(\mu, \delta)$ that always pursue the investment-based strategy converge to $a_{trap} > 1$.

Suppose $a \in [a_{ng}, a_{trap}]$. Then if $\delta < \delta^*(\mu)$, the equilibrium involves investment-based strategy for $a < a_r(\mu, \delta) < a_{trap}$, and the innovation-based strategy for $a > a_r(\mu, \delta)$, and convergence to the frontier where $\delta^*(\mu)$ is given by (28). If $\delta > \delta^*(\mu)$, the economy always pursues the investment-based strategy and converges to a_{trap}.

4.3 Welfare analysis

In this subsection, we compare the laissez-faire equilibrium with the refinancing policy that maximizes social welfare. We will see that the economy with no moral hazard pursues an investment-based strategy (refinancing of low-skill managers) for too long, while an economy with moral hazard may have too much or too little refinancing relative to the social optimum.

More formally 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

[14] The exception is that because high competition increases a_{ng}, it makes stagnation traps more likely. Thus the statement in the text applies to economies with $a >> a_{ng}$.

[15] 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.

$C_t+\sum_{j=1}^{\infty}\beta^t C_{t+j}$, where $C_t=\zeta L A_t-\int_0^1 k_t(\nu)\,d\nu$ is equal to net output minus investment at date t with

$$\int_0^1 k_t(\nu)\,d\nu=\begin{cases}\left(\frac{1+\phi}{2}\right)\kappa\bar{A}_{t-1} & \text{if } R_t=1\\ \frac{\lambda+\phi(2-\lambda)}{2}\kappa\bar{A}_{t-1} & \text{if } R_t=0.\end{cases}$$

As before, we start with an allocation where prices $p_t(\nu)$ satisfy (2), the wage rate, w_t, is given by (6), high-skill experienced managers are refinanced, exactly as in an equilibrium allocation, but we now suppose that the decision to refinance low-skill experienced managers, R_t, is made to maximize welfare. In other words, the planner only controls the refinancing decision, R_t.

To gain some intuition, it is useful to start by characterizing the choice of a "myopic planner" who disregards future generations, i.e., $\beta=0$. The myopic planner chooses the refinancing policy at t so as to maximize total consumption at t. The myopic planner refinances low-skill managers 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.,

$$a_{mfb}\equiv\frac{\eta(1-\sigma)-(1-\phi)\kappa/\zeta L}{\sigma\lambda\gamma}. \tag{29}$$

Note here that the expression of a_{mfb} is identical to the expression of $a_r(\mu=0,\delta)$ (see equation (17)), except that here ζ replaces δ in (17). Recall that because of the appropriability effect, $\zeta>\delta$. This implies that $a_{mfb}>a_r(\mu=0,\delta)$, i.e., the planner puts more weight on the benefits of innovation than the equilibrium allocation. Therefore, the planner will choose refinancing (an investment-based strategy) over a larger range of a's. The planner's choice can also 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, we have $a_{mfb}<\hat{a}$. Thus the myopic planner sets $a_{mfb}\in(a_r(\mu=0,\delta),\hat{a})$.

Now, consider a non-myopic planner who also cares about future consumption, i.e., she has $\beta>0$. The non-myopic planner realizes that, by increasing the no-refinancing 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, $a_{fb}>a_{mfb}$, which, a fortiori, satisfies $a_{fb}>a_r(\mu=0,\delta)$, proving that the equilibrium switch to an investment-based strategy occurs too soon. 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

$$a_r(\mu = 0, \delta) < a_{mfb} < a_{fb} < \hat{a}.$$

By continuity, the same inequality holds in economies with sufficiently low μ's. The analysis therefore establishes:[16]

Proposition 5 The welfare-maximizing policy is $R_t = 1$ and an investment-based strategy for all $a_{t-1} < a_{fb}$, and $R_t = 0$ and an innovation-based strategy for all $a_{t-1} > a_{fb}$ where $a_{fb} \in (a_{mfb}, \hat{a})$ with $\hat{a}$ and a_{mfb} given by (26) and (29).

The laissez-faire economy with sufficiently small μ switches to an innovation-based strategy ($R_t = 0$) too soon relative to the welfare-maximizing allocation, i.e., $a_r(\mu = 0, \delta) < a_{fb}$. An economy with sufficiently high μ and δL, on the other hand, switches to an innovation-based strategy too late.

The last part of the proposition simply follows from the second part of Proposition 3, where we show that with sufficiently high μ and δL, we have $a_r > \hat{a}$, combined with the observation that $a_{fb} < \hat{a}$.

5 Microfoundations

The previous sections outlined a simple model to analyze the costs and benefits of investment-based and innovation-based strategies as a function of an economy's distance to the world technology frontier. The important trade-off was between exploiting the experience of existing managers vs. selecting younger more skilled managers or managers who might be better matched to new tasks. An important element was the feature that experienced managers can carry out greater investments (hence the term "investment-based strategy" for the case where experienced managers are refinanced). To simplify the analysis, this aspect was introduced as an assumption. However, there is a natural reason for why experienced managers will invest more in the presence of moral hazard: incentive compatibility problems restrict the size of investments, and retained earnings

[16] Note also that the same argument as in Proposition 5 applies if we were to compare the laissez-faire economy to the unconstrained first best, where the planner also controls pricing decisions. The unconstrained planner would set monopoly distortions to zero, so ζ would reach its highest possible value, $(1-\alpha)/\alpha$, and the planner would have a greater incentive to choose an investment-based strategy.

of experienced managers relax the credit constraints imposed by incentive problems and enable them to undertake greater investments.

In this section, we briefly outline why, once we introduce a choice of project size, we should expect experienced managers to run large projects, while young managers run small projects. This analysis is somewhat more involved than the one in the preceding sections, which motivates our choice of imposing this differential project size as an assumption for the basic model.

Consider the same economy as above with the key difference that all managers now have a choice between two project sizes: small and large. We assume that small projects cost $k_t(\nu) = \phi\kappa\bar{A}_{t-1}$, while large projects cost $k_t(\nu) = \kappa\bar{A}_{t-1}$. Next, we also allow experienced managers to be potentially more productive, though the results also go through even when they are equally productive as inexperienced managers (but in that case the investment-based strategy would not be an equilibrium in the absence of moral hazard, nor would it ever be chosen by the growth-maximizing or welfare-maximizing planners). More specifically, in equation (9), let $s_t(\nu)$, the productivity a manager-firm match, be given as:

$$s_t(\nu) = \begin{cases} \hat{\sigma} & \text{if manager=experienced and project=small} \\ 1 & \text{if manager=experienced and project=large} \\ \varepsilon\hat{\sigma} & \text{if manager=inexperienced and project=small} \\ \varepsilon & \text{if manager=inexperienced and project=large} \end{cases}$$

where $\hat{\sigma} < 1$ and $\varepsilon \leq 1$, i.e., $\varepsilon = 1$, the case in which experienced and inexperienced managers have the same productivity is allowed. This setup therefore implies that both project size and managerial experience (firm-specific experience) contribute to the productivity of the project (in other words, to advances in productivity). Moreover, we assume that

$$\delta L\eta - \kappa > \hat{\sigma}\delta L\eta - \phi\kappa,$$

so, absent moral hazard problems, even with low-skill experienced managers it would be profitable to run larger projects. Without this assumption, low-skill experienced managers would never run large projects. The rest of the setup is unchanged, and for brevity, here we focus on the case with moral hazard.

To determine whether young managers run small or large projects, we need to compare capitalists' returns from these two options (individual rationality constraints of young managers are slack with both small and large projects). Comparing these returns, we find that capitalists will choose the small project if:

$$(1-\mu)\varepsilon\delta L(\eta + \lambda\gamma a_{t-1}) - \kappa < (1-\mu)\varepsilon\hat{\sigma}\delta L(\eta + \lambda\gamma a_{t-1}) - \phi\kappa. \tag{30}$$

The left-hand side is profits from the large project, multiplied by $1-\mu$ (since a fraction μ of the profits have to be paid to the manager for incentive compatibility), minus the cost of the project. Notice that profits are multiplied by ε, which captures the fact that young managers are potentially less productive. The right-hand side is for a small project, and profits are multiplied by $\varepsilon\hat{\sigma}$, since now the manager is inexperienced and is running a small project. Moreover, the project cost is multiplied by ϕ, because the project is small. Notice that (30) can only be true if $\phi < \hat{\sigma}$, that is if there are decreasing returns to scale so that the costs of the project increase faster than the returns from the larger scale. Thus, we are implicitly assuming that there are decreasing returns to scale in project size.

Since the left-hand side of (30) increases in a_{t-1} less than the right-hand side, then if (30) holds for $a_{t-1} = 1$, it will hold for all a_{t-1}. To simplify the discussion, we now impose the following condition

$$(1-\mu)\,\varepsilon\delta L\,(\eta+\lambda\gamma) - \kappa < (1-\mu)\,\varepsilon\hat{\sigma}\delta L\,(\eta+\lambda\gamma) - \phi\kappa, \tag{31}$$

which ensures that (30) holds for $a_{t-1} = 1$, and , hence, for all a_{t-1}. As a result, when (31) holds, capitalists will always choose the small project with young managers. This condition is satisfied if μ is sufficiently large (as long as $\phi < \hat{\sigma}$).

Now, let us look at the choice of project size with experienced managers. To simplify the discussion, we will introduce an additional assumption, similar to (20) above, but somewhat stronger:

$$\phi\kappa < \frac{1+r}{1+g}\hat{\sigma}\mu\delta L\varepsilon\eta < \kappa. \tag{32}$$

Similar to our previous assumption (20), this condition states that retained earnings are less than the cost of the large project. But it also requires that they are greater than the cost of the small project.

Next, we must compare the return to capitalists from large and small projects with an experienced low-skill manager. Capitalists prefer the large project with experienced managers when

$$(1-\mu)\,\delta L\eta - \kappa + \frac{1+r}{1+g}\hat{\sigma}\mu\delta L\varepsilon\eta > (1-\mu)\,\hat{\sigma}\delta L\eta. \tag{33}$$

The left-hand side is capitalists' profits from a large project with a low-skill experienced manager, $\delta L\eta$, times $1-\mu$ (since this fraction has to be paid to the manager for incentive compatibility), minus the difference between the cost of the project and the retained earnings that the manager contributes, i.e., $\kappa - \frac{1+r}{1+g}\hat{\sigma}\mu\delta L\varepsilon\eta$. The right-hand side is simply profits from the small project times $1-\mu$. The cost of the small project is financed entirely by the manager from retained earnings.

When assumptions (31), (32) and (33) are satisfied, young managers run small projects while experienced managers run large projects.[17] Intuitively, capitalists are discouraged from financing large projects because the form of the incentive compatibility problem (that the manager can appropriate a fraction μ of the returns) implies greater payments to managers from larger projects. Incentive problems are therefore introducing "credit constraints," and these constraints are especially severe on young managers, reducing their level of investment. However, for experienced managers there is a countervailing effect: these managers have retained earnings, and they can use these retained earnings to bear part of the cost of the project, making larger projects more attractive for capitalists. In other words, retained earnings relax the "credit constraints" introduced by the incentive problems.

Notice that moral hazard together with choice of project size (and the assumptions (31), (32), and (33)) implies exactly the configuration we imposed in the previous sections: young managers run smaller projects, while experienced managers run larger projects and, everything else equal, they generate more revenues. This indicates that the results here will be very similar to those in the previous sections. In fact, they are identical.

To see this, define $\sigma \equiv \varepsilon\hat{\sigma}$, and the comparison between financing a young manager and a low-skill experience manager boils down to the comparison of (21) and (22), and yields the critical threshold of (23) as in Section 2.

Proposition 6 In the economy with project size choice and moral hazard, as long as conditions (31), (32) and (33) are satisfied, the equilibrium is identical to that in Proposition 2 with $\sigma \equiv \varepsilon\hat{\sigma}$.

Therefore, this extended model gives exactly the same results as our basic model in

[17] It is straightforward to verify that these three conditions can be easily satisfied. For example, take $\frac{1+r}{1+g}\hat{\sigma}\mu\delta L\varepsilon\eta \to \kappa$, which ensures both (32) and (33), and then reduce ϕ to satisfy (31).

In addition, we have to make sure that it is in the interest of the entrepreneur to inject his retained earnings. Since the capitalist is making the offer, this simply requires that individual rationality constraint of the entrepreneur is satisfied, or in other words:

$$\mu\delta L\eta - \frac{1+r}{1+g}\hat{\sigma}\mu\delta L\varepsilon\eta > w_t = (1-\alpha)\,\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}a_t.$$

where the left-hand side is the net return from running the large project, taking into account the costs, and the right-hand side is the return from quitting and working for the market wage. As usual, we assume that this individual rationality constraint is satisfied, which again requires μ to be large enough.

Section 2. In the rest of the paper, we work with the simpler model of Section 2.

6 Policy, appropriate institutions and political economy traps

The analysis so far established that:

1. The equilibrium allocation, the growth-maximizing allocation and the social optimum all involve an investment-based regime with high investment and long-term relationships, followed by an innovation-based regime with lower investment, shorter relationships, younger firms and more selection.

2. Unless incentive/credit market problems are sufficiently severe and the economy is highly non-competitive, the equilibrium switch to an innovation-based strategy will happen too soon. In other words, economies farther away from the frontier might have a tendency to "invest too little" and grow "too slowly".

These observations raise the possibility of useful policy interventions along the lines suggested by Gerschenkron: relatively backward economies intervening to increase investment in order to ensure faster adoption of technologies and development. In this section, we discuss possible policies to foster growth, how they can be interpreted as corresponding to "appropriate institutions" for countries at different stages of development, and how political economy considerations might turn appropriate institutions into "inappropriate institutions" that generate traps.

6.1 Policy and appropriate institutions

Consider an equilibrium allocation with $a_r(\mu, \delta) < a_{fb}$. A policy intervention that encourages greater investment will increase welfare and growth over a certain range. There are a number of different policies that 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 subsidizing a fraction s of the cost of investment. If s is chosen appropriately, the economy can be induced to switch from an investment-based strategy to an innovation-based strategy exactly at a_{fb} or (at $\hat{a}$, depending on the purpose of policy). Alternatively, the government could use investment subsidies only for existing firms, with similar results. An additional role of investment subsidies is that they would reduce a_{ng}, the stagnation threshold, thus making stagnation less likely.

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, in much of the discussion we focus on another potential policy instrument which affects the equilibrium threshold $a_r(\mu, \delta)$, the extent of anti-competitive policies, such as entry barriers, merger policies etc.. Nevertheless, all of the discussion applies to investment subsidies for all or only existing firms.

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 $a_r(\mu, \delta) < a_{fb}$, policies that restrict competition will close the gap between the equilibrium threshold and the social optimum (or the growth-maximizing point). Although restricting competition creates static losses, in the absence of feasible tax/subsidy policies this may be the best option available for encouraging faster growth and technological convergence. Similar to investment subsidies, a higher δ (or a higher χ) also reduces $a_{ng}(\mu, \delta)$ and the range of stagnation.

The situation where the government chooses a less competitive institutional environment in a relatively backward economy in order to encourage more investment, long-term relationships and faster technological convergence is reminiscent to Gerschenkron's analysis. Appropriate institutions for relatively backward economies may then be thought to correspond to those that create a less competitive and perhaps "more rigid" environment, and encourage longer-term relationships and greater investment. This is also, in some sense, similar to the famous "infant-industry" arguments that call for protection and government support for industries at early stages of development.

But our analysis also reveals that such institutions/policies limiting competition, and similarly investment subsidies, are harmful for economies closer to the world technology frontier. Appropriate institutions for early stages of development are *inappropriate* for an economy close to the world technology frontier. Therefore, any economy that adopts such institutions must then abandon them at some point; otherwise, it will end up in a non-convergence trap.

A sequence of optimal 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 the change in policy. To the extent that economic power buys political power, they will 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 certain groups oppose the change in policy, and the economy ends up in a non-convergence trap because, at early stages of development, it adopted appropriate institutions.

We now build a simple political economy model where special interest groups, depending on the economic power, may capture politicians. Our basic political-economy model is a simplified version of the special-interest-group model of Grossman and Helpman (1997, 2001) combined with our growth setup.

6.2 Political economy traps

Suppose that competition policy (the "institutional" environment), χ , is determined in each period by a politician (or government) who cares about the current welfare of living agents, but is also sensitive to bribes—or 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 (i.e., the planner does not have a long horizon), 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, and the timing structure simplifies the algebra below.

In this formulation, the parameter H may 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. We will refer to H as the honesty parameter of politicians. When H is greater, the political system is less corruptible. This formulation is similar to that in Grossman and Helpman (1997, 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. We also assume that only capitalists can organize as interest groups, so the only group with the capability to bribe politicians are mature capitalists.[18]

To simplify the analysis further, we assume that the institutional choice facing the politician is between two policies, low and high competition, or between "competitive" and "anti-competitive" policies, i.e., $\chi_t \in \left\{\underline{\chi}, \bar{\chi}\right\}$ where $\alpha^{-1} \geq \bar{\chi} > \underline{\chi}$. We set, by analogy, $\delta_t = (\chi_t - 1)\,\chi_t^{-\frac{1}{1-\alpha}} \in \{\underline{\delta}, \overline{\delta}\}$ and $\zeta_t = (\chi_t - \alpha)\,\chi_t^{-\frac{1}{1-\alpha}}/\alpha \in \{\underline{\zeta}, \overline{\zeta}\}$. 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.).

We start our analysis by characterizing the policy that would be chosen by an *honest politician* who will never be influenced by bribes (i.e., $H = \infty$). First, note that the honest politician will always choose competitive policies ($\chi = \underline{\chi}, \delta = \underline{\delta}$) for $a_{t-1} \leq a_r\,(\mu, \underline{\delta})$, since over this range, even with relatively high competition there is refinancing, i.e., $R_t = 1$, so anti-competitive policies would simply create static distortions without affecting equilibrium refinancing decisions.

Will an honest politician choose anti-competitive policies for any $a > a_r\,(\mu, \underline{\delta})$? In this range, anti-competitive policies may create a trade-off: they lead to monopoly price distortions, but they may encourage the investment-based strategy, which for $a < a_{mfb}$, yields greater current consumption than the innovation-based strategy, where a_{mfb} is defined by (29). It is straightforward to verify that the honest politician will choose anti-competitive policies if and only if $a_{t-1} \leq a_{wm}$, where a_{wm} is such that:[19]

$$a_{wm} \equiv \frac{\left(\bar{\zeta}\,(1+\sigma) - \underline{\zeta}\,(\lambda + \sigma\,(2-\lambda))\right)\eta - (1-\phi)\,(1-\lambda)\,\kappa/L}{\lambda\gamma\left(\underline{\zeta}\,(1+\sigma\,(2-\lambda)) - \bar{\zeta}\,(1+\sigma)\right)}. \tag{34}$$

Thus a_{wm} is the threshold of the distance to frontier such that low competition and $R_t = 1$ give the same level of current consumption as greater competition and $R_t = 0$. Honest politicians will prefer low competition when $a \in [a_r\,(\mu, \underline{\delta})\,, a_{wm}]$, when this set is nonempty. The reason why anti-competitive policies cease to be desirable when $a > a_{wm}$ is that the benefits from these policies decline as the economy gets closer to the frontier.

Next consider the competition policy set by a politician who responds to bribes (i.e.,

[18] The qualitative results would not change if we allowed mature managers to contribute to the anti-competitive lobby.

[19] a_{wm} is derived by equating consumption under (i) refinancing and low competition, $\bar{\zeta}$, and (ii) no refinancing and high competition, $\underline{\zeta}$. Formally, we set:

$$\bar{\zeta}L\left(\frac{1+\sigma}{2}\,(\eta + \lambda\gamma a_{wm})\right) - \frac{1+\phi}{2}\kappa = \underline{\zeta}L\frac{1}{2}\,((\lambda + \sigma + (1-\lambda)\,\sigma)\,\eta + (1 + \sigma + (1-\lambda)\,\sigma)\,\lambda\gamma a_{wm}) - \frac{\lambda + \phi\,(2-\lambda)}{2}\kappa,$$

Simplifying this expression gives (34).

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, $\chi_t = \overline{\chi}$, rather than competitive policies, $\chi_t = \underline{\chi} < \overline{\chi}$.[20]

We assume that agents cannot borrow to pay bribes, so the amount of bribes that they can pay will be also limited by their current income. This assumption introduces the link between economic power and political power in our context: *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 $\delta_{t-1} \in [\underline{\delta}, \bar{\delta}]$ was 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:

$$B^C(\delta_{t-1}, a_{t-1}) = \delta_{t-1}(1-\mu)\sigma(L\eta + \lambda\gamma a_{t-1}) - \phi\kappa. \tag{35}$$

The maximum bribes capitalists will pay are therefore:

$$B(\delta_{t-1}, a_{t-1}) = \min\left\langle B^W(a_{t-1}), B^C(\delta_{t-1}, a_{t-1})\right\rangle.$$

We focus on economies where capitalists are credit constrained in the range of interest. Sufficiently small values of σ guarantee that this is the case. Thus, from now on, $B(\delta_{t-1}, a_{t-1}) = B^C(\delta_{t-1}, a_{t-1})$. 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, $\overline{\chi}$, for welfare-maximizing reasons, she will be induced to change the policy to $\overline{\chi}$ if and only if bribes are sufficient to cover the dishonesty cost, HA_{t-1}, or if and only if:

$$B^C(\delta_{t-1}, a_{t-1}) \geq Ha_{t-1}.$$

Using (35), we can rewrite this inequality as

$$\delta_{t-1}(1-\mu)\sigma L(\eta + \lambda\gamma a_{t-1}) - \phi\kappa \geq Ha_{t-1}. \tag{36}$$

[20] Let $R(\delta, a) \in \{0, 1\}$ denote the refinancing decision conditional on the policy δ and distance from frontier a. Then, the maximum bribe that capitalists are willing to pay is given by

$$\begin{aligned} B^W(a) \;=\; & \left(\bar{\delta}\left(1 - R\left(\bar{\delta}, a\right)\right) - \underline{\delta}\left(1 - R\left(\underline{\delta}, a\right)\right)\right)(1-\mu)L\left((\lambda + (1-\lambda)\sigma)\eta + (1 + (1-\lambda)\sigma)\lambda\gamma a\right) + \\ & \left(\bar{\delta}R\left(\bar{\delta}, a\right) - \underline{\delta}R\left(\underline{\delta}, a\right)\right)(1-\mu)L(\eta + \lambda\gamma a) - \left(R\left(\bar{\delta}, a\right) - R\left(\underline{\delta}, a\right)\right)(1 - (\lambda + (1-\lambda)\phi))\kappa. \end{aligned}$$

It can be shown that $B^W(a)/a$ is a continuously decreasing function of a.

We define a_L and a_H as the unique values of a_{t-1} such that (36) holds with equality for, respectively, $\delta_{t-1} = \bar{\delta}$ and $\delta_{t-1} = \underline{\delta}$. We thus have:

$$a_L \equiv \frac{\bar{\delta}\left(1-\mu\right)\sigma L\eta - \phi\kappa}{H - \lambda\gamma\bar{\delta}\left(1-\mu\right)\sigma L} \text{ and } a_H \equiv \frac{\underline{\delta}\left(1-\mu\right)\sigma L\eta - \phi\kappa}{H - \lambda\gamma\underline{\delta}\left(1-\mu\right)\sigma L}.$$

The politicians will be bribed to maintain the anti-competitive policy, $\overline{\chi}$, as long as $a_{t-1} \leq a_L$. Similarly, they will be bribed to switch from competitive to the anti-competitive policies when $a_{t-1} \leq a_H$.

It is immediate to see that $a_L > a_H$, since 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 are more likely to secure the policy that they prefer. Note that both cut-offs, a_L and a_H, are decreasing functions of H, which captures the fact that more honest politicians will be harder to convince to pursue the policy preferred by capitalists.

Figure 3 HERE

Now consider Figure 3.[21] For $a \leq a_H$, the politician is successfully bribed and anti-competitive policies prevail. If $a \geq a_L$, there is no bribe, and the politician chooses the welfare-maximizing policy. Finally, if $a \in (a_H, a_L)$, 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. If competition is initially high, capitalists do not make as much profits and do have not enough funds to buy politicians, so there is no effective lobbying activity, and equilibrium policies are competitive.

Next consider the evolution of an economy with initial level of technology a_0 satisfying $a_0 < a_H$. Irrespective of past competition policies, the capitalist lobby is wealthy enough to buy the anti-competitive policy $\bar{\chi}\ (\bar{\delta})$. In earlier stages of development $(a < \hat{a})$, the only effect of the lobbying activity is a static distortion that reduces consumption, but it has no effect on innovation and growth. In some intermediate stage of development where $a \in (a_r(\mu, \underline{\delta}), \hat{a})$, the anti-competitive policy actually becomes growth-enhancing. When $a \in \left(\hat{a}, a_r\left(\mu, \bar{\delta}\right)\right)$, however, the industrial policy resulting from lobbying activities

[21] Figure 3 uses a parameter configuration such that $a_H < a_r(\mu, \underline{\delta}) < a_{wm} < a_L$, while Figure 4 assumes $a_H < a_r(\mu, \underline{\delta}) < a_{trap} < a_L$. It is straightforward construct numerical examples with plausible parameters where $a_H < a_r(\mu, \underline{\delta}) < a_{wm} < a_{trap} < a_L$, thus satisfying both parameter configurations. Details are available upon request.

is harmful for growth, as well as reducing the level of consumption through the static distortion. Growth slowdowns and the economy may even get stuck into a non-convergence trap.

We refer to this case as a "political economy trap", since the reason why the non-convergence trap emerges is the ability of the capitalist lobby to bribe politicians. Consider an economy starting at some level $a_0 < a_H$, such that the capitalistic lobby can initially bribe politicians. Two conditions are sufficient for a political economy trap to arise:

1.
$$\underline{\delta} < \delta^*(\mu) < \overline{\delta}, \tag{37}$$
where, recall, $\delta^*(\mu)$ was defined as the cut-off competition level such that $a_r(\mu, \delta^*(\mu)) = a_{trap}$ defined in (28). Under this assumption, the anti-competitive policy, $\overline{\chi}$, leads to a non-convergence trap, where low-skill managers are always refinanced and the economy pursues an investment-based strategy. The competitive policy, $\underline{\chi}$, would have instead ensured convergence to the world technology frontier.

2.
$$a_{trap} < a_L, \tag{38}$$
This condition implies that when the economy reaches a_{trap}, and convergence comes to a halt, the anti-competitive lobby continues to prevent the change of policy that would be necessary to induce further convergence.

These two conditions, (37) and (38), are more likely to be satisfied when H is low, i.e., when the political system is more corruptible. Therefore, political economy traps are more likely in societies with weak political institutions, and such institutions might have to be more careful in pursuing government interventions.

Figure 4 HERE

Figure 4 describes how the trap arises diagrammatically. The policy choice is endogenous, and the lobbying activity implies low competition for all $a \leq a_L$. If the economy ever reached a state $a = a_L$, it would switch to high competition and an innovation-based strategy, and would eventually attain full convergence to the world technology frontier. But this stage is never reached since convergence stops at $a = a_{trap} < a_L$.

Another, possibly more interesting, case is when the economy starts with $a_0 \in (a_H, a_r(\mu, \underline{\delta}))$ and $\chi = \underline{\chi}$, i.e., competition is high. In this case, capitalists do not initially have enough funds to bribe politicians to reduce competition. Thus, without a change in competition policy, the capitalist lobby will never be effective. However, as long as $a_{wm} > a_r(\mu, \underline{\delta})$, when a increases above $a_r(\mu, \underline{\delta})$, politicians will choose to reduce competition *in order* to create welfare gains for the citizens. Once competition is reduced, capitalists become richer, and now they have enough funds to successfully bribe politicians to keep competition low. This case, therefore, illustrates how a well-meaning (but shortsighted) attempt to introduce appropriate institutions may lead to a political economy trap.

Finally, note also that when $a_H < a_{trap}$, a temporary improvement in policy might have long-run policy and economic benefits. In particular, if the adverse effects of lobbying activity could be prevented for even just one period (e.g., by the election of an exceptionally honest politician), the economy could escape from the trap. The honest politician would choose competitive policies, and this would destroy the ability of capitalists to lobby against competition in the future. So, even a temporary improvement in "political institutions" would lead to permanent changes in "economic institutions" (here the degree of competition in the product market).[22]

This discussion establishes:

Proposition 7 Suppose that competition policy is decided by a sequence of politicians with honesty parameter H, and bribes by the lobby of capitalists. Then, there exists a cut-off level a_L, which is decreasing in H, such that the politician will always be bribed into maintaining a low level of competition if $a < a_L$.

When parameters are such that (i) $\delta^*(\mu) < \overline{\delta}$ where $\delta^*(\mu)$ is defined by (28); and (ii) $a_{trap} < a_L$, then an economy starting at $a_0 < \max \langle a_r(\mu, \underline{\delta}), a_L \rangle$ will be locked-in into a non-convergence trap, characterized by the anti-competitive policy $\overline{\delta}$ $(\overline{\chi})$, and bribes to politicians from the capitalist lobby. Such political economy traps are more likely in economies where H is small.

[22]Note that this extreme result hinges on the two-period nature of our model. If agents live for more periods, and the capitalists own other assets, other reforms may be necessary to curb the power of insiders. Redistribution and reduction of income or wealth inequality may be necessary to make such a reform sustainable. Nevertheless, clearly the feature that current policies affect profits and therefore the capitalist lobby's ability to influence policy in the future is more general than the 2-period model here

7 Conclusion

There are certain marked differences in the economic organization of technological leaders and technological followers. While technological leaders often feature younger firms and greater churning, technological followers emphasize investment and long-term relationships. In other words, while technological leaders follow an innovation-based strategy, technological followers adopt an investment-based strategy of growth.

In this paper, we have proposed a model which accounts for this pattern, and also evaluates the pros and cons of investment-based and innovation-based strategies. In our economy, managers engage both in copying and adopting technologies from the world frontier and in innovation activities. The selection of high-skill managers is more important for innovation activities. As the economy approaches the technology frontier, selection becomes more important. As a result, countries that are far away from the technology frontier pursue an investment-based strategy, with long-term relationships, high average size and age of firms, large average 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 managers.

The sequence of investment-based strategy followed by an innovation-based strategy is not only a feature of the equilibrium, but also of the socially-planned economy. However, societies 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, makes the switch too soon. Whereas the presence of retained earnings that incumbent managers 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 existing firms may be useful because they encourage 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 managerial selection. This means that policies encouraging investment-based strategies might also lead to non-convergence traps.

The optimal policy sequence for economic growth 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 maintain these policies. Moreover, these groups, in our model the capitalists, will be politically powerful precisely because they have economically benefited from the less-competitive policies in place. Therefore, the model illustrates how a well-meaning attempt to speed up convergence may lead to a political economy trap. Interestingly, such traps are more likely when the underlying political institutions are weak, making politicians easier to capture. 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 is a 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, to impose effective constraints on politicians and elites, such government intervention may lead to the capture of politicians by groups that benefit from government 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] Abernathy, William (1980) *The Productivity Dilemma,* Baltimore John Hopkins

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

[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] Acemoglu, Daron and Fabrizio Zilibotti (1999) "Information Accumulation in Development" *Journal of Economic Growth*, vol. 4, pp. 5-38.

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

[6] Arrow, Kenneth (1974) *The Limits of Organization*, New York, Norton.

[7] Bahk B. H. and Michael Gort (1993) "Decomposing learning by doing in new plants" *Journal of Political Economy* Vol 101, pp 561-583.

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

[9] Benkard, Lenier (2000) "Learning and Forgetting: the Dynamics of Aircraft Production" *American Economic Review*, volume 90,1034-1054.

[10] Bond, Stephen and Meghir, Costas (1994) "Dynamic investment models and the firm's financial policy", *Review of Economic Studies*, 61, 197-222.

[11] 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.

[12] Carlin, Wendy and Colin Mayer, (2002), "Finance, Investment and Growth", Mimeo University College London.

[13] 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.

[14] Dunne, Timothy, Mark Roberts and Larry Samuelson (1989) “The Growth and Failure of U.S. Manufacturing Plants” *Quarterly Journal of Economics*, volume 104,671-698.

[15] Evans, Peter, (1998), *Embedded Autonomy: States and Industrial Transformation*, Princeton University Press.

[16] Francois, Patrick and Joanne Roberts (2001), “Contracting Productivity Growth” CentER Working Paper No.35, Tilburg University.

[17] Freeman, Christopher (1982) *The economics of industrial innovation*, Cambridge, MA, MIT Press.

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

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

[20] 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

[21] Grossman, Gene and Elhanan Helpman, (1999), “Competing for Endorsements”, *American Economic Review*, vol.89, No. 3, pp.501-524

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

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

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

[25] Hirsch, W. Z. (1952) “Manufacturing progress functions” *Review of Economics and Statistics* Vol 34, pp 143-155.

[26] Hoshi, Takeo and Anil Kashyap (2001) *Corporate financing and governance in Japan : the road to the future*, Cambridge, Massachusetts, MIT Press.

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

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

[29] Johnson, Chalmers, (1982), *MITI and the Japanese Miracle: The Growth of Industrial Policy, 1925-75*, Stanford University Press.

[30] Irwin, Douglas A. (2002) "Did Import Substitution Promote Growth in the Late Nineteenth Century?" National Bureau of Economic Research, Working Paper No. 8751.

[31] Irwin, Douglas A. and Peter Klenow (1994) "Learning-by-Doing Spillovers in the Semiconductor Industry" *Journal of Political Economy*, volume 102, 1200-1227.

[32] King, Robert G. and Ross Levine, (1993), *Quarterly Journal of Economics*, vol. 108, No. 3. pp. 717-737.

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

[34] La Porta, Rafael, Florencio López-de-Silanes, Andrei Shleifer and Robert Vishny (1998) "Law and Finance", *Journal of Political Economy,* vol.106, No. 6, pp. 1113-55.

[35] Lieberman, M. B. (1984) "The learning curve and pricing curve in chemical processing industries" *Rand Journal of Economics* Vol 15, pp 213-28

[36] Martimort, David and Thierry Verdier (2001) "Monopolies life cycle, bureaucratization and Schumpeterian growth ", London : Centre for Economic Policy Research.

[37] Nelson, Richard R. and Sidney G. Winter (1982) *An Evolutionary Theory of Economic Change*, Cambridge, MIT Press.

[38] Nunn, Nathan and Daniel Trefler, (2002) "The Political Economy of Tariffs and Growth"University of Toronto, mimeo.

[39] Ohashi, Hiroshi (2002) "Learning-by-Doing, Export Subsidies and Japanese Steel Miracle" University of British Columbia mimeo.

[40] O'Rourke, Kevin (2000) "Tariffs and Growth in Late 19th-century" *Economic Journal*, volume 110, 456-483.

[41] Rajan, Raghuram G. and Luigi Zingales (1998) "Financial Dependence and Growth', *American Economic Review,* vol. 88, pp. 559-586.

[42] Rosenberg, Nathan (1982) *Inside the Black Box: Technology and Economics*, Cambridge, Cambridge University Press.

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

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

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

[46] Vogel, Ezra F. (1991) *The Four Little Dragons: the Spread of Industrialization in East Asia*, Cambridge, MA, Harvard University press.

[47] Wade, Robert (1990) *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization*, Princeton, Princeton University Press.

7.1 Appendix A: determination of $\underline{\mu}$

Since $a_r(\mu = 0, \delta) < \hat{a}$, the incentive compatibility of the old unsuccessful starts binding before the incentive compatibility of the young employed in old firms. The condition for the incentive compatibility of the old unsuccessful to bind is:

$$\mu \delta L \eta \bar{A}_{t-1} > (1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}} A_t = w_t \tag{39}$$

where the wage, w_t, depends, through A_t, on $R_t \in \{0, 1\}$.

If $R_t = 1$, then (39) reads (after dividing both terms of the inequality by $\bar{A}_{t-1}$):

$$\mu \delta L \eta > (1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}} \frac{1+\sigma}{2}(\eta + \lambda\gamma a_{t-1})$$

If $R_t = 0$, then (39) reads

$$\mu \delta L \eta > (1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}} \frac{1}{2}\left((\lambda + \sigma + (1-\lambda)\sigma)\eta + (1 + \sigma + (1-\lambda)\sigma)\lambda\gamma a_{t-1}\right) \tag{40}$$

These two equations define, respectively, the following thresholds

$$\begin{aligned} a_{R=1} &= \frac{\mu\delta L - (1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}\frac{1+\sigma}{2}}{\lambda\gamma(1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}\frac{1+\sigma}{2}}\eta \\ a_{R=0} &= \frac{\mu\delta L - (1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}\frac{(\lambda+\sigma+(1-\lambda)\sigma)}{2}}{\lambda\gamma(1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}\frac{(1+\sigma+(1-\lambda)\sigma)}{2}}\eta \end{aligned}$$

The incentive compatibility constraints do not bind as long as $a_r(\mu = 0, \delta) > \max\{a_{R=1}, a_{R=0}\}$, in which case the allocation is as in the case where $\mu = 0$. Since $a_r(\mu = 0, \delta) < \hat{a}$, then $a_{R=1} > a_{R=0}$. The threshold $\underline{\mu}$ is therefore given by equating $a_r(\mu = 0, \delta) = a_{R=1}$. This yields

$$\frac{\underline{\mu}\delta L - (1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}\frac{1+\sigma}{2}}{\lambda\gamma(1-\alpha)\alpha^{-1}\chi^{-\frac{\alpha}{1-\alpha}}\frac{1+\sigma}{2}}\eta = \frac{(1-\sigma)\eta - (1-\phi)\frac{\kappa}{\delta L}}{\sigma\lambda\gamma}$$

and, solving

$$\underline{\mu} = (1-\alpha)\alpha^{-1}\chi^{\frac{-\alpha}{1+\alpha}}(1+\sigma)\left(\frac{\delta L - \frac{\kappa}{\eta}(1-\phi)}{2\sigma}\right)$$

Note that, when $\mu < \underline{\mu}$, the incentive compatibility condition of the experienced managers may bind at some low levels of a_{t-1} such that $a_{t-1} < a_r(\mu = 0, \delta)$. But this has no effect on the refinancing decision nor on the determination of the threshold $a_r(\mu = 0, \delta)$, since when the economy approaches the relevant threshold all incentive compatibility conditions are slack.

7.2 Appendix B: Details with more general costs

In this appendix, we briefly generalize the results to the case when investment costs take the form $k_t(\nu) = \kappa \bar{A}_{t-1}^{\rho} A_{t-1}^{1-\rho}$ for experienced managers, and $k_t(\nu) = \kappa \phi \bar{A}_{t-1}^{\rho} A_{t-1}^{1-\rho}$ for all others. It is clear that this modification does not affect the growth maximizing threshold, $\hat{a}$. The comparison in the case of no moral hazard changes simply to:

$$\delta L \eta - \kappa a_{t-1}^{1-\rho} \text{ vs. } \delta \sigma L (\eta + \lambda \gamma a_{t-1}) - \phi \kappa a_{t-1}^{1-\rho}.$$

Equating these two terms, we obtain a new equation defining $a_r(\mu = 0, \delta)$:

$$\delta L \eta - \kappa a_r(\mu = 0, \delta)^{1-\rho} = \delta \sigma L (\eta + \lambda \gamma a_r(\mu = 0, \delta)) - \phi \kappa a_r(\mu = 0, \delta)^{1-\rho}.$$

Although this equation has no closed-form solution, $a_r(\mu = 0, \delta)$ has exactly the same comparative statics as (17) in the text. It is also straightforward to verify that $a_r(\mu = 0, \delta) < \hat{a}$.

The analysis of the case with moral hazard is similar, and defines an equilibrium cutoff parallel to $a_r(\mu, \delta)$ given by (23). Condition (20), which ensures that retained earnings are not sufficient to pay to full cost of the project, now is somewhat more complicated, and requires

$$\frac{1+r}{1+g} \mu \delta L \sigma \eta < \kappa (a_r(\mu, \delta))^{1-\rho}.$$

Given this assumption, $a_r(\mu, \delta)$ is a solution to the equation:

$$(1-\mu) \delta L \eta - \kappa (a_r(\mu, \delta))^{1-\rho} - \frac{1+r}{1+g} \mu \delta L \sigma \eta = (1-\mu) \delta L \sigma (\eta + \lambda \gamma a_r(\mu, \delta)) - \phi \kappa (a_r(\mu, \delta))^{1-\rho},$$

and has the same comparative statics as $a_r(\mu, \delta)$ in (23). In particular, it is monotonically increasing in δ, and may be increasing or decreasing in μ. As in the text, because of the rents created by moral hazard, $a_r(\mu, \delta)$ can be larger than $\hat{a}$. The results about non-convergence traps immediately generalize to this case

FIGURE 1

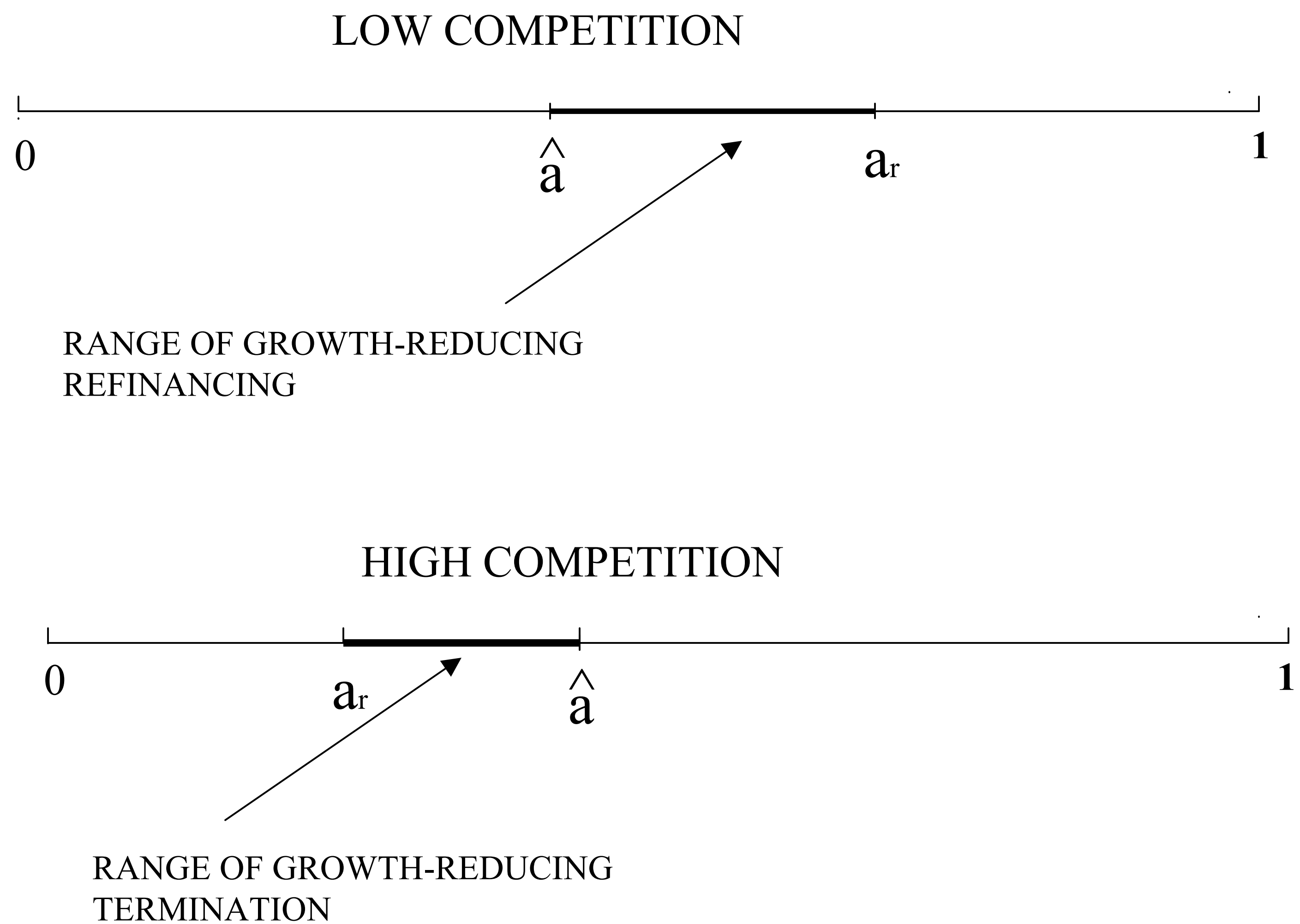
LOW COMPETITION
0
$\hat{a}$
a_r
1
RANGE OF GROWTH-REDUCING REFINANCING
HIGH COMPETITION
0
a_r
$\hat{a}$
1
RANGE OF GROWTH-REDUCING TERMINATION

FIGURE 2

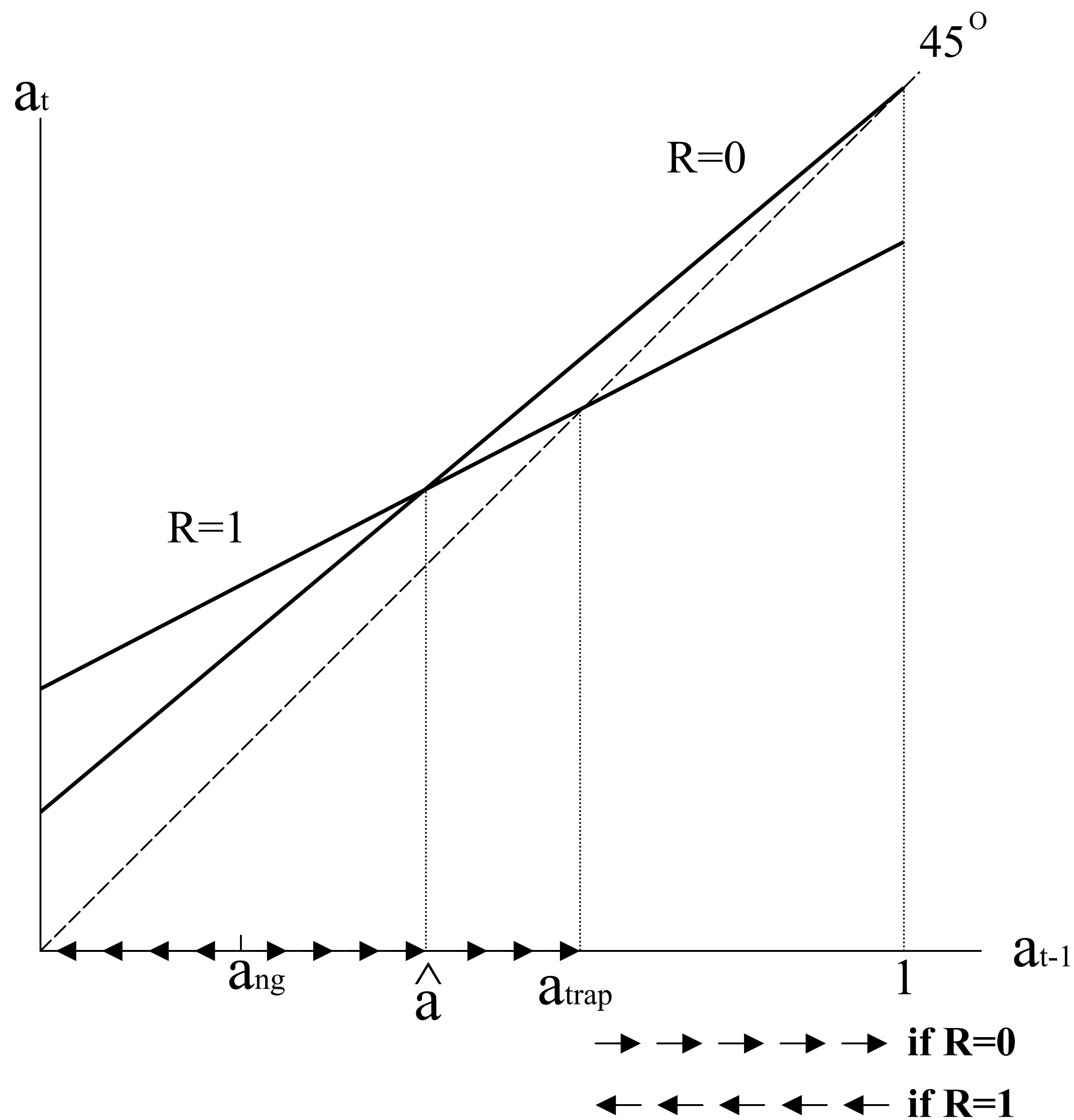

FIGURE 3

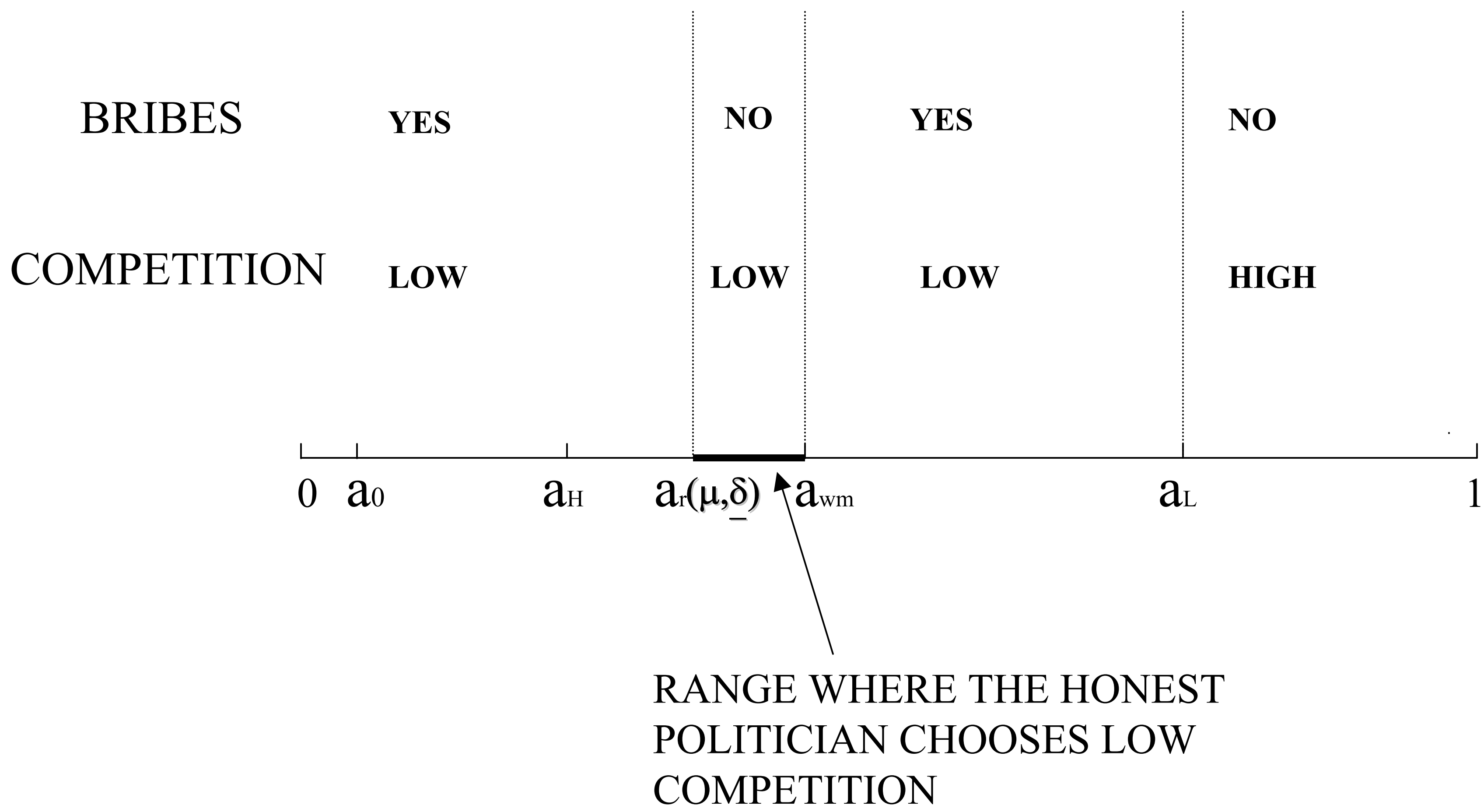
BRIBES
YES
NO
YES
NO
COMPETITION
LOW
LOW
LOW
HIGH
0
a_0
a_H
$a_r(\mu,\underline{\delta})$
a_{wm}
a_L
1
RANGE WHERE THE HONEST POLITICIAN CHOOSES LOW COMPETITION

FIGURE 4