Asymmetric Growth and Institutions in an Interdependent World*

Daron Acemoglu
MIT

James A. Robinson
Harvard

Thierry Verdier
Paris School of Economics-Ecole des Ponts Paris-Tech

April 2014.

Abstract

We argue that in a technologically interconnected world, the world equilibrium may be asymmetric, involving different economic institutions and technology levels for different countries. In our model, all countries benefit and potentially contribute to advances in the world technology frontier. A greater gap of incomes between successful and unsuccessful entrepreneurs (thus greater inequality) increases entrepreneurial effort and hence a country's contribution to the world technology frontier. We show that, under plausible assumptions, the world equilibrium is asymmetric: some countries will opt for a type of "cutthroat" capitalism that generates greater inequality and more innovation and will become the technology leaders, while others will free-ride on the cutthroat incentives of the leaders and choose a more "cuddly" form of capitalism. Paradoxically, those with cuddly reward structures, though poorer, may have higher welfare than cutthroat capitalists; but in the world equilibrium, it is not a best response for the cutthroat capitalists to switch to a more cuddly form of capitalism. We also show that domestic constraints from social democratic parties or unions may be beneficial for a country because they prevent cutthroat capitalism domestically, instead inducing other countries to play this role.

JEL Classification: O40, O43, O33, P10, P16.

Keywords: cutthroat capitalism, economic growth, inequality, innovation, interdependences, technological change.

*We thank Pascual Restrepo for superb research assistance. We thank the editor Harald Uhlig and four anonymous referees for useful comments and suggestions. We also benefited from the suggestions of seminar participants at Brown University, Copenhagen University, El Rosario University in Bogotá and 2011 Latin American and Caribbean Economic Association meetings. We thank James Conran for bringing Hayek's quote to our attention. Acemoglu and Robinson gratefully acknowledge support from the Canadian Institute for Advanced Research.

"Not only are the countries of the West richer because they have more advanced technological knowledge, but they have more advanced technological knowledge because they are richer. And the free gift of the knowledge that has cost those in the lead much to achieve enables those who follow to reach the same level at a much smaller cost. Indeed, so long as some countries lead, all the others can follow, although the conditions for spontaneous progress may be absent in them. That even countries or groups which do not possess freedom can profit from many of its fruits is one of the reasons why the importance of freedom is not better understood. For many parts of the world the advance of civilization has long been a derived affair, and, with modern communications, such countries need not lag very far behind, though most of the innovations may originate elsewhere. How long has Soviet Russia or Japan been living on an attempt to imitate American technology! So long as somebody else provides most of the new knowledge and does most of the experimenting, it may be possible to apply all this knowledge deliberately in such a manner as to benefit most of the members of a given group at about the same time and to the same degree. But, though an egalitarian society could advance in this sense, its progress would be essentially parasitical, borrowed from those who have paid the cost." Friedrich von Hayek, *The Constitution of Liberty*, p. 42, Routledge Classics Edition, 2006.

1 Introduction

The costs and benefits of the American (or more broadly Anglo-Saxon) economic system compared to its European counterpart are much debated. To its proponents, US institutions, which tolerate or even encourage greater economic inequality,[1] are at the root of its innovative economy, technological leadership and high level of per capita income.[2] To its critics, the inequality and the economic uncertainty they create, coupled with its more limited social safety net, more than offset any efficiency gains that US economic system may create (and many also doubt that there are such major efficiency gains).[3] Implicit in much of this debate is the notion that countries can switch to whichever type of economic system is superior—provided that this is known and other political economy constraints are overcome.

[1] For example, the top 1% of earners in the United States account for about 23.5% of national income, while the same number is 10% in Sweden, 11.1% in Spain and 12.7% in Germany, and excluding capital gains, 9.25% in France, 8.3% in Finland, 6.1% in Denmark, 8.5% in Norway, and 9.9% in Italy (Atkinson, Piketty and Saez, 2011).

[2] Though there is a debate on whether the United States is more innovative than Scandinavian and continental European nations, and whether it has been losing its technology leadership, our reading of the evidence is that the US technology leadership is fairly clear. In an earlier version of our paper (Acemoglu, Robinson and Verdier, 2012), we provided evidence showing that the United States generates more patents and more importantly, a higher fraction of US patents are more highly cited than Scandinavian ones, and this gap increases when we focus on the very top of the citation distribution. Acemoglu, Akcigit and Celik (2014) construct several indices of creative innovations, and with all of them the United States is ahead of Scandinavia and continental Europe.

[3] In addition to the lower social welfare because of greater inequality and economic uncertainty, Americans also work more hours, which could also reduce social welfare. Schor (1993) was among the first to point out the comparatively much greater hours that American workers work. Blanchard (2004) has more recently argued that Americans may be working more than Europeans because they value leisure less.

A more sophisticated version of this view is developed in the literature on "varieties of capitalism," pioneered by Hall and Soskice (2001). These authors argue that a successful capitalist economy need not give up on social insurance to achieve rapid growth. They draw a distinction between a Coordinated Market Economy (CME) and a Liberal Market Economy (LME), and suggest that both can have high incomes and similar growth rates, but CMEs have more social insurance and less inequality. Different societies develop these different models for historical reasons, and once set up institutional complementarities make it very difficult to switch from one model to another (and piecemeal reforms may backfire). Nevertheless, a LME could turn itself into a CME by simultaneously reforming many interwoven economic institutions, and if it succeeded in doing so, it would lose little in terms of income and growth, and gain significantly in terms of welfare.

In this paper, we suggest that in an interconnected world, such a switch may have much more far-reaching consequences because there are reasons for the world equilibrium to be *asymmetric*. It may be precisely the more "cutthroat" American economic institutions that make it possible for more "cuddly" economic institutions to emerge in other parts of the world, for example in Scandinavia or other parts of continental Europe.

The basic idea we propose is simple and echoes Hayek's quote we started with. The main building block of our model is technological interdependence across countries: innovations, particularly by the most technologically advanced countries, contribute to the world technology frontier, and other countries can build on this frontier.[4] We combine this with the idea that innovations require incentives for workers and entrepreneurs. The well-known incentive-insurance trade-off (e.g., Holmstrom, 1979) implies that a society strongly encouraging innovation will have greater inequality (and a weaker safety net). Crucially, however, in a world with technological interdependences, when one society (or a small subset) is at the technological frontier and is contributing disproportionately to its advancement, the incentives for others to do so will be weaker. This is because innovation incentives of economies at the world technology frontier will create higher growth by advancing the frontier, while strong innovation incentives by followers will only create a "level effect" since the world technology frontier is already being pushed forward by advanced economies. This logic underlies the asymmetric nature of world equilibrium. Because innovation is associated with more high-powered incentives, the technology leader(s) will have to sacrifice insurance and equality. The followers, on the other hand, can best respond to the technology leader's advancement of the world technology frontier by ensuring better insurance to their population—e.g., in the form of a better safety net, a welfare state and greater equality. Notably, the followers, though technologically less advanced and poorer in terms of income per capita, may achieve *higher* welfare because of the better risk-sharing and insurance they provide to their citizens.

[4] Such knowledge spillovers are consistent with broad patterns in the data and are often incorporated into models of world equilibrium growth. See, Coe and Helpman (1995) and Keller (2001), Botazzi and Peri (2003), and Griffith, Redding and Van Reenen (2005) for some of the cross-industry evidence, and see, among others, Nelson and Phelps (1966), Howitt (2000), and Acemoglu, Aghion and Zilibotti (2006) for models incorporating international spillovers.

The bulk of our paper formalizes these ideas using a simple (canonical) model of world equilibrium with technology transfer. Our model is a version of Romer's (1990) endogenous technological change model with multiple countries (as in Acemoglu, 2009, Chapter 18). R&D investments within each economy advance that economy's technology, but these build on the world technology frontier. Incorporating Gerschenkron's (1962) famous insight, countries that are further behind the world technology frontier have an "advantage of backwardness" in that there is more unused knowledge at the frontier for them to build upon (see also Nelson and Phelps, 1966). We depart from this framework only in one dimension: by assuming plausibly that there is a moral hazard problem for workers (entrepreneurs). For successful innovation, agents need to be given incentives, which will be at the cost of consumption insurance. A fully forward-looking (country-level) social planner chooses a *reward structure*, which corresponds to levels of consumption for successful and unsuccessful outcomes for entrepreneurs, and shapes innovation incentives.

To start with, we focus on the case in which the world technology frontier is advanced only by the technology leader. Then, under some natural assumptions, the leader adopts a "cutthroat" reward structure, with high-powered incentives for success, while other countries free-ride on this leader's innovations and choose a more egalitarian, "cuddly," reward structure, at least once they reach a certain level of income and technology.[5] In the long-run, all countries grow at the same rate, but those with cuddly reward structures are poorer. As already noted above, these follower countries may have higher welfare than the cutthroat leader, however.

Our model therefore challenges the idea implicit in much of the debate on US vs. European (or Scandinavian) capitalism. Under the assumptions of our model, which we view as a fairly natural approximation to reality, we *cannot* all be like the Scandinavians (or like continental Europeans) because it is not an equilibrium for the cutthroat leader, "the United States," to also adopt such a reward structure. This is because if, given the strategies of other countries, the cutthroat leader did so, this would reduce the growth rate of the entire world economy, discouraging the adoption of the more egalitarian reward structure. In contrast, followers are still happy to choose more egalitarian reward structures because this choice, though making them poorer, does not permanently reduce their growth rates, which are determined by the growth rate of the world technology frontier.

This result makes it clear that the egalitarian reward structures in the follower countries are made possible by the technological externalities created by the cutthroat technology leader. So interpreting the empirical patterns in light of our theoretical framework, one may claim (with all the usual caveats of course) that, for example, the more harmonious and egalitarian Scandinavian societies are made possible because they are able to benefit from and free-ride on the knowledge externalities created by the cutthroat American institutions.

[5] In barticular, somewhat reminiscent of the path pursued by countries such as South Korea or Taiwan and consistent with a version of the modernization theory (e.g., Lipset, 1959), our model implies that some follower countries may first adopt cutthroat reward structures for rapid convergence, and then start building a safety net and more egalitarian institutions once they have approached the world technology frontier sufficiently and reached a certain level of income.

Our results extend straightforwardly to the case in which the world technology frontier is affected by all countries, provided that the function aggregating the innovation decisions of all countries into the world technology frontier is sufficiently convex. This requirement is natural since without convexity of this aggregator, innovations by less advanced countries would be as important for world technological progress as those by more advanced countries, removing the economic forces that underpin an asymmetric world equilibrium.

Finally, we consider an extension in which we introduce domestic politics as a constraint on the behavior of the social planner. We do this in a simple reduced-form, assuming that in some countries there is a strong social democratic party (or labor movement) ruling out reward structures that are very unequal. We show that if two countries start at the same initial technology level, the social democratic party in country 1 may prevent cutthroat capitalism in that country, inducing a unique equilibrium in which country 2 is the one adopting the cutthroat reward structure. In this case, however, this is a significant advantage, because if the two countries start at the same level, the cutthroat country always has lower welfare. Therefore, a social democratic party, by constraining the actions of the social planner, can act as a commitment to egalitarianism, inducing an equilibrium in which the country in question becomes the beneficiary from the asymmetric world equilibrium. This result has the flavor of the domestic political conflicts in one country being "exported" to another, as the strength of the social democratic party in country 1 makes the poor in country 2 suffer more—as country 2 in response adopts a more cutthroat reward structure.

The role of several simplifying assumptions in our analysis should be recognized at the outset (and will be discussed more later). First and most importantly, linking technological change to the financial rewards to successful innovation is a simplification. Many important innovations are produced without high-powered incentives, and there are reasons why innovation may be encouraged by greater equality (for example, because of better risk sharing or because compressed wage structures encourage technology adoption, e.g., Acemoglu, 2003). These considerations notwithstanding, private innovation naturally responds to profit incentives (a feature that is the bedrock of the canonical endogenous technological change model, which we utilize in this paper). Second, in contrast to our model, Scandinavian and continental European countries are clearly not ex ante identical to the United States, and may have chosen more redistributive policies not only—not mainly—as a result of the trade-off between innovation incentives and social insurance, but because of their political history or because of greater taste for redistribution or concerns of fairness among their voters. Naturally, this does not invalidate our analysis, and such differences can be readily incorporated into the preferences of the social planner without major changes in the formal analysis (but, of course, the resulting equilibrium will then be even more likely to be asymmetric). More interestingly, our analysis in Section 6 shows that in a global economy there will be a natural complementarity between this type of preference for redistribution and equilibrium reward structures. For example, even a weak preference for redistribution might serve as a selection device, in the same way that a strong social democratic

party does in Section 6, ensuring that countries with greater preference for redistribution end up as institutional and technological followers, potentially with positive effects on their citizens' welfare. Finally, as we discuss further in Section 4.8, interdependence across countries may not be purely technological. Most interestingly, when countries trade, those with cutthroat incentives may specialize in different sectors than those operating under cuddly reward structures, providing an additional channel for asymmetric equilibria in an interconnected world.

The rest of the paper proceeds as follows. In the next section, we discuss the related literature. Section 3 introduces the economic environment. Section 4 presents the main results of the paper. Section 5 presents two important generalizations of these results. Section 6 shows how domestic political economy constraints can be advantageous for a country because they prevent it from adopting a cutthroat reward structure. Section 7 provide historical and case study evidence illustrating the main mechanism proposed in this paper, focusing primarily on innovation and technological spillovers in the pharmaceutical industry. Section 8 concludes, and the proofs of the main results are provided in the Appendix.

2 Related Literature

Our paper is related to several different literatures. First, as already noted, the issues we discuss are at the core of the "varieties of capitalism" literature in political science, e.g., Hall and Soskice (2001) which itself builds on earlier intellectual traditions offering taxonomies of different types of capitalism (Cusack, 2009) or welfare states (Esping-Anderson, 1990). A related argument is developed in Aoki (2001), who also notes the possibility that international interconnections might be at the root of this type of institutional diversity (see also Rodrik, 2008). As mentioned above, Hall and Soskice (2001) argue that while both CME and LMEs are innovative, they innovate in different ways and in different sectors. LMEs are good at "radical innovation" characteristic of particular sectors, like software development, biotechnology and semiconductors, while CMEs are good at "incremental innovation" in sectors such as machine tools, consumer durables and specialized transport equipment (see Taylor, 2004, and Akkermans, Castaldi, and Los, 2009, for assessments of the empirical evidence on these issues). This literature has not considered that growth in an CME might critically depends on innovation in the LMEs and on how the institutions of CMEs are influenced by this dependence. Most importantly, to the best of our knowledge, the point that the world equilibrium may be asymmetric, and different types of capitalism are chosen as best responses to each other, is new and does not feature in this literature.

Second, there is a related literature in economics, focusing on the causes of institutional differences across developed economies, and on why the US lacks a European style welfare state and why Europeans work less. Acemoglu and Pischke (1998) suggest an explanation for differences in labor market institutions between the US and Germany based on multiple equilibria in turnover, information and training investments. Landier (2005) develops a similar model to account for differences in entrepreneurial risk-taking between the US and France. Bénabou and

Tirole (2006) develop a model in which self-fulfilling beliefs about justice and fairness can lead to divergent redistributive policies across countries. Bénabou's (2000) model is also related as it generates multiple equilibria, one with high inequality and low redistribution and another with low inequality and high redistribution, because redistribution can contribute to growth in the presence of capital market imperfections (see also Saint-Paul and Verdier, 1993, and Moene and Wallerstein, 1997). Bénabou's (2006) focus is also closely related to ours, particularly when he asks "what joint configuration of technology, inequality and policy are feasible in the long run? ... How does the [international] diffusion of technology affect nations ability to maintain their own redistributive institutions and social structures?".[6] None of this work, however, contains the core idea of this paper that the institutions of one country interact with those of another and that even with identical fundamentals asymmetric equilibria are the norm not an exception to explain.

Third, the idea that institutional differences may emerge endogenously depending on the distance to the world technology frontier has been emphasized in past work, for example, in Acemoglu, Aghion and Zilibotti (2006) (see also Krueger and Kumar, 2004). Nevertheless, this paper and others in this literature obtain this result from the domestic costs and benefits of different types of institutions (e.g., more or less competition in the product market), and the idea that activities leading to innovation are more important close to the world technology frontier is imposed as an assumption. In our model, this latter feature is endogenized in a world equilibrium, and the different institutions emerge as best responses to each other. Put differently, the distinguishing feature of our model is that the different institutions emerge as an asymmetric equilibrium of the world economy—while a symmetric equilibrium does not exist.

Finally, our results also have the flavor of "symmetry breaking" as in several papers with endogenous location of economic activity (e.g., Krugman and Venables, 1996, Matsuyama, 2002, 2005, 2013) or with endogenous credit market frictions (Matsuyama, 2004, 2007). These papers share with ours the result that similar or identical countries may end up with different choices and welfare levels in equilibrium, but the underlying mechanism and the focus are very different.[7]

[6] Another branch of the literature has emphasized the role of differences in political systems. These include: electoral systems with proportional representation, characteristic of continental, Europe may lead to greater redistribution (Alesina, Glaeser and Sacerdote, 2001, Milesi-Ferretti, Perotti and Rostagno, 2002, Persson and Tabellini, 2003); or the federal nature of the US may lower redistribution (Cameron, 1978, Alesina, Glaeser and Sacerdote, 2001); or the greater ethnic heterogeneity of the US may reduce the demand for redistribution (Alesina, Glaeser and Sacerdote, 2001); or greater social mobility in the US may mute the desire for redistributive taxation (Bénabou and Ok, 2001, Alesina and La Ferrara, 2005), or redistribution may be greater in Northern Europe because of higher levels of social capital and trust (Algan, Cahuc and Sangnier, 2011).

[7] There is also a connection between our work and the literature on "dependency theory" in sociology, developed, among others, by Cardoso and Faletto (1979) and Wallerstein (1974-2011). (We thank Leopoldo Fergusson for pointing out this connection.) This theory argues that economic development in "core" economies, such as Western European and American ones, takes place at the expense of underdevelopment in the "periphery," and that these two patterns are self-reinforcing. In this theory, countries such as the United States that grow faster are the winners from this asymmetric equilibrium. In our theory, there is also an asymmetric outcome, though the mechanisms are very different, and indeed ours is more a model of "reverse dependency theory," since it is the Scandinavian or European "periphery" which, via free-riding, is benefiting from this asymmetric equilibrium.

3 Model

In this section, we describe the economic environment, which combines two components: the first is a standard model of endogenous technological change with knowledge spillovers across J countries—and in fact closely follows Chapter 18 of Acemoglu (2009). The second introduces moral hazard on the part of entrepreneurs, thus linking entrepreneurial innovative activity of an economy to its reward structure. We then introduce "country social planners" who choose the reward structures within their country in order to maximize discounted welfare.

3.1 Economic Environment

Consider an infinite-horizon economy consisting of J countries, indexed by $j = 1, 2, ..., J$. Each country is inhabited by non-overlapping generations of agents who live for a period of length Δt, work, produce, consume and then die. A continuum of agents, with measure normalized to 1, is alive at any point in time in each country, and each generation is replaced by the next generation of the same size. We will consider the limit economy in which $\Delta t \to 0$, represented as a continuous time model.

The aggregate production function at time t in country j is

$$Y_j(t) = \frac{1}{1-\beta}\left(\int_0^{N_j(t)} x_j(\nu,t)^{1-\beta} d\nu\right) L_j^{\beta}, \tag{1}$$

where L_j is labor input, $N_j(t)$ denotes the number of machine varieties (or blueprints for machine varieties) available to country j at time t. In our model, $N_j(t)$ will be the key state variable and will represent the "technological know-how" of country j at time t. We assume that technology diffuses slowly and endogenously across countries as will be specified below. Finally, $x_j(\nu, t)$ is the total amount of machine variety ν used in country j at time t. To simplify the analysis, we suppose that x depreciates fully after use, so that the x's are not additional state variables. Crucially, blueprints for producing these machines, captured by $N_j(t)$, live on, and the increase in the range of these blueprints will be the source of economic growth.

Each machine variety in economy j is owned by an "entrepreneur," who sells machines embodying this technology at the profit-maximizing (rental) price $p_j^x(\nu, t)$ within the country (there is no international trade). This monopolist entrepreneur can produce each unit of the machine at a marginal cost of ψ in terms of the final good. Without any loss of generality, we normalize $\psi \equiv 1 - \beta$.

Suppose that each entrepreneur in this economy exerts effort $e_{j,i}(t) \in \{0, 1\}$ to invent a new machine. Effort $e_{j,i}(t) = 1$ costs $\gamma > 0$ units of time, while $e_{j,i}(t) = 0$ has no time cost. Thus, entrepreneurs who exert effort consume less leisure. We also assume that entrepreneurial success is risky. When the entrepreneur exerts effort $e_{j,i}(t) = 1$, he is "successful" with probability q_1 and unsuccessful with the complementary probability. If he exerts effort $e_{j,i}(t) = 0$, he is successful with the lower probability $q_0 < q_1$. Throughout we assume that effort choices are private information.

We assume the utility function of entrepreneur/worker i takes the form

$$U\left(C_{j,i}\left(t\right),e_{j,i}\left(t\right)\right)=\frac{\left[C_{j,i}\left(t\right)\left(1-\gamma e_{j,i}\left(t\right)\right)\right]^{1-\theta}}{1-\theta}, \tag{2}$$

where $\theta \geq 0$ (and $\theta \neq 1$) is the coefficient of relative risk aversion. This form of the utility function ensures balanced growth.[8]

We assume that entrepreneurs can simultaneously work as workers (so that there is no occupational choice). This implies that each individual receives wage income as well as income from entrepreneurship, and also implies that $L_j = 1$ for $j = 1, \ldots, J$.

An unsuccessful entrepreneur does not generate any new ideas (blueprints), while a successful entrepreneur in country j generates

$$\eta N\left(t\right)^{\phi} N_j\left(t\right)^{1-\phi}$$

new ideas for machines, where $N(t)$ is an index of the world technology frontier, to be endogenized below, and $\eta > 0$ and $\phi > 0$ are assumed to be common across the J countries. This form of the innovation possibilities frontier implies that the technological know-how of country j advances as a result of the R&D and other technology-related investments of entrepreneurs in the country, but the effectiveness of these efforts also depends on how advanced the world technology frontier is relative to this country's technological know-how. When it is more advanced, an innovation will lead to more rapid progress, and the parameter ϕ measures the extent of this.

Given the likelihood of success by entrepreneurs as a function of their effort choices and defining $e_j(t) = \int e_{j,i}(t)\, di$, technological advance in this country can be written as:

$$\dot{N}_j\left(t\right)=\left(q_1 e_j\left(t\right)+q_0(1-e_j\left(t\right))\right)\eta N\left(t\right)^{\phi} N_j\left(t\right)^{1-\phi}. \tag{3}$$

We also assume that monopoly rights over the initial set of ideas are randomly allocated (independent of effort) to some of the current entrepreneurs, so that they are also produced monopolistically.[9]

The world technology frontier is assumed to be given by

$$N\left(t\right)=G(N_1\left(t\right),...,N_J\left(t\right)), \tag{4}$$

where G is a linearly homogeneous function. We will examine two special cases of this function. The first is

$$G(N_1\left(t\right),...,N_J\left(t\right))=\max\left\{N_1\left(t\right),...,N_J\left(t\right)\right\}. \tag{5}$$

which implies that the world technology frontier is given by the technology level of the most advanced country, the technology leader, and all other countries benefit from the advances of

[8] When $\theta = 1$, we set the utility function (2) to $\ln C_{j,i}(t) + \ln(1 - \gamma e_{j,i}(t))$. (We do not include the -1 sometimes included in the numerator of this class of utility functions in order to simplify notation). All of our results apply to this log case also, but in what follows we often do not treat this case separately to save space.

[9] The alternative is to assume that existing machines are produced competitively. This has no impact on any of the results in the paper, and would just change the value of B in (10) below.

this technological leader. The second is a more general convex aggregator

$$G(N_1(t), ..., N_J(t)) = \frac{1}{J^{\frac{\sigma}{\sigma-1}}} \left[\sum_{j=1}^{J} N_j(t)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{6}$$

with $\sigma < 0$. The term $1/J^{\frac{\sigma}{\sigma-1}}$ ensures that the convex aggregator is homogeneous of degree 1 in the number of countries. As $\sigma \uparrow 0$ (6) converges to (5). For much of the analysis, we focus on the simpler specification (5), though at the end of the next section we show that our general results are robust when we use (6) with σ sufficiently close to 0.

3.2 Reward Structures

Entrepreneurial effort levels will depend on the *reward structure* in each country, which determines the relative rewards to successful entrepreneurship. In particular, suppressing the reference to country j to simplify notation for this subsection, let $\tilde{R}^s(t)$ denote the time t entrepreneurial income for successful entrepreneurs and $\tilde{R}^u(t)$ for unsuccessful entrepreneurs. Thus the total income of a worker/entrepreneur is

$$R^i(t) = \tilde{R}^i(t) + w(t),$$

where $i \in \{s, u\}$, and $w(t)$ is the equilibrium wage at time t.[10] In what follows, it is total income, R^i, rather than just the entrepreneurial component of income, $\tilde{R}_i$, that matters for effort decisions. The reward structure can then be summarized by the ratio $r(t) \equiv R^s(t)/R^u(t)$. When $r(t) = 1$, there is perfect consumption insurance at time t, but this generates effort $e = 0$. Instead, to encourage $e = 1$, $r(t)$ needs to be above a certain threshold, which we characterize in the next section.

This description makes it clear that each country will have a choice between two styles of capitalism: "cutthroat capitalism" in which $r(t)$ is chosen above a certain threshold, so that entrepreneurial success is rewarded while failure is at least partly punished, and "cuddly capitalism" in which $r(t) = 1$, so that there is perfect equality and consumption insurance, but this comes at the expense of lower entrepreneurial effort and innovation.

Throughout we assume that the sequence of reward structures in country j, $[r_j(t)]_{t=0}^{\infty}$ is chosen by its country-level social planner. This assumption enables us to construct a simple game between countries (in particular, it enables us to abstract from within-country political economy issues until later). Limiting the social planner to only choose the sequence of reward structures is for simplicity and without any consequence.[11]

[10] Thus both $\tilde{R}_u(t)$ and $\tilde{R}_s(t)$ include the rents that entrepreneurs make in expectation because of existing ideas being randomly allocated to them.

[11] It is straightforward to see that, instead of directly choosing the reward structure, the social planner can also achieve the same allocation by using taxes and subsidies conditional on success.

If, in addition, we allow the social planner to set prices that remove the monopoly markup, this would only change the value of the constant B in (10) below, with no impact on any of our results.

We assume that each country social planner maximizes discounted welfare of the citizens in that country using the following preferences

$$\int_0^\infty e^{-\rho t} \left[\int \left(\frac{\left[C_{j,i}(t)\left(1-\gamma e_{j,i}(t)\right)\right]^{1-\theta}}{1-\theta} \right) di \right] dt, \tag{7}$$

where ρ is the discount rate that the social planner applies to future generations, $\left[C_{j,i}(t)\left(1-\gamma e_{j,i}(t)\right)\right]^{1-\theta}/(1-\theta)$ denotes the utility of agent i in country j alive at time t (and thus the inner integral averages across all individuals of that generation). Thus, in this formulation, θ, in addition to being the coefficient of relative risk aversion of agents, captures the aversion of the social planner to inequality both within a cohort and between cohorts. Note also that (7) gives a well-defined preferences both when $\theta < 1$ and when $\theta > 1$.

We also consider the following Epstein-Zin-style preferences (Epstein and Zin, 1989), which separate the coefficient of relative risk aversion from the parameter determining cross-cohort comparisons. For this purpose, define

$$\mathcal{W}(t) = \left[\int \left(\left[C_{j,i}(t)\left(1-\gamma e_{j,i}(t)\right)\right]^{1-\theta} \right) di \right]^{\frac{1}{1-\theta}}.$$

Then these more general preferences are given by

$$\mathcal{W} = \frac{1}{1-\lambda} \int_0^\infty e^{-\rho t} \mathcal{W}(t)^{1-\lambda} dt. \tag{8}$$

Here θ is the coefficient of relative risk aversion, and $\lambda \neq 1$ measures the social planner's aversion to inequality between cohorts (and is thus similar to the inverse of the intertemporal elasticity of substitution). Clearly, (8) contains (7) as a special case (setting $\theta = \lambda$), and is also well-defined for all values of θ and λ except 1.

4 Main Results

In this section, we present our main results focusing on the "max" specification of the world technology frontier given by (5) and the standard preferences for country social planners given by (7). We present generalizations of our results to the cases where preferences are given by (8) and the world technology frontier is as in (6) in Section 5.

4.1 Cutthroat and Cuddly Reward Structures

We now define the *cutthroat* and *cuddly* reward structures more formally. Consider the reward structures that ensure effort $e = 1$ at time t. This will require that the incentive compatibility constraint for entrepreneurs be satisfied at t, or in other words, expected utility from exerting effort $e = 1$ should be greater than expected utility from $e = 0$. Using (2), this requires

$$\frac{1}{1-\theta}\left(q_1 R^s(t)^{1-\theta} + (1-q_1) R^u(t)^{1-\theta}\right)(1-\gamma)^{1-\theta} \geq \frac{1}{1-\theta}\left(q_0 R^s(t)^{1-\theta} + (1-q_0) R^u(t)^{1-\theta}\right),$$

where recall that $R^s(t)$ is the income and thus the consumption of an entrepreneur/worker conditional on successful innovation, and $R^u(t)$ is the income level when unsuccessful, and this expression takes into account that high effort leads to success with probability q_1 and low effort with probability q_0, but with high effort the total amount of leisure is only $1-\gamma$. Rearranging this expression, we obtain

$$\begin{aligned} r(t) \equiv \frac{R^s(t)}{R^u(t)} &\geq \left(\frac{(1-q_0)-(1-q_1)(1-\gamma)^{1-\theta}}{q_1(1-\gamma)^{1-\theta}-q_0} \right)^{\frac{1}{1-\theta}} \\ &= \left(1+\frac{1-(1-\gamma)^{1-\theta}}{q_1(1-\gamma)^{1-\theta}-q_0} \right)^{\frac{1}{1-\theta}} \equiv A. \end{aligned} \tag{9}$$

Clearly, the expression A defined in (9) measures how "high-powered" the reward structure needs to be in order to induce effort, and will thus play an important role in what follows.

The next assumption, which will be maintained throughout our analysis, ensures that, both when $\theta < 1$ and when $\theta > 1$, high effort requires entrepreneurs to be given incentives. In particular, it implies that $A > 1$.[12]

Assumption 1:

$$\min \left\{ q_1(1-\gamma)^{1-\theta}-q_0 \ , \ (1-q_0)-(1-q_1)(1-\gamma)^{1-\theta} \right\} > 0.$$

Since the (country) social planner maximizes average utility, she would like to achieve as much consumption insurance as possible subject to the incentive compatibility constraint (9), which implies that she will satisfy this constraint as equality. In addition, $R^s(t)$ and $R^u(t)$ must satisfy the resource constraint at time t. Using the expression for total output and expenditure on machines provided in the Appendix, this implies

$$q_1 R^s(t) + (1-q_1)R^u(t) = BN_j(t)$$

where

$$B \equiv \frac{\beta(2-\beta)}{1-\beta}, \tag{10}$$

and we are using the fact that in this case, all entrepreneurs will exert high effort, so a fraction q_1 of them will be successful. Combining this expression with (9), we obtain

$$R^s(t) = \frac{BA}{q_1 A + (1-q_1)} N_j(t) \text{ and } R^u(t) = \frac{B}{q_1 A + (1-q_1)} N_j(t). \tag{11}$$

The alternative to a reward structure that encourages effort is one that forgoes effort and provides full consumption insurance—i.e., the same level of income to all entrepreneur/workers of $R^o(t)$, regardless of whether they are successful or not. In this case, the same resource constraint implies

$$R^o(t) = BN_j(t). \tag{12}$$

[12] In particular, when $\theta < 1$, $1 + \frac{1-(1-\gamma)^{1-\theta}}{q_1(1-\gamma)^{1-\theta}-q_0}$ is greater than 1 and is raised to a positive power, while when $\theta > 1$, it is less than 1 and it is raised to a negative power.

Given these expressions, the expected utility of entrepreneurs/workers under the "cutthroat" and "cuddly" incentives, denoted respectively by $s = c$ and $s = o$, can be rewritten as

$$\begin{aligned} W_j^c(t) &\equiv \mathbb{E}\left[U\left(C_j^c(t), e_j^c(t)\right)\right] = \frac{\left(q_1 R^s(t)^{1-\theta} + (1-q_1) R^u(t)^{1-\theta}\right)(1-\gamma)^{1-\theta}}{1-\theta}, \\ W_j^o(t) &\equiv \mathbb{E}\left[U\left(C_j^o(t), e_j^o(t)\right)\right] = \frac{R_0(t)^{1-\theta}}{1-\theta}. \end{aligned}$$

Now using (11) and (12), we can express these expected utilities as:

$$W_j^c(t) = \omega_c N_j(t)^{1-\theta} \text{ and } W_j^o(t) = \omega_o N_j(t)^{1-\theta}, \tag{13}$$

where

$$\omega_c \equiv \frac{\left(q_1 A^{1-\theta} + (1-q_1)\right)(1-\gamma)^{1-\theta}}{\left(q_1 A + (1-q_1)\right)^{1-\theta}} \frac{B^{1-\theta}}{1-\theta} \text{ and } \omega_o \equiv \frac{B^{1-\theta}}{1-\theta}. \tag{14}$$

It can be verified that $\omega_c < \omega_o$, though when $\theta > 1$, it important to observe that we have $\omega_c < \omega_o < 0$. Moreover , we have: (i) ω_c, and thus $(\omega_c/\omega_o)^{1/(1-\theta)}$, is decreasing in A, defined in (9), (since a higher A translates into greater consumption variability); (ii) A is increasing in γ (and thus $(\omega_c/\omega_o)^{1/(1-\theta)}$ is decreasing in γ), which compensates for the higher cost of effort (see Lemma 2 in the Appendix); and (iii) A is non-monotone in θ (because a higher coefficient of relative risk aversion also reduces the disutility of effort).

Two more observations are useful. First, the key expressions ω_c and ω_o depend on agents' coefficient of risk aversion, θ. This will continue to be the case when we adopt the more general preferences given in (8) for the social planner. Second, without loss of any generality, we focus on cutthroat and cuddly incentives for a given entrepreneur throughout (and thus on ω_c and ω_o), since it is optimal to provide cutthroat incentives to the entrepreneur if the social planner wishes to induce effort, and cuddly incentives and thus perfect insurance otherwise. Therefore, the most general reward structure that needs to be considered is one that gives cutthroat incentives to some fraction of the agents and cuddly incentives to the rest at any point in time, and this is the reward structure we consider in our analysis.

4.2 World Equilibrium Given Reward Structures

We next first characterize the dynamics of growth in this world for given reward structures. The next result establishes that when all countries choose asymptotically time-invariant (but potentially different from each other) reward structures, a well-defined world equilibrium exists and involves all countries growing at *the same rate*, set by the rate of growth of the world technology frontier. This growth rate is determined by the innovation rates (and thus reward structures) of either all countries (with (6)) or the leading country (with (5)). In addition, differences in reward structures determine the relative income of each country in the asymptotic equilibrium.

Proposition 1 *Suppose that the reward structure for each country is asymptotically time-invariant (i.e., for each j, $R_j^s(t)/R_j^u(t) \to r_j$). Then starting from any initial condition*

$(N_1(0), ..., N_J(0))$, the world economy converges to a unique stationary distribution $(n_1^*, ..., n_J^*)$, with $n_j(t) \equiv N_j(t)/N(t)$ and $\dot{N}(t)/N(t) = g^*$. In addition, $(n_1^*, ..., n_J^*)$ and g^* are functions of $(r_1, ..., r_J)$. Moreover, with the max specification of the world technology frontier, (5), g^* is a function of only the most innovative country's reward structure, r_ℓ.

Suppose without loss of any generality that $n_\ell^* \geq n_j^*$ for all j. Then, as $\phi \to 0$, for any j with $n_j^* < n_\ell^*$, $n_j^*/n_\ell^* \to 0$.

Proof. The proof of this proposition follows from the material in Chapter 18 of Acemoglu (2009) with minor modifications and is omitted to save space. ■

The process of technology diffusion ensures that all countries grow at the same rate, even though they choose (asymptotically) different reward structures. Countries that do not encourage innovation may first fall behind, but given the form of technology diffusion in equation (3), the advances in the world technology frontier pulls them to the same growth rate as those that provide greater inducements to innovation. The proposition, which will be used in Proposition 5, also shows that in the special case where (5) applies, it will be only innovation and the reward structure in the technologically most advanced country that determines the world growth rate, g^*.

The last part of the proposition is a consequence of the fact that, from (3), as $\phi \to 0$, the world economy approaches a collection of linear endogenous growth economies, each growing at a different rate.

In what follows, we will see that equilibria involve asymptotically constant reward structures, thus the conclusion from Proposition 1 applies and ensures that all countries grow at the same rate regardless of their asymptotic reward structures.

4.3 Equilibrium Reward Structures

We now characterize the equilibrium of the game between country social planners. Throughout, we focus on (pure-strategy) Markov Perfect Equilibria (though, as mentioned above, we do allow for mixed reward structures). The Markovian restriction implies that strategies at time t are only condition on payoff-relevant variables, which are the vector of technology levels, $N_1(t), ..., N_J(t)$.

Country social planners can provide cutthroat reward structures to some of their entrepreneurs while choosing a cuddly reward structure to the rest.[13] Let us define $u_j(t) \in [0,1]$ as the fraction of entrepreneurs receiving a cutthroat reward structure. Clearly, $u_j(t) = 0$ at all points in time corresponds to a cuddly reward structure and $u_j(t) = 1$ for all t corresponds to a cutthroat reward structure throughout.

Slightly generalizing equation (13), we can write average utility in country j at time t as $\omega(u_j(t))N_j(t)^{1-\theta}$, where

$$\omega(u_j(t)) = \omega_o(1 - u_j(t)) + \omega_c u_j(t). \tag{15}$$

[13] It is straightforward to see that it is never optimal to give any entrepreneur any other reward structures than perfect insurance or the cutthroat reward structure that satisfies the incentive compatibility constraint as equality

In addition, from (3), the growth rate of technology of country j adopting a reward structure summarized by $u_j(t) \in [0,1]$, is

$$\dot{N}_j(t) = g(u_j(t))N(t)^{\phi} N_j(t)^{1-\phi}, \tag{16}$$

where

$$g(u_j(t)) = g_o(1 - u_j(t)) + g_c u_j(t), \tag{17}$$

with

$$g_o \equiv q_0\eta \text{ and } g_c \equiv q_1\eta > g_o$$

corresponding, respectively, to the growth rates from fully cuddly and cutthroat reward structures. The fact that $g_c > g_o$ reiterates that a country choosing a cutthroat reward structure will have a faster growth.

4.4 Main Result

In this subsection, we focus on the world technology frontier given by (5), and also assume that at the initial date, there exists a single country ℓ that is the technology leader, i.e., a single ℓ for which $N_\ell(0) = \max\{N_1(0), ..., N_J(0)\}$. Let us also define the relative technology of country j at time t as $n_j(t) \equiv N_j(t)/N(t)$.

We next introduce three assumptions. The first ensures that the cutthroat growth rate, g_c, is not so high as to lead to infinite welfare for the country social planners and will also be maintained throughout (without explicitly being stated):

Assumption 2:

$$\rho - (1-\theta) g_c > 0.$$

The next assumption ensures that if the world consists of a single country, then that country would prefer cutthroat incentives to cuddly incentives.

Assumption 3:

$$\frac{\omega_c}{\rho - (1-\theta)g_c} > \frac{\omega_o}{\rho - (1-\theta)g_o}.$$

More specifically, this assumption ensures that growth at the rate g_c with cutthroat incentives (and less risk sharing) is preferred to growth at the rate g_o with cuddly incentives (and full risk sharing). An important implication of this assumption is that the technology leader will prefer cutthroat to cuddly incentives.

Finally, to start with, we restrict attention to a specific equilibrium selection rule, which imposes that the same country, ℓ, *remains the technology leader throughout.* This is stated in the next assumption.

Assumption 4: $N_\ell(t) = \max\{N_1(t), ..., N_J(t)\}$ for all t.

The selection rule encapsulated in Assumption 4 greatly simplifies the exposition, enabling us to focus on a single equilibrium. Without this selection rule, as we discuss below, there are other equilibria, though arguably the one we focus on here, where the initial technology leader remains so throughout, is the most natural among these. We discuss under what conditions there will be equilibrium uniqueness without this condition in Section 4.7.

The next proposition gives our main result.

Proposition 2 *Suppose country social planners maximize (7), the world technology frontier is given by (5), and Assumptions 1-4 hold. Let*

$$\widetilde{m} \equiv (1-\theta)\frac{(\omega_o - \omega_c)\, g_c + (g_c - g_o)\,\omega_c}{(\omega_o - \omega_c)\,(\rho + \phi g_c)}. \tag{18}$$

Then the world equilibrium is characterized as follows. The leader country ℓ always chooses cutthroat rewards, i.e., $u_\ell(t) = 1$ for all t. For each follower $j \neq \ell$, we have:

1. *If*

$$\widetilde{m} \;< \frac{g_o}{g_c}, \tag{19}$$

   there exist $\overline{m} < g_o/g_c$ and $0 \leq T < \infty$ such that for $n_j(0) < \overline{m}^{1/\phi}$, the reward structure of country j is cutthroat (i.e., $u_j(t) = 1$) for all $t < T$, and cuddly (i.e., $u_j(t) = 0$) for all $t \geq T$; for $n_j(0) \geq \overline{m}^{1/\phi}$, the reward structure of country j is cuddly (i.e., $u_j(t) = 0$) for all t. Regardless of the initial condition, in this case $n_j(t) \to (g_o/g_c)^{1/\phi}$.

2. *If*

$$\frac{g_o}{g_c} < \widetilde{m} < 1, \tag{20}$$

   there exists $0 < T < \infty$ such that for $n_j(0) < \widetilde{m}^{1/\phi}$, the reward structure of country j is cutthroat (i.e., $u_j(t) = 1$) for all $t < T$, and then at $t = T$ when $n_j(T) = \widetilde{m}^{1/\phi}$, the country adopts a "mixed" reward structure and stays at $n_j(t) = \widetilde{m}^{1/\phi}$ (i.e., $u_j(t) = u_j^* \in (0,1)$) for all $t \geq T$; for $n_j(0) > \widetilde{m}^{1/\phi}$, the reward structure of country j is cuddly (i.e., $u_j(t) = 0$) for all $t < T$, and then at $t = T$ when $n_j(T) = \widetilde{m}^{1/\phi}$, the country adopts a mixed reward structure and stays at $n_j(t) = \widetilde{m}^{1/\phi}$ (i.e., $u_j(t) = u_j^* \in (0,1)$) for all $t \geq T$.

3. *If*

$$\widetilde{m} > 1, \tag{21}$$

   then the reward structure of country j is cutthroat for all t (i.e., $u_j(t) = 1$ for all t).

The proof of this proposition, like the proof of all of the remaining results in this paper, is provided in the Appendix. Here, we provide a discussion and interpretation of the results. We start with the three phase diagrams shown in Figures 1, 2 and 3. We have $n_j(t)^\phi$ on the horizontal axis and $\kappa_j(t)$, which is a transform of the co-state variable (multiplier) in the dynamic optimization problem of country j social planner's problem, on the vertical axis. The value $\widetilde{\kappa}$

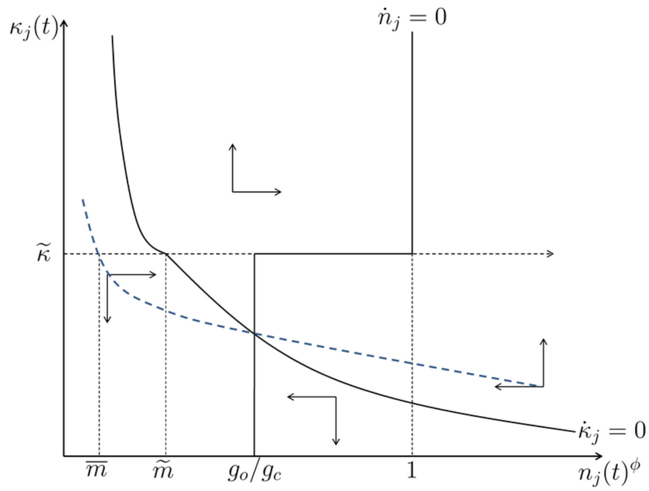

Figure 1: Phase diagram: part 1 of Proposition 2.

makes the social planner indifferent between cuddly and cutthroat incentives (and $\widetilde{m}$ defined in (18) is the corresponding technology level as shown in the figure). The shape of the locus for $\dot{n}_j(t) = 0$ and $\dot{\kappa}_j(t) = 0$ highlights the fact that depending on whether κ is greater or less than $\widetilde{\kappa}$, we will have cutthroat or cuddly incentives, thus changing the dynamics of the system.

Figure 1 depicts part 1 of the proposition, where $\widetilde{m} < g_o/g_c$ and the world equilibrium is asymmetric as emphasized in the Introduction. In particular, in this configuration, the leader, country ℓ, always chooses cutthroat rewards, i.e., $u_\ell(t) = 1$. Moreover, as $t \to \infty$, all followers choose cuddly rewards, i.e., $u_j(t) = 1$ for all $j \neq \ell$, and approach an income level equal to a fraction g_o/g_c of the leader's income level.[14]

The asymmetric nature of the equilibrium here is a consequence of the diffusion of new technologies across countries. Assumption 3 ensures that when all other countries are choosing cuddly incentives, country ℓ prefers to choose cutthroat incentives. In this case, $\frac{\omega_c}{\rho-(1-\theta)g_c}$ is the discounted value of country ℓ from a cutthroat reward structure, ensuring an own (and world growth rate) of g_c. In contrast, because all other countries are choosing a cuddly reward structures, if country ℓ were also to do so, the world economy would only grow at the rate $g_o < g_c$, yielding a discounted value of $\frac{\omega_o}{\rho-(1-\theta)g_o}$ to country ℓ from a cuddly reward structure. Assumption 3 ensures that country ℓ prefers the first option. This comparison reflects the fact that, when all other countries are choosing cuddly reward structures, the incentives provided by country ℓ have a *growth effect* on the world economy (and thus on itself).

However, given the diffusion of technology across countries, once country ℓ chooses cutthroat incentives, the choice of reward structure for other countries has only a *level effect*: a country choosing a cutthroat reward structure would increase its position relative to country ℓ but, in view of Proposition 1, would not change its long-run growth rate. Condition (19) then ensures that in the limit this level effect is more than compensated for by the better risk sharing offered

[14] In addition, below $\overline{m}$, followers start with $u_j(t) = 0$ and then switch to $u_j(t) = 1$ at $\overline{m}$. The figure makes clear that such a $\overline{m}$ will exist if the stable arm intersects the line at $\widetilde{\kappa}$.

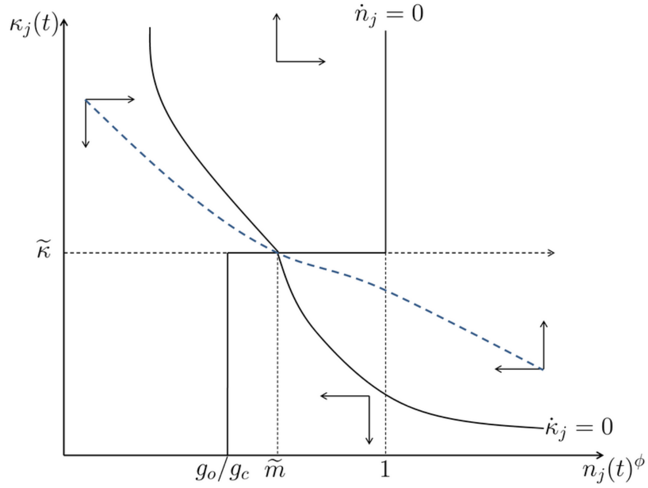

Figure 2: Phase diagram: part 2 of Proposition 2.

by the cuddly reward structure. This contrast between the growth effect of the reward structure of the leader and the level effect of the reward structure of followers is at the root of the asymmetric equilibrium (and the non-existence of a symmetric equilibrium).

It is also worth noting that, without Assumption 4, there could in principle be other equilibria, with some other country playing the role of the leader ℓ and choosing a cutthroat reward structure, while others choose cuddly reward structures. These equilibria are ruled out by the selection rule in Assumption 4, which imposes that the same country remains the technology leader throughout (since in such a situation the country adopting cutthroat incentives would ultimately overtake ℓ). We discuss conditions for uniqueness without Assumption 4 below.

Part 2 of the proposition, where $1 > \widetilde{m} > g_o/g_c$ and the steady state involves mixed rewards, is illustrated in Figure 2, which shows how, for these parameter values, followers will adopt mixed reward structures when they are close to the income level of the leader. With such reward structures some entrepreneurs are made to bear risk, while others are given perfect insurance—and thus are less innovative. This enables them to reach a growth rate between that implied by a fully cuddly reward structure and the higher growth rate of the cutthroat reward structure.

Finally, part 3 of the proposition, corresponding to the case where $\widetilde{m} > 1$, is shown in Figure 3 and involves "institutional" and technology convergence. Now the unique equilibrium path features cutthroat reward structures also for the followers. When this is the case, technology spillovers ensure not only the same long-run growth rate across all countries but convergence in income and technology levels. This contrasts with the patterns in the other cases where countries maintain their different institutions (reward structures), and as a result, they reach the same growth rate, but their income levels do not converge.

We next illustrate some of the patterns that emerge from this proposition more explicitly. The dynamic path in the first part of the proposition is illustrated in Figure 4. Follower j starts with a cutthroat reward structure because it is so far behind the leader that rapid convergence is

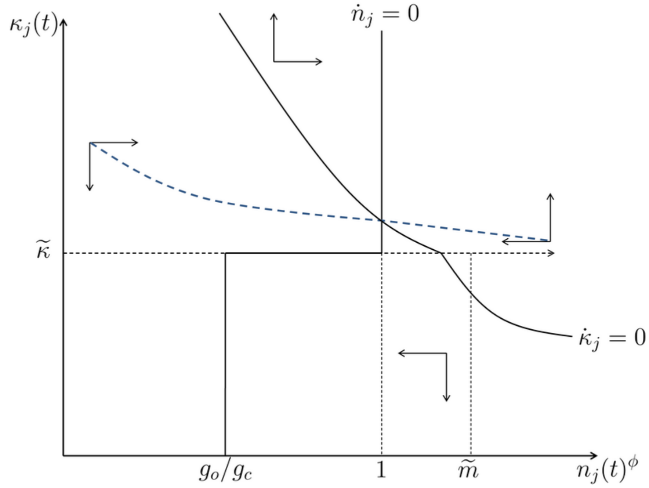

Figure 3: Phase diagram: part 3 of Proposition 2.

beneficial and this is achieved best by adopting cutthroat incentives. However, after $n_j(t)$ exceeds $\overline{m}^{1/\phi}$, so that the follower comes closer to the leader, it adopts a cuddly reward structure. With this cuddly reward structure, the follower converges to a position relative to the leader given by $(g_o/g_c)^{1/\phi}$. This range captures the central asymmetric nature of equilibrium in our model: because the leader is choosing a cutthroat reward structure and pushing the world technology frontier, it makes sense for the follower to only provide cuddly incentives and benefit from better risk sharing. The cost of this is lack of full convergence to the leader's level. But from the welfare point of view, this is more than made up by the better risk sharing within the follower country.

The dynamic path for a follower country starting below $\overline{m}^{1/\phi}$ in this part of the proposition also provides a way of interpreting the growth and social trajectories followed by countries such as South Korea and Taiwan, which adopted institutions with little in the way of a safety net during their early phases of convergence, but then started building a welfare state. From the viewpoint of our model, this path is explained by the high returns to these countries from rapid innovation early on combined with the asymmetric nature of the asymptotic world equilibrium, in which they can build on the technological advances made by the leader country (though of course many other political economy factors played a defining role in the transitions of South Korea and Taiwan towards more democratic institutions, which went hand-in-hand with the development of a safety net and better social insurance in these countries).

Figure 5 illustrates the second part of the proposition, where the asymptotic reward structure of followers is a mixed one where some fraction of its agents receive cutthroat incentives while the rest receive cuddly incentives. The growth rate of this economy at any point in time is thus greater than the one that would have followed from fully cuddly incentives. As a result, this country becomes asymptotically richer relative to one that would have chosen a fully cuddly reward structure. In particular, its asymptotic reward structure is chosen such that the proportional gap between the this country and the leader remains exactly at $\widetilde{m}^{1/\phi}$.

Finally, Figure 6, corresponding to part 3, illustrates the complete convergence between the

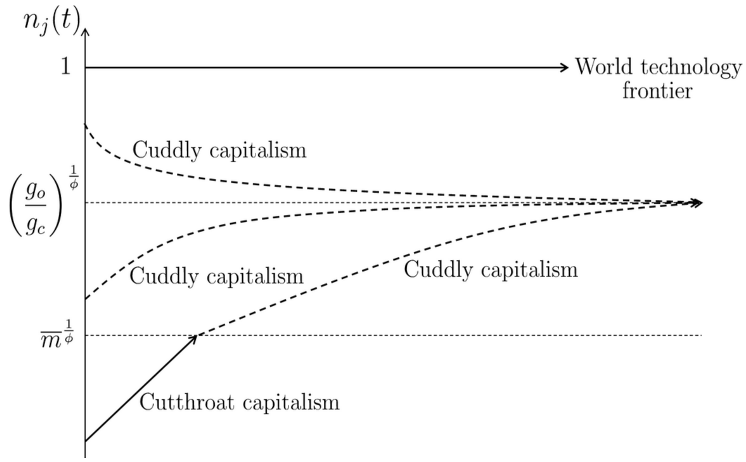

Figure 4: Growth dynamics: part 1 of Proposition 2.

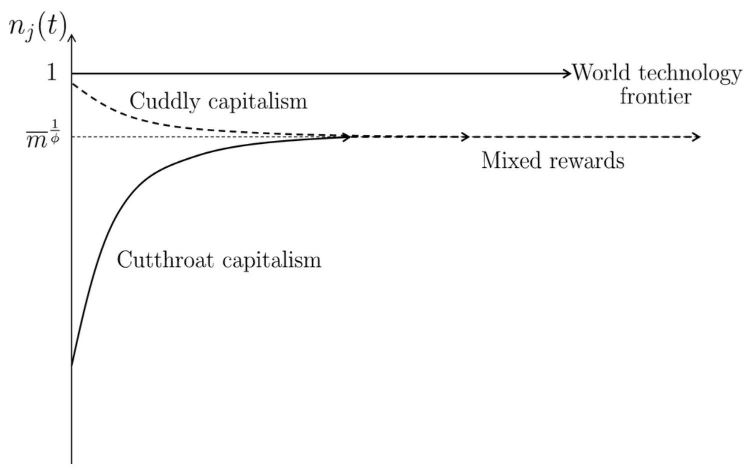

Figure 5: Growth dynamics: part 2 of Proposition 2.

leader and the followers. This is the only case in which the world equilibrium is symmetric, and the proposition makes it clear that the sufficient condition to rule out this case is $\widetilde{m} < 1$, which will be the case, for example, when the cost of effort, γ, or the extent of spillovers, ϕ, is large (see Proposition 3 below).

4.5 Comparative Statics

The next proposition provides a number of comparative static results indicating when an asymmetric world equilibrium is more likely to emerge.

Proposition 3 *The world equilibrium is more likely to be asymmetric (corresponding to part 1 or part 2 of Proposition 2, or equivalently $\widetilde{m} < 1$), when γ is large (the cost of innovation is large) and when ϕ is large (spillovers are large). Also for $\theta < 1$ and γ sufficiently small, the world equilibrium is more likely to be asymmetric when θ is higher (risk aversion is large).*

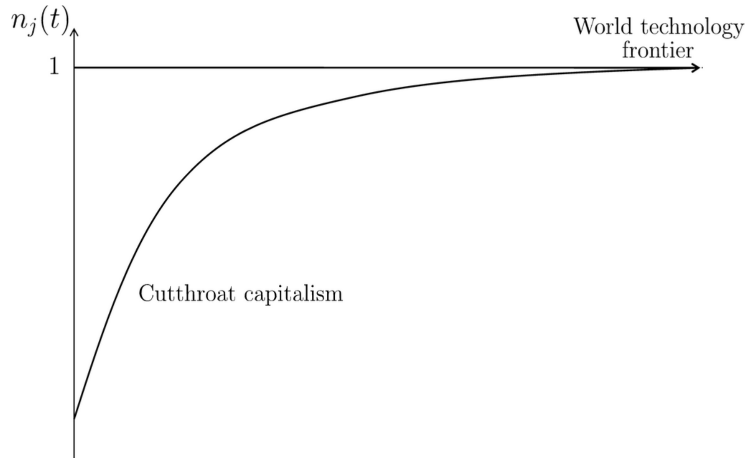

Figure 6: Growth dynamics: part 3 of Proposition 2.

These comparative statics are intuitive. When γ is large, cutthroat incentives become less attractive, though Assumption 3 still ensures that the technology leader, country ℓ, prefers cutthroat incentives. When ϕ is large, the relative gap between the technology leader and followers using cuddly incentives, $(g_o/g_c)^{1/\phi}$, is smaller, encouraging cuddly reward structures. Finally, a higher θ implies greater risk aversion and thus increases the benefits from a cuddly reward structure, which provides consumption insurance, making an asymmetric world equilibrium with followers adopting cuddly reward structures more likely. Nevertheless, because in our baseline model θ also determines the willingness of the social planner to trade-off consumption between cohorts, this result requires us to focus on the case where $\theta < 1$ (we separate these two notions in Section 5).

4.6 Welfare

We next look at the welfare implications of the asymmetric equilibrium we have characterized so far (in particular, parts 1 and 2 of Proposition 2). The most interesting result concerning welfare is that, even though the technological leader, country ℓ, starts out ahead of others and chooses a "growth-maximizing" strategy, average welfare (using the social planner's discount rate) may be lower in that country than in the follower countries choosing a cuddly reward structure. This result is contained in the next proposition and its intuition captures the central economic force of our model: followers are both able to choose an egalitarian reward structure providing perfect insurance to their entrepreneur/workers and benefit from the rapid growth of technology driven by the technology leader, country ℓ, because they are able to free-ride on the cutthroat reward structure in country ℓ, which is advancing the world technology frontier. In contrast, country ℓ, as the technology leader, must bear the cost of high risk for its entrepreneur/workers. The fact that followers prefer to choose the cuddly reward structure implies that, all else equal, country ℓ would have also liked to but cannot do so, because it realizes that if it did, the growth rate of world technology frontier would slow down—while followers know that the world technology

frontier is being advanced by country ℓ and can thus free-ride on that country's cutthroat reward structure.

Proposition 4 *Suppose that society is in part 1 or part 2 of Proposition 2. Then there exists* $\delta > 0$ *such that for all* $n_j(0) > 1 - \delta$*, welfare (at time* $t = 0$*) in country* j *is higher than welfare in country* ℓ*.*

An important implication of this result is that, all else equal, there are benefits to cuddly reward structures. So if we compare an unequal society, with seemingly high-powered incentives, such as the United States, with societies with more egalitarian income distributions and stronger safety nets, such as the Scandinavian countries, welfare will tend to be higher in the latter (provided that they are not too far behind the United States technologically). But importantly, in view of Proposition 2, it is not an equilibrium for the United States to also adopt cuddly incentives, because what enables the rest of the countries to enjoy the benefits of cuddly reward structures is the rapid innovation and used by US cutthroats incentives, and if the United States also switched to cuddly reward structures, this would slow down the world growth rate.

4.7 Uniqueness without the Selection Rule

The analysis so far has proceeded under Assumption 4, which imposed the selection rule that country ℓ remains the technology leader. Absent this selection rule, it is clear that the equilibrium characterized in Proposition 2 may not be unique. Consider, for example, the case in which the world consists of two countries starting with exactly the same level of technology. Then the logic of Proposition 2 (under the same conditions there) ensures that there does not exist a symmetric equilibrium (under parts 1 and 2 of this proposition, i.e., when $\widetilde{m} < 1$). Yet not only are there two mirror-image equilibria, one in which the first country is the technology leader and the other one in which the second country is, but in fact, there may be a myriad of other equilibria in which technology leadership switches between the two countries one or more times over time.[15]

Nevertheless, in this subsection we show that if the technology leader, again denoted by ℓ, is sufficiently ahead of followers, the asymmetric equilibrium characterized in Proposition 2 is unique. The reasoning is simple but important: if followers are significantly behind country ℓ, then any profile of reward structures that involves country ℓ losing the technology leadership will involve the world technology frontier, and thus country ℓ, growing at the slower rate g_o rather than g_c for an extended period of time. Then under Assumption 3, country ℓ would in fact prefer to adopt a cutthroat reward structure. If, in addition, the asymmetric equilibrium in which country ℓ chooses a cutthroat reward structure and the followers choose cuddly reward structures leads to an asymptotic world equilibrium in which the followers still remain significantly behind country ℓ, then the unique equilibrium will involve ℓ choosing a cutthroat reward structure and

[15] Equilibria in which there are switches in technology leadership will be related to the literature on leapfrogging (e.g., Brezis, Krugman and Tsiddon, 1993), though, in contrast to this literature, there is no natural reason why followers should overtake leaders in our model.

the followers choosing cuddly reward structures throughout. This result thus highlights that in an important set of cases—where the technology leader is sufficiently ahead of the rest—the selection assumption we have imposed so far is unnecessary. Otherwise, the selection assumption, Assumption 4, is important for ensuring uniqueness (but the equilibrium we characterized in Proposition 2 of course remains an equilibrium and is arguably the most "focal" one among all equilibria).

In the next proposition, we establish a sufficient condition for the uniqueness of the equilibrium in which the initial technology leader always remains the leader. Rather than stating this result for follower countries that are significantly behind the leader, we equivalently take the gap between the leader and the follower as given, and we establish uniqueness for sufficiently small technology spillovers, ϕ (when ϕ is very small, it will take a very long time for a follower adopting cutthroat incentives to catch up with the leader even if the leader chooses cuddly incentives).

Proposition 5 *Suppose that all of the hypotheses of Proposition 2 hold, $\widetilde{m} < g_o/g_c$ (so that we are in part 1 of Proposition 2), and $n_\ell(0) > n_j(0)$ for all $j \neq \ell$. Then there exists $\bar{\phi} \in (0,1)$ such that for $\phi < \bar{\phi}$, the equilibrium is unique and involves country ℓ adopting a cutthroat reward structures throughout and all other countries choosing a cuddly reward structure asymptotically (i.e., $u_\ell(t) = 1$ for all t and $u_j(t) = 0$ for all $j \neq \ell$ and for t large enough).*

It is also useful to note briefly another reason why equilibria might be unique. If the model is extended so that countries have different sizes, the usual scale effect (e.g., Romer, 1990) implies that the growth rate of the world economy (particularly with the world technology frontier given in (5)) would depend on the size of the technology leader. Then, with a logic similar to that in Proposition 5, under some conditions the largest country would wish to be the technology leader by choosing cutthroat incentives because technology leadership by smaller country would reduce the growth rate of the world technology frontier. This economic force would also resolve the multiplicity problem (without Assumption 4).

4.8 Discussion of Modeling Assumptions

Here we discuss the role and interpretation of some of the modeling assumptions we have adopted so far. Those already discussed, such as Assumptions 1-4, will not be discussed further.

Five assumptions deserve to be highlighted.

The first is that our agents are short-lived. This assumption is adopted for simplicity. Clearly, if agents themselves are forward-looking with the same preferences as (8), but are only given a short-term incentives, nothing in the analysis changes. However, with such long-lived agents, optimal contracts designed to deal with the moral hazard problem would be more complicated, involving rewards given as a function of the entire history of success and failure in innovation. Our assumption abstracts from such dynamic incentives which are not central to our focus.

The second assumption implicit in our approach concerns the specific form of the moral hazard problem whereby greater innovation effort follows from less risk sharing (more "high-powered" rewards). Though this is natural, it should be noted that there are alternatives. For

example, one could formulate a model in which entrepreneurs take greater (socially efficient) risks when there is better risk sharing. This might follow from the presence of partially uninsured income risk for entrepreneurs. Though this is an interesting avenue to pursue, what we have focused on is the canonical moral hazard problem highlighting the risk-reward trade-off in risky activities (entrepreneurship). We believe that investigating the implications of partially uninsured income risk for risk-taking is an interesting area for research and the exact implications of risk sharing for innovative activities needs to be investigated both theoretically and empirically in future work.

Third, the assumption that there are only two levels of effort also significantly simplifies our analysis. Without this assumption, the main economic forces in our model would still lead to an asymmetric equilibrium with one country becoming the technology leader and choosing to induce higher effort than the follower countries (when the world technology frontier takes the form of (5); otherwise, there will typically be several countries playing the role of the technology leader). But the degree of asymmetry would change over time because the effort level in all countries would change away from the asymptotic equilibrium. In addition to making the analysis somewhat more involved, this would also make it harder to map the model to the "cutthroat" and "cuddly" styles of capitalism. We have therefore simplified the analysis and the discussion by focusing on two levels of effort.

Fourth, the global linkages in our model are purely technological, whereas in reality there are trade-induced and financial linkages also. Though we believe the technological linkages we have emphasized, which imply that when one country is advancing the world technology frontier this creates the possibility for others to free-ride on this effort, is important, some of these other linkages also lead to asymmetric world equilibria. For instance, when countries trade, there may be complementary reasons for countries to end up with different reward structures; some sectors may benefit from cutthroat reward structures while others benefit from the better risk-sharing resulting from cuddly were structures. In a world equilibrium with trade, some countries will specialize in sectors benefiting from cutthroat incentives and may then find it beneficial to provide such incentives, while others opt for cuddly incentives and specialize in sectors that benefit from such cuddly incentives.[16]

Fifth, we have also simplified the analysis by assuming that reward structures are determined to maximize discounted utility not by domestic political economy considerations. This assumption is also adopted for simplicity, and in reality, political economy considerations are central, and different reward structures will create different winners and losers. Nevertheless, abstracting from these considerations has enabled us to clearly delineate the key force leading to an asymmetric world equilibrium. In Section 6, we show how domestic political constraints, modeled in a simple manner, interact with the forces we have highlighted so far and play the role of selecting which country will be a technology leader with cutthroat incentives.

[16] Relatedly, see Acemoglu and Ventura (2001) for terms of trade effects linking growth and welfare across countries, Chatterjee (2010) for a model of asymmetric policies creating comparative advantage, and Levchenko (2007) on the linkages between institutions and trade.

5 Extensions

In this section, we extend our main results in two dimensions. First, we characterize the world equilibrium under the more general preferences introduced in (8). Second, we show that they hold with the general convex form of the world technology frontier given in (6).

5.1 General Preferences

In this subsection, we return to the more general preferences introduced in (8), which separate the coefficient of relative risk aversion from the parameter that determines the willingness of the social planner to trade-off consumption differences across cohorts. This requires us to modify Assumptions 2 and 3 their equivalents in this case where $\theta \neq \lambda$.

Assumption 2′:

$$\rho - (1-\lambda)\, g_c > 0.$$

Notice that once $\theta \neq \lambda$, what is relevant to ensure that the social planner's discounted utility is bounded away from infinity is λ, since it regulates how future increases in utility are valued today.

Assumption 3′:

$$\left(\frac{\omega_c}{\omega_o}\right)^{\frac{1}{1-\theta}} > \left(\frac{\rho-(1-\lambda)g_c}{\rho-(1-\lambda)g_o}\right)^{\frac{1}{1-\lambda}}.$$

Proposition 6 *Suppose country social planners maximize (8), the world technology frontier is given by (5) and Assumptions 1,2′, 3′ and 4 hold, $\lambda > \theta$ and $\phi > \frac{(1-\theta)(1-\lambda)}{(\lambda-\theta)}$. Let*

$$\widetilde{m}_o \equiv \frac{(1-\theta)(g_c-g_o)\omega_o + (1-\lambda)g_o(\omega_o-\omega_c)}{(\omega_o-\omega_c)(\rho+\phi g_c)},$$

and

$$\widetilde{m}_c \equiv \frac{(1-\theta)(g_c-g_o)\omega_c + (1-\lambda)g_c(\omega_o-\omega_c)}{(\omega_o-\omega_c)(\rho+\phi g_c)}.$$

Then the world equilibrium is characterized as follows. The leader country ℓ always chooses cutthroat rewards, i.e., $u_\ell(t) = 1$ for all t. For each follower $j \neq \ell$, we have:

1. *If*

$$\frac{g_o}{g_c} > \widetilde{m}_o > \widetilde{m}_c, \tag{22}$$

there exist $\underline{\underline{m}}$, $\underline{m}$, T ,$T' \geq 0$ such that for $n_j(0) < \underline{\underline{m}}^{1/\phi}$, the reward structure of country j is cutthroat (i.e., $u_j(t) = 1$) for all $t < T$, and then at $t = T$, we have $n_j(T) = \underline{\underline{m}}^{1/\phi}$ and country j adopts a "mixed" reward structure until T' (i.e., $u_j(t) \in (0,1)$) for all $T' > t \geq T$. Then at $t = T'$, we have $n_j(T') = \underline{m}^{1/\phi}$ and country j switches to a cuddly reward structure (i.e., $u_j(t) = 0$) for all $t \geq T'$, and $n_j(t) \to (g_o/g_c)^{1/\phi}$.

2. *If*

$$\frac{g_o}{g_c} < \widetilde{m}_o \text{ and } \widetilde{m}_c < 1, \tag{23}$$

then there exist $\underline{\underline{m}}$ *and* $\underline{m}$ *such that for* $\underline{m}^{1/\phi} < n_j(0) < \underline{\underline{m}}^{1/\phi}$*, the reward structure of country* j *is mixed (i.e.,* $u_j(t) \in (0,1)$*) for all* t*, and* $(m_j(t), u_j(t)) \to (m^*, u^*)$*. If* $n_j(0) < \underline{\underline{m}}^{1/\phi}$*, then country* j *first adopts a cutthroat reward structure (i.e.,* $u_j(t) = 1$*) until some* $T \geq 0$*, and then switches to a mixed reward structure, again converging to a unique* (m^*, u^*)*, and if* $n_j(0) > \underline{m}^{1/\phi}$*, then country* j *first adopts a cuddly reward structure (i.e.,* $u_j(t) = 0$*) until some* $T' \geq 0$*, and then switches to a mixed reward structure, again converging to a unique* (m^*, u^*)*.*

3. *If*

$$\widetilde{m}_o > \widetilde{m}_c > 1, \tag{24}$$

then for any $n_j(0) < 1$*, the reward structure of country* j *is cutthroat for all* t *(i.e.,* $u_j(t) = 1$ *for all* t*).*

The economics of the results in this proposition are very similar to those in Proposition 2, and the three parts of this proposition are the analogs of the three parts in Proposition 2. The new features involve a clearer illustration of the roles of the coefficient of relative risk aversion, θ, and the willingness of the social planner to substitute consumption across cohorts given by, λ, and somewhat richer dynamics where mixed rewards can arise along the transition path (as opposed to Proposition 2 where mixed rewards only emerge at the end of a transition path).

5.2 General Convex Aggregators for World Technology Frontier

We next show that the main result of this section holds with general aggregators of the form (6) provided that these aggregators are sufficiently "convex," i.e., putting sufficient weight on technologically more advanced countries (we also return to the baseline preferences given by (7), but the same result holds with (8)). The main difference from the rest of our analysis is that with such convex aggregators, the world growth rate is no longer determined by the reward structure (and innovative activities) of a single technology leader, but by a weighted average of all economies. Nevertheless, the same economic forces our present because the convexity of these aggregators implies that the impact on the world growth rate of a change in the reward structure of a technologically advanced country would be much larger than that of a backward economy, and this induces the relatively advanced economies to choose cutthroat reward structures, while relatively backward countries can free-ride and choose cuddly reward structures, safe in the knowledge that their impact on the long-run growth rate of the world economy (and thus their own growth rate) will be small.

Proposition 7 *Suppose that Assumptions 1 and 2 hold, that* $\omega_o/\omega_c > g_o/g_c$ *and that the world technology frontier is given by (6). Then there exist* $\bar{\sigma} < 0$*,* $\bar{\rho} > 0$*,* $\overline{\theta} < 1$ *and* $\bar{\gamma} < 1$ *such that when* $\sigma \in (\bar{\sigma}, 0)$*,* $\rho \leq \bar{\rho}$ *,* $\theta > \overline{\theta}$*, and* $\gamma > \bar{\gamma}$*, there is no symmetric world equilibrium with all*

countries choosing the same reward structure. Instead, there exists $T < \infty$ such that for all $t > T$, a subset of countries will choose a cutthroat reward structure while the remainder will choose a cuddly or mixed reward structure.

Observe that the assumption that $\rho \leq \bar{\rho}$ and $\omega_o/\omega_c > g_o/g_c$ replaces Assumption 3 for this case, and $\gamma > \bar{\gamma}$ ensures that a symmetric equilibrium in which all countries adopt cutthroat incentives does not exist. The condition that σ has to be above some $\bar{\sigma} < 0$ is also intuitive; if σ approaches $-\infty$ (so that the world technology frontier becomes linear), technologically more advanced and more backward economies have similar contributions to the world technology frontier, removing the economic rationale for an asymmetric equilibrium where the contributions of the more advanced economies to world technology enable the rest to choose cuddly incentives. Finally the condition $\theta > \overline{\theta}$ implies that the risk sharing problem has to be sufficiently important to ensure that a cuddly strategy is attractive for some countries. Note also that we are not imposing Assumption 4 in this case because this proposition does not characterize the full equilibrium dynamics, where Assumption 4 was previously used; rather, it shows that asymptotically some countries will adopt cutthroat incentives while the rest do not (which is true without relying on Assumption 4).

6 Equilibrium under Domestic Political Constraints

In this section, we focus on the world economy with two countries, j and j', and we also relax Assumption 4, so that asymmetric equilibria in which countries that initially start out behind later become the technology leader are possible. We also simplify the discussion by assuming that $n_{j'}(0) = n_j(0)$ and also by assuming that the world technology frontier is given by (5). This implies that there are two asymmetric equilibria, one in which country j is the technology leader and j' the follower, and vice versa. We also suppose that the social planner in country j is subject to domestic political constraints imposed by a social democratic party or a labor movement, which prevent the ratio of rewards between successful and unsuccessful entrepreneurs from exceeding some amount ζ. There are no domestic constraints in country j'. If $\zeta \geq A$, then domestic constraints have no impact on the choice of country j, and there continue to be two asymmetric equilibria.

Suppose instead that $\zeta < A$. This implies that because of domestic political constraints, it is impossible for country j to adopt a cutthroat strategy regardless of the strategy of country j'. This implies that of the two asymmetric equilibria, the one in which country j adopts a cutthroat reward structure disappears, and the unique equilibrium becomes the one in which country j' adopts the cutthroat strategy and country j chooses an egalitarian structure. However, from Proposition 2 above, this implies that country j will now have higher welfare than in the other asymmetric equilibrium (which has now disappeared). This simple example thus illustrates how domestic political constraints, which restrict the amount of inequality in society, can create an advantage in the world economy.

We next show that this result generalizes to the case in which the two countries do not start with the same initial level of technology.

Proposition 8 *Suppose that the world technology frontier is given by (5), Assumptions 1-3 hold, and $\widetilde{m} < 1$ as in parts 1 and 2 of Proposition 2 so that the world equilibrium is asymmetric. Suppose also that there are two countries j and j' with initial (relative) technology levels $n_{j'}(0) \geq n_j(0)$ (which is without loss of any generality). Then:*

1. *There exists $\delta > 0$ such that for all $n_j(0) > 1 - \delta$, there are two two types of asymmetric equilibria, one type in which asymptotically country j adopts a cutthroat reward structure and country j' adopts a cuddly reward structure, and vice versa.*

2. *If domestic constraints imply that country j cannot adopt a cutthroat reward structure, then the unique equilibrium is the one in which country j' adopts a cutthroat reward structure and country j asymptotically adopts a cuddly reward structure. The equilibrium welfare of country j is greater than that of country j'.*

An interesting implication of this result is that country j, which has a stronger social democratic party, benefits in welfare terms by having both equality and rapid growth, but "exports" its potential labor conflict to country j', which now has to choose a reward structure with significantly greater inequality.

In the context of the comparison of the US to Scandinavian economies, the latter clearly have a history of stronger labor movement and social democratic party, suggesting that this might have been one of the factors influencing the specific pattern of asymmetric world equilibrium that has developed over the last several decades. Friedman (2010), for example, provides an overview of existing cross-national historical data on union density which shows that in 1928, just prior to the date when the Swedish Social Democrats took power (1932), its unionization rate was 32%. In Denmark, this was 39.7% and in Norway 17.4%. In the US this number was 9.9% in 1928. It is not just the extent of unionization, but how it is organized that mattered. American Unions like those of Britain tended to be along craft lines with multiple unions in a single firm, making it much more difficult for the labor movement to act collectively. In Scandinavia, unions were organized by industry and were much more encompassing. Finally, the rise of social democratic parties which cemented the Scandinavian model into place had its roots in the late 19th century (see Lundberg and Åmark, 2001, for Sweden, Baldwin, 1992, more generally), and its particular universalistic and tax financed nature was a result of what Gourevitch (1986) calls the "Red-Green Coalition," which linked poor rural peasants with urban industrial workers (see also Baldwin, 1992). In the US similar debates took place and the coalitions that formed within the Populist and Progressive movements had Red-Green elements. Though they also managed to push progressive reforms, such as the introduction of the income tax in 1913, these movements were weaker and failed to unite with a factious labor movement. While the 1920s saw left-wing political parties come to power in Denmark (the Social Democrats came to power in 1924) and

Norway (Norwegian Labor Party formed its first government in 1928), left-wing US political parties lost their power during this period. Thus just at the critical juncture where many of the institutions of 20th century developed countries states were formed, the strength of labor movements and left-wing political parties was much greater in Scandinavia than in the US.

7 Case Study Evidence from the Pharmaceutical Industry

In this section we discuss the industrial organization of the global pharmaceutical sector, which, we argue, is consistent with the major assumptions and predictions of the model. We seek to establish four main claims: (i) that there are large cross-national spillovers in pharmaceutical research and development; (ii) that in spite of the joint, global fixed cost nature of R&D costs in the pharmaceutical sector, there are large, persistent differences in drug prices/markups between the US and other OECD countries which, similar to our distinction between cuddly and cutthroat reward structures, translate into different rewards to firms offering new drugs. Moreover, these differences seem to arise mainly from the absence or presence of drug price controls; (iii) that "cuddly" countries, with drug price controls that result in cheaper domestic access to drugs, contribute less to global new drug discoveries than does the "cutthroat" US; and (iv) that politicians in the US and in other OECD countries are aware of and seek to maintain this discriminatory pricing arrangement, suggesting an asymmetric global political economy equilibrium along the lines of our model.

Research and development plays a central role in the pharmaceutical industry, with R&D accounting for roughly $800 million per major new drug according to one study (DiMasi et al., 2003) and roughly 30 percent of total costs by some estimates (Danzon, 1997). Existing evidence suggests that not only the overall amount of innovation, but the direction of innovation in the sector strongly responds to profits incentives (Acemoglu and Linn, 2004, Finkelstein, 2004). Investments by firms in R&D have large cross-national spillovers, because once a new drug is discovered no marginal R&D costs are involved in bringing the drug to market in specific countries (Danzon, 2003). For this reason, the pharmaceutical industry is dominated by multi-national companies, which market new drugs globally, either directly or indirectly through licensing. The pharmaceutical sector is one of the most prone to free-riding on innovation.[17]

Consumers pay very different costs across countries for similar drugs, mainly as a result of government policy. Golec et al. (2006) demonstrate, using drug price indices, that between 1986 and 2004, after adjusting for inflation, drug prices remained stable in the European Union but rose dramatically in the United States. The difference in prices for the exact same drugs between the neighboring US and Canada is a particularly stark illustration of this phenomenon. Quon et al (2005), find that brand-name medications are approximately 24% cheaper on the websites of Canadian pharmaceutical retailers than on the websites of US drug chain pharmacies. These higher drug prices in the US are widely argued to be the result of a Congressional ban on the

[17] Mansfield (1986) suggests that innovations are protected by patents to a much greater extent in the pharmaceutical industry than in any other sector.

federal government negotiating drug prices with companies, whereas all other OECD countries have adopted drug price controls (Scherer, 2004). Cross-border price disparities are preserved via laws that prohibit the re-importation of prescription medicines from other countries; nonetheless, there is a large illegal cross-border drug re-importation trade (Bhosle and Balakrishnan, 2007).

The trade-off between "the affordability of drugs and technological progress" (Scherer, 2004) is widely acknowledged in the pharmaceutical industry. In our model, countries at the world technology frontier disproportionately contribute to the incentives for innovation while countries behind the frontier can free ride upon these incentives and enjoy "cuddlier" domestic institutions. While the issue is difficult to assess empirically, given the multi-national nature of pharmaceutical companies, there is prima facie empirical evidence that US consumers contribute disproportionately to new drug development via higher drug prices. Since US-based firms tend to sell disproportionately in the US market compared to EU firms, it is possible to compare industries to assess relative rates of profitability and innovation. Golec et al. (2006) report that between 1993-2004, U.S.-based pharmaceutical firms averaged 17.1 percent profitability (in terms of operating income divided by total assets) while EU-based firms averaged 12.2 percent profitability. In the same period US firms spent 9.1% of the value of their total assets on R&D, while EU firms spent 8.9%. They also report data on the aggregate size of R&D investments, noting that between 1993 and 2004, US firms' assets and R&D investments have grown at a much faster pace than EU firms'. To the extent that these differences reveal the impact of EU price controls upon innovation, they are likely underestimates, since EU-based firms also market drugs in the more profitable US market.

Additional evidence links profitability and drug prices to innovation intensity. Danzon and Wang (2004) analyze the launch of 85 new drugs in 25 major markets and find that while drugs reach all major markets they are launched earlier in countries– notably the US – with higher expected drug prices. Scherer (2001) finds that within the US, the single largest market for U.S. drug companies' products, R&D outlays closely track profits over time, both in term of long-term trends and short-term deviations from trend, a finding echoed by Giaccotto et al. (2005).

Kneller (2010) reports that, between 1998 and 2007, the majority of the 252 new drugs approved by the Food and Drug Administration in the United States, the single largest market for pharmaceutical products were patented by US-based firms and research entities.

Finally, there is evidence that policy makers in the US and in other OECD countries are aware of this discriminatory pricing arrangement but nevertheless see it as their best response to preserve the arrangement. This is seen clearly in the efforts in both the US and Canada to prevent the equalization of drug prices via cross-border trade. Though more than 70% of American above the age of fifty report that they would buy drugs from Canada if it were allowed (Choudhry and Detsky, 2005), prescription drug re-importation is illegal in the US (Bhosle and Balakrishnan, 2007). The Canadian Pharmacists' Association supports a ban of the exportation of drugs from Canada to the US on the grounds that "allowing Canadian

price-controlled medicines to be exported to the United States will damage the Canadian drug supply and could very well lead to increased drug prices for Canadians". Proposals in the US, meanwhile, that the federal government negotiate drug prices with pharmaceutical companies have been consistently opposed by conservative politicians and defeated on the grounds that it would stifle innovation and competition (Moffitt, 2013). Strikingly, the US did not implement drug price controls or negotiation as part of the adoption of the Affordable Healthcare Act. In other OECD countries, governments utilize their bargaining power as managers of national healthcare schemes to negotiate and control drug prices. Though it is difficult to separate absence of price controls in the US from the lobbying of rent-seeking pharmaceutical companies (Marmor and Hacker, 2005), it is clear that politicians have been more successful in leveraging these arguments to block drug price controls in the US than politicians in other OECD countries.

Thus, consistent with the model, we see that in a sector where global free-riding on innovation is a salient dynamic, some countries adopt relatively more "cutthroat" or "cuddly" incentives for pharmaceutical innovation via price control policies. Consistent with the notion of a stable, asymmetric institutional equilibrium in our model, policy makers in both cuddly and cutthroat systems appear to understand the economic forces that make this discriminatory pricing arrangement stable.

8 Conclusion

In this paper, we have taken a first step towards a systematic investigation of institutional choices in an interdependent world—where countries trade or create knowledge spillovers on each other. Focusing on a model in which all countries benefit and potentially contribute to advances in the world technology frontier, we have suggested that the world equilibrium may necessarily be asymmetric. In our model economy, because effort by entrepreneurs is private information, a greater gap of incomes between successful and unsuccessful entrepreneurs—thus greater inequality—increases innovative effort and a country's contributions to the world technology frontier. Under plausible assumptions, in particular with sufficient risk aversion and a sufficient return to entrepreneurial effort, some countries will opt for a type of "cutthroat" capitalism that generates greater inequality and more innovation and will become the technology leaders, while others will free-ride on the cutthroat incentives of the leaders and choose a more "cuddly" form of capitalism.

We have also shown that, somewhat paradoxically, starting with similar initial conditions, those that choose cuddly capitalism, though poorer, will be better off than those opting for cutthroat capitalism. Nevertheless, this configuration is an equilibrium because cutthroat capitalists cannot switch to cuddly capitalism without having a large impact on world growth, which would ultimately reduce their own welfare.

This perspective therefore suggests that the diversity of institutions we observe among relatively advanced countries, ranging from greater inequality and risk-taking in the United States to the more egalitarian societies supported by a strong safety net in Scandinavia, rather than

reflecting differences in fundamentals between the citizens of these societies, may emerge as a mutually self-reinforcing world equilibrium. If so, in this equilibrium, "we cannot all be like the Scandinavians," because Scandinavian capitalism depends in part on the knowledge spillovers created by the more cutthroat American capitalism.

Clearly, the ideas developed in this paper are speculative. We have theoretically shown that a specific type of asymmetric equilibrium emerges in the context of a canonical model of growth—with knowledge spillovers combined with moral hazard on the part of entrepreneurs. Whether these ideas contribute to the actual divergent institutional choices among relatively advanced nations is largely an empirical question. We hope that our paper will be an impetus for a detailed empirical study of these issues.

Appendix: Proofs of Main Results from the Text

Derivation of Equation (12) . To derive (12), we need to characterize the equilibrium prices and quantities in country j as a function of current technology $N_j(t)$. This follows directly from Chapter 18 of Acemoglu (2009). Here it suffices to note that the final good production function (1) implies iso-elastic demand for machines with elasticity $1/\beta$, and thus each monopolist will charge a constant monopoly price of $\psi/(1-\beta)$, where recall that ψ is the marginal cost in terms of the final good of producing any of the machines given its blueprint (invented or adapted from the world technology frontier). Our normalization that $\psi \equiv 1-\beta$ then implies that monopoly prices and equilibrium quantities are given by $p_j^x(\nu,t)=1$ and $x_j(\nu,t)=L_j=1$ for all j, ν and t. This gives that total expenditure on machines in country j at time t will be $X_j(t)=(1-\beta)\,N_j(t)$, while total gross output is

$$Y_j(t)=\frac{1}{1-\beta}N_j(t).$$

Therefore, total net output, left over for distributing across all workers/entrepreneurs is $NY_j(t) \equiv Y_j(t)-X_j(t)=BN_j(t)$, with $B \equiv \beta(2-\beta)/(1-\beta)$ as in equation (10) in the text, leading to (12). ■

Proof of Proposition 2. We now present a proof of our main result, Proposition 2. Suppose country ℓ adopts a cutthroat reward structure at all times, i.e., $u_\ell(t)=1$ for all t. Then the problem of follower country j's social planner can then be written as

$$\begin{aligned}
\mathcal{W}_j(N_j(t),N_\ell(t)) &= \max_{u_j(\cdot)\in[0,1]}\int_t^\infty e^{-\rho(\tau-t)}\omega(u_j(\tau))N_j(\tau)^{1-\theta}\,d\tau \\
\text{such that } \dot{N}_j(\tau) &= g((u_j(\tau))\,N_\ell(\tau)^\phi\,N_j(\tau)^{1-\phi}, \\
\text{with } N_\ell(\tau) &= N(t)\,e^{g_c(\tau-t)} \text{ (for } \tau\ge t).
\end{aligned}$$

With a change of variable for $m_j \equiv (N_j/N_\ell)^\phi \le 1$, this can be written as:

$$\begin{aligned}
\mathcal{W}_j(m_j(t)) &= N_\ell(t)\max_{u(.)\in[0,1]}\int_t^\infty e^{-(\rho-(1-\theta)g_c)(\tau-t)}\omega(u(\tau))m_j(\tau)^{\frac{1-\theta}{\phi}}\,d\tau \qquad \text{(A1)}\\
\dot{m}_j(\tau) &= \phi\left[g(u(\tau))-g_c m_j(\tau)\right].
\end{aligned}$$

The solution to this problem would be the "open loop" best response of follower j to the evolution of the world technology frontier driven by the technology leader, ℓ. The Markov Perfect Equilibrium corresponds to the situation in which all countries use "closed loop" strategies. However, given our selection rule, the same country, ℓ, remains the leader and adopts a cutthroat reward structure (ensured by Assumption 3), the open and the closed loop solutions coincide, because under this scenario, country ℓ always adopts a cutthroat reward structure, regardless of the strategies of other countries. Hence we can characterize all equilibria by deriving the solution to (A1).

Let us next define the current-value Hamiltonian, suppressing the country index j to simplify notation,

$$H(m\left(t\right), u\left(t\right), \mu\left(t\right)) = \omega(u\left(t\right))m\left(t\right)^{\frac{1-\theta}{\phi}} + \mu\left(t\right)\phi\left[g(u\left(t\right)) - g_c m\left(t\right)\right],$$

where $\mu(t)$ is the current-value co-state variable. We next apply the Maximum Principle to obtain a candidate solution. This implies for the control variable (reward structure) $u\left(t\right)$ the following bang-bang form:

$$u\left(t\right)\begin{cases} =1 & \Psi\left(t\right)<0 \\ \in[0,1] \quad \text{if} & \Psi\left(t\right)=0 \\ =0 & \Psi\left(t\right)>0 \end{cases} \tag{A2}$$

where $\Psi\left(t\right)$ is the switching function:

$$\Psi\left(t\right) \equiv \left(\omega_o - \omega_c\right) m\left(t\right)^{\frac{1-\theta}{\phi}} - \mu\left(t\right)\phi\left[g_c - g_o\right]. \tag{A3}$$

In addition,

$$\begin{aligned} \dot{m}\left(t\right) &= \phi\left[g(u\left(t\right)) - g_c m\left(t\right)\right] \quad \text{with } m\left(0\right)>0 \text{ given} \\ \dot{\mu}\left(t\right) &= \left(\rho-(1-\theta)g_c+\phi g_c\right)\mu\left(t\right) - \frac{1-\theta}{\phi}m\left(t\right)^{\frac{1-\theta}{\phi}-1}\omega\left(u\left(t\right)\right), \end{aligned} \tag{A4}$$

and the transversality condition,

$$\lim_{t\to\infty} e^{-(\rho-(1-\theta)g_c)t}\mu(t) = 0. \tag{A5}$$

Now differentiating (A3) and combining it with (A4), we have

$$\dot{\Psi}(t) = \left(\rho-(1-\theta)g_c+\phi g_c\right)\Psi(t) + (\omega_o-\omega_c)(\rho+\phi g_c)m(t)^{\frac{1-\theta}{\phi}-1}\left(\widetilde{m}-m(t)\right), \tag{A6}$$

where $\widetilde{m}$ is given by (18) in the statement of Proposition 2. Integrating (A6), we obtain

$$\Psi(t) = (\omega_o-\omega_c)(\rho+\phi g_c)\int_t^\infty e^{-((\rho-(1-\theta)g_c+\phi g_c))(\tau-t)}m(\tau)^{\frac{1-\theta}{\phi}-1}\left(m(\tau)-\widetilde{m}\right)d\tau. \tag{A7}$$

Note also a special feature of this problem. We have that

$$\frac{\partial\left[\omega\left(u(t)\right)m\left(t\right)^{\frac{1-\theta}{\phi}}\right]/\partial u}{\partial\left[g(u\left(t\right))-g_c m\left(t\right)\right]/\partial u} = \frac{\omega_c-\omega_o}{g_c-g_o}m\left(t\right)^{\frac{1-\theta}{\phi}}$$

is independent of $u(t)$. Therefore, from Proposition 2 of Spence and Starrett (1975), whenever a candidate solution that reaches a steady-state in finite time exists, this defines a Most Rapid Approach Path (MRAP) that gives the unique global maximum.

In this light, now consider first part 2 of the proposition, where $1 > \widetilde{m} > g_o/g_c$. Then

$$u(t) = \begin{cases} 0 & \text{if } m(t) > \widetilde{m} \\ u^* & \text{if } m(t) = \widetilde{m} \\ 1 & \text{if } m(t) < \widetilde{m} \end{cases}$$

with u^* given such that $\widetilde{m} = g(u^*)/g_c$ satisfies (A2) and defines a MRAP (note that when $m(t) = \widetilde{m}$, $\dot{m}(t) = 0$ and $\Psi(t) = 0$), and is thus the unique global maximum. This establishes part 2 of the proposition.

Next consider part 3 where $\widetilde{m} > 1$. In this case, $\Psi(t) < 0$ for all t (regardless of initial conditions), and thus $u(t) = 1$ for all t defines a MRAP, establishing part 3 of the proposition.

Finally, consider part 1, where $\widetilde{m} < g_o/g_c$. If $m(0) \geq g_o/g_c$, then (A4) implies that $m(t) \geq g_o/g_c > \widetilde{m}$ for all t. Hence (A7) implies that $\Psi(t) > 0$ for all t, and thus $u(t) = 0$ for all t (which also implies from (A4) that $m(t)$ is monotonically decreasing towards g_o/g_c). This again defines a MRAP, yielding the desired result.

The only remaining case is where $\widetilde{m} < g_o/g_c$ and $m(0) < g_o/g_c$. Note that in this case the solution must have $u(t) = 0$ or $u(t) = 1$ for almost all times (since $m(t) = \widetilde{m}$ for all t is not feasible). We now prove, with the help of the next lemma, that in this case the unique optimal path is given by part 1 of the proposition, i.e., $u(t) = 1$ for all $t \leq T$ and $u(t) = 0$ for all $t > T$ for some T (or conversely for $n(t) \geq \overline{m}^{1/\phi}$ for some $\overline{m}$).

Lemma 1 *When $\widetilde{m} < g_o/g_c$, there exists $T < \infty$ such that $u(t) = 1$ for $t < T$ and $u(t) = 0$ for $t > T$.*

Proof. First, $u(t) = 1$ for all t is not optimal in view of the fact that $\widetilde{m} < g_o/g_c$. Therefore, there exists at least some interval in which $u(t) = 0$. Take $[T_1, T_2)$ to be the the first such interval.

If $T_2 = \infty$, the lemma is proved. To obtain a contradiction that $T_2 < \infty$. We then can show that this leads to a contradiction.

Note first that $\Psi(t)$ defined by (A7) is continuously differentiable. Moreover, by definition, $\Psi(T_1) = \Psi(T_2) = 0$ and $\Psi(t) < 0$ for $t \in (T_1, T_2)$. This implies that

$$m(t) = m(T_1)\, e^{-\phi g_c (t - T_1)} + \frac{g_o}{g_c}\left(1 - e^{-\phi g_c (t - T_1)}\right) \text{ for } t \in [T_1, T_2).$$

Once again because $\widetilde{m} < g_o/g_c$, $u(t) = 1$ at all time $t \geq T_2$ is not optimal, and thus there exists $T_3 < \infty$ such that $u(t) = 1$ and thus $\Psi(t) > 0$ for $t \in [T_2, T_3]$, and also $\Psi(T_2) = \Psi(T_3) = 0$, and $\Psi(T_3 + \epsilon) < 0$ for $\epsilon > 0$ small enough. Hence

$$m(t) = m(T_2)\, e^{-\phi g_c (t - T_2)} + 1 - e^{-\phi g_c (t - T_2)} \text{ for } t \in [T_2, T_3). \tag{A8}$$

Combining this with (A6), we have

$$\begin{aligned}
\dot{\Psi}(T_1) &= (1-\theta)(\omega_c-\omega_o)(\rho+\phi g_c)m(T_1)^{\frac{1-\theta}{\phi}-1}\left[\widetilde{m}-m(T_1)\right] \\
\dot{\Psi}(T_2) &= (1-\theta)(\omega_c-\omega_o)(\rho+\phi g_c)m(T_2)^{\frac{1-\theta}{\phi}-1}\left[\widetilde{m}-m(T_2)\right] \\
\dot{\Psi}(T_3) &= (1-\theta)(\omega_c-\omega_o)(\rho+\phi g_c)m(T_3)^{\frac{1-\theta}{\phi}-1}\left[\widetilde{m}-m(T_3)\right].
\end{aligned}$$

Now, if $m(T_1) > g_o/g_c$, $m(t)$ is decreasing in t for $t \in [T_1, T_2]$ with $m(T_2) > g_o/g_c$. This implies $\widetilde{m} < g_o/g_c < m(T_2) < m(T_1)$, and thus $\dot{\Psi}(T_1) < 0$ and $\dot{\Psi}(T_2) < 0$, which contradicts $\Psi(T_1) = \Psi(T_2) = 0$.

If, instead, $m(T_1) < g_o/g_c$, then $m(t)$ is increasing in t for $t \in [T_1, T_2]$ with $m(T_2) < g_o/g_c$. In addition:

1. If $\widetilde{m} < m(T_1) < m(T_2)$, then $\dot{\Psi}(T_1) < 0$ and $\dot{\Psi}(T_2) < 0$, leading to a contradiction.

2. If $m(T_1) < m(T_2) < \widetilde{m} < g_o/g_c$, then $\dot{\Psi}(T_1) > 0$ and $\dot{\Psi}(T_2) > 0$, yielding another contradiction.

3. If $m(T_1) < \widetilde{m} < m(T_2) < g_o/g_c$, then $u(t) = 1$ and $\Psi(t) < 0$ for $t \in (T_2, T_3)$, and thus $\Psi(T_2) = \Psi(T_3) = 0$, and (A8) implies that $m(t)$ is increasing on $[T_2, T_3]$ and $m(T_2) < m(T_3)$. But this implies $\dot{\Psi}(T_2) < 0$ and $\dot{\Psi}(T_3) < 0$, which gives a contradiction combined with $\Psi(T_2) = \Psi(T_3) = 0$, establishing the lemma.

■

This lemma this implies that in the case where $\widetilde{m} < g_o/g_c$, the equilibrium will involve $u(t) = 1$ for all $t \leq T$ and $u(t) = 0$ for all $t > T$. Then from (A4) evaluated with $u(t) = 1$, we define $m(T) = \overline{m}$ to complete the proof of the proposition.

Finally, note that the phase diagrams in Figures 1, 2 and 3 can be derived straightforwardly from the expressions here, but to save space, we will derive them instead as a special case of the phase diagrams introduced in the proof of Proposition 6. ■

Proof of Proposition 3. The condition that $\widetilde{m} < 1$ can be written as

$$\widetilde{m} = \frac{1-\theta}{1-\Omega}\frac{g_c - \Omega g_o}{\rho + \phi g_c} < 1, \tag{A9}$$

where

$$\Omega \equiv \frac{\omega_c}{\omega_o} \equiv \frac{\left[q_1 A^{1-\theta} + (1-q_1)\right](1-\gamma)^{1-\theta}}{\left(q_1 A + (1-q_1)\right)^{1-\theta}} > 0,$$

with A given as in equation (9). First note that Ω does not depend on g_o and thus a higher g_o makes reduces the left-hand side of (A9) and makes this inequality more likely to hold, establishing the second claim. Next Ω also does not depend on ϕ and thus a higher ϕ makes (A9) more likely to hold as well, establishing the third claim.

To prove the first claim, observe that

$$\frac{\partial \widetilde{m}}{\partial \gamma} = \frac{1}{\rho + \phi g_c}\frac{\Omega\left(g_c - g_o\right)}{(1-\Omega)^2}\left[(1-\theta)\frac{\partial\Omega/\partial\gamma}{\Omega} + A(1-\theta)\frac{\partial\Omega/\partial A}{\Omega}\frac{\partial A/\partial\gamma}{A}\right].$$

We next show that the term in square brackets on the right-hand side is negative, establishing that $\widetilde{m}$ is decreasing in γ and thus the first claim in the proposition.

Lemma 2 *Suppose that Assumption 1 holds. Then*

$$(1-\theta)\frac{\partial\Omega/\partial\gamma}{\Omega}<0,\ (1-\theta)\frac{\partial\Omega/\partial A}{\Omega}<0,\ \textit{and}\ \frac{\partial A/\partial\gamma}{A}>0.$$

Proof. *Straightforward differentiation gives*

$$(1-\theta)\frac{\partial\Omega/\partial\gamma}{\Omega}=-\frac{(1-\theta)^2}{1-\gamma}<0,$$

$$(1-\theta)\frac{\partial\Omega/\partial A}{\Omega}=\frac{(1-\theta)^2 q_1(1-q_1)(A^{-\theta}-1)}{\left[q_1A^{1-\theta}+(1-q_1)\right]\left[q_1A+(1-q_1)\right]}<0,$$

and

$$\frac{\partial A/\partial\gamma}{A}=\frac{(1-\gamma)^{-\theta}\left[q_1-q_0\right]}{\left[1-q_0-(1-q_1)(1-\gamma)^{1-\theta}\right]\left[q_1(1-\gamma)^{1-\theta}-q_0\right]}>0.$$

■

The proof of the second claim is similar and is omitted.

We next turn to the third claim. We have

$$\frac{\partial\widetilde{m}/\partial\theta}{\widetilde{m}}=\frac{1}{1-\theta}\left[-1+\frac{\partial\Omega}{\partial\theta}\frac{(g_c-g_o)(1-\theta)}{(g_c-\Omega g_o)(1-\Omega)}\right],$$

$$\begin{aligned}\frac{\partial\Omega/\partial\gamma}{\Omega}&=-\log(1-\gamma)-\frac{q_1A^{1-\theta}\log A}{q_1A^{1-\theta}+(1-q_1)}-\log(q_1A+(1-q_1))\\&\quad+(1-\theta)\frac{\partial A}{\partial\theta}\frac{q_1(1-q_1)(A^{-\theta}-1)}{\left[q_1A^{1-\theta}+(1-q_1)\right]\left[q_1A+(1-q_1)\right]},\end{aligned}$$

and

$$\begin{aligned}(1-\theta)\frac{\partial A/\partial\theta}{A}&=\frac{(q_1-q_0)(1-\gamma)^{1-\theta}\log(1-\gamma)}{\left[(1-q_0)-(1-q_1)(1-\gamma)^{1-\theta}\right]\left[q_1(1-\gamma)^{1-\theta}-q_0\right]}\\&\quad+\frac{1}{(1-\theta)}\left[\log\left((1-q_0)-(1-q_1)(1-\gamma)^{1-\theta}\right)-\log\left(q_1(1-\gamma)^{1-\theta}-q_0\right)\right].\end{aligned}$$

Consider next a first-order Taylor expansion of A around $\gamma=0$, which gives $\log A\backsimeq\frac{\gamma}{q_1-q_0}$ and $(1-\theta)\frac{1}{A}\frac{\partial A}{\partial\theta}\backsimeq$constant$\cdot\gamma^2$. Therefore, ignoring second-order terms in γ, we have that around $\gamma=0$,

$$\frac{\partial\Omega/\partial\theta}{\Omega}\backsimeq\gamma-\frac{q_1\gamma}{q_1-q_0}-\log(1+\frac{q_1\gamma}{q_1-q_0})\backsimeq\gamma-\frac{2q_1\gamma}{q_1-q_0}=-\frac{(q_1+q_0)\gamma}{q_1-q_0}<0.$$

Therefore, there exists a value $\overline{\gamma}>0$ such that for $\gamma<\overline{\gamma}$, $\frac{\partial\Omega/\partial\gamma}{\Omega}<0$ and thus $\frac{\partial\widetilde{m}}{\partial\theta}<0$ when $\theta<1$, establishing the fourth claim. ■

Proof of Proposition 4. Consider the case where $n_\ell(0)=n_j(0)$. Then the result follows immediately from the proof of Proposition 2. In particular, recall that in part 1 or part 2 of that

proposition, the maximization problem of the social planner of country $j \neq \ell$ has a strictly higher value with cuddly reward structures (asymptotically) than with a cutthroat reward structure. If country j were to choose a cutthroat structure, it would have exactly the same welfare as country ℓ, and thus at $n_\ell(0) = n_j(0)$, country j has strictly higher welfare than country ℓ. Next by continuity, this is also true for $n_j(0) > 1 - \delta$ for δ sufficiently small and positive. ■

Proof of Proposition 5. To simplify notation, focus on the case with two countries, $\ell = 1$ and $j = 2$. Suppose that there exists another equilibrium than the one characterized in Proposition 2 in which country $\ell = 1$ still remains the leader throughout. This means that either country 1 adopts $u_1(t) = 1$ throughout and country 2 adopts a different strategy than in Proposition 2, or that country 1 adopts $u_1(t) = 0$ for some interval (with positive measure). But the first possibility is ruled out by the proof of Proposition 2, while the second one would be contradicted by the fact that under Assumption 3 and the hypothesis that it is always the leader, country 1 would obtain strictly lower welfare if $u_1(t) = 0$ for some (positive measure) interval.

Next consider the case where there exists another equilibrium in which country $\ell = 1$ ceases to be the technology leader at some point T. We now characterize the lowest possible value of T, which results when country $\ell = 1$ adopts $u_1(t) = 0$ and country 2 adopts $u_2(t) = 1$ until T. This is given as the solution to

$$\begin{gathered} N_1(T) = N_2(T), \text{ where} \\ N_1(t) = N_1(0)e^{g_o t} \\ N_2(t) = \left(N_2(0)^\phi + \tfrac{g_c}{g_o}\left(e^{\phi g_o t} - 1\right)(N_1(0))^\phi\right)^{\frac{1}{\phi}}, \end{gathered}$$

where these expressions follow directly from (16). Solving these three equations together gives us the date at which switches from country 1 to country 2 is

$$T = \frac{1}{\phi g_o} \ln\left(\frac{\frac{g_c}{g_o} - n(0)^\phi}{\frac{g_c}{g_o} - 1}\right). \tag{A10}$$

This equation implies that for $\phi \to 0$, we have $T \to \infty$.

Note next that welfare of country $\ell = 1$ at time $t = 0$ can then be written as

$$\mathcal{W}_1(0) = \left(1 - e^{-(\rho-(1-\theta)g_o)T}\right)\frac{\omega_o}{\rho - (1-\theta)g_o} + e^{-\rho T}\mathcal{W}_1(T),$$

where $\mathcal{W}_1(T)$ is the continuation utility of country 1 after T where the technology leadership shifts to country 2. Clearly $\mathcal{W}_1(T) \leq \omega_o/(\rho - (1-\theta)g_c)$. In view of this and of Assumption 3, when T is large enough (or equivalently $\phi < \bar{\phi}$), $\mathcal{W}_1(0)$ is strictly less than

$$\mathcal{W}_1^c(0) = \left(1 - e^{-(\rho-(1-\theta)g_c)T}\right)\frac{\omega_c}{\rho - (1-\theta)g_c} + e^{-(\rho-(1-\theta)g_c)T}\frac{\omega_c}{\rho - (1-\theta)g_c}$$

(where Assumption 3 ensures that the first term of $\mathcal{W}_1(0)$ is strictly smaller than the first term of $\mathcal{W}_1^c(0)$, and the hypothesis that T is large enough ensures that the second term of $\mathcal{W}_1(0)$ is small and thus cannot make up the difference). The hypothesis that $\phi < \bar{\phi}$ also ensures that

the same conclusion applies even after country 2 is arbitrarily close to its BGP level of relative technology, n^* (which it will approach as $u_1(t) = 1$ and $u_2(t) = 0$ for all t). Therefore, country 1 would be strictly better off choosing $u_1(t) = 1$ for all t. The same argument applies for any profile in which the switch happens at a later date than T, thus establishing the desired result.
■

Proof of Proposition 6. First consider country ℓ who is the technology leader at time $t = 0$. Given the selection rule implied by Assumption 4 that this country will remain the technology leader and (5), the world frontier technology is the same as this country's level of technology, i.e., $N(t) = N_\ell(t)$. Then (16) implies that

$$\frac{\dot{N}_\ell(t)}{N_\ell(t)} = g(u(t)).$$

Then Assumption 3 implies that country ℓ always prefers fully cutthroat incentives, i.e., $u_\ell(t) = 1$.

Next, let us focus on the problem of a follower country $j \neq \ell$ and drop the subscript j. This can be written as:

$$\max_{u(t)} \tfrac{1}{1-\lambda} \int_0^\infty e^{-\rho t}[(1-\theta)\omega(u(t))]^{\frac{1-\lambda}{1-\theta}} n(t)^{1-\lambda} N_\ell(t)^{1-\lambda} dt \quad \text{s.t:} \quad \begin{array}{r} \frac{\dot{n}(t)}{n(t)} = g(u(t))n(t)^{-\phi} - g_c, \\ \dot{N}_\ell(t) = g_c N_\ell(t) \\ n(0) = N(0)/N_\ell(0) \text{ given.} \end{array} \tag{A11}$$

To simplify the algebra, it is useful to consider the same change of variable as in the proof of Proposition 2, $m(t) \equiv (N(t)/N_\ell(t))^\phi$. Then (A11) can equivalently be written as

$$\max_{u(t)} \tfrac{1}{1-\lambda} \int_0^\infty e^{-\rho t}[(1-\theta)\omega(u(t))]^{\frac{1-\lambda}{1-\theta}} m(t)^{\frac{1-\lambda}{\phi}} N_\ell(t)^{1-\lambda} dt \quad \text{s.t:} \quad \begin{array}{r} \dot{m}(t) = \phi[g(u(t)) - g_c m(t)], \\ \dot{N}_\ell(t) = g_c N_\ell(t) \\ m(0) = (N(0)/N_\ell(0))^\phi \text{ given.} \end{array}$$

The current value Hamiltonean for this problem can be written as

$$H = \frac{1}{1-\lambda}\left[(1-\theta)\omega(u(t))\right]^{\frac{1-\lambda}{1-\theta}} m(t)^{\frac{1-\lambda}{\phi}} + \mu(t)\phi[g(u(t)) - g_c m(t)].$$

Consider the candidate solution given by the Maximum Principle, i.e., as a solution to the following equations:

$$\begin{aligned} \frac{\partial H}{\partial u} &= \Psi(t) \equiv (\omega_c - \omega_o)\left[(1-\theta)\omega(u(t))\right]^{\frac{\theta-\lambda}{1-\theta}} m(t)^{\frac{1-\lambda}{\phi}} + \mu(t)\phi(g_c - g_o) = 0 \text{ for } 0 \le u(t) \le 1 \\ \dot{m}(t) &= \phi[g(u(t)) - g_c m(t)] \qquad \text{(A12)} \\ \dot{\mu}(t) &= (\rho - (1-\lambda)g_c + \phi g_c)\mu(t) - \frac{1}{\phi}\left[(1-\theta)\omega(u(t))\right]^{\frac{1-\lambda}{1-\theta}} m(t)^{\frac{1-\lambda}{\phi}-1}, \end{aligned}$$

together with the transversality condition, which takes the form

$$\lim_{t\to\infty} e^{-(\rho-(1-\lambda)g_c)t}\mu(t) = 0.$$

If the first condition cannot be satisfied for interior $u(t)$, we have a corner solution at 0 or 1.

Let us next use another change of variable and define the following modified multiplier (co-state variable):

$$\kappa(t) \equiv \mu(t) m(t)^{\frac{\lambda-1}{\phi}}.$$

Then we can write, from the first line of (A12),

$$\Psi(t) = m(t)^{\frac{1-\lambda}{\phi}} \left[(\omega_c - \omega_o) \left[(1-\theta)\omega(u(t)) \right]^{\frac{\theta-\lambda}{1-\theta}} + \kappa(t)\phi(g_c - g_o) \right].$$

This immediately implies the following optimal control as a function of the multiplier κ (where we suppress the time argument from now on):

$$u(\kappa) = \begin{cases} 1 & \text{if } \kappa \geq \widetilde{\kappa}_1 \\ 0 & \text{if } \kappa \leq \widetilde{\kappa}_0 \\ u^*(\kappa) & \text{if } \widetilde{\kappa}_1 > \kappa > \widetilde{\kappa}_0 \end{cases}, \tag{A13}$$

where

$$\begin{aligned} \widetilde{\kappa}_1 &\equiv \frac{(\omega_c - \omega_o) \left[(1-\theta)\omega_c \right]^{\frac{\theta-\lambda}{1-\theta}}}{\phi(g_c - g_o)}, \\ \widetilde{\kappa}_o &\equiv \frac{(\omega_c - \omega_o) \left[(1-\theta)\omega_o \right]^{\frac{\theta-\lambda}{1-\theta}}}{\phi(g_c - g_o)}, \text{ and} \\ u^*(\kappa) &\equiv \frac{\omega_o}{\omega_o - \omega_c} - \frac{1}{1-\theta} \frac{1}{\omega_o - \omega_c} \left(\frac{\phi(g_c - g_o)}{\omega_o - \omega_c} \kappa \right)^{\frac{1-\theta}{\theta-\lambda}}. \end{aligned} \tag{A14}$$

Then substituting for (A13) into the differential equations for $\dot{m}$ and $\dot{\mu}$, and using the definition of κ, we obtain the law of motion of the system consisting of (m, κ) as

$$\begin{aligned} \dot{\kappa} &= (\rho - (1-\lambda)g_c + \phi g_c)\kappa - \frac{1}{\phi}((1-\theta)\omega(u(\kappa)))^{\frac{1-\lambda}{1-\theta}} \frac{1}{m} - (1-\lambda)(g_o + (g_c - g_o)u(\kappa)) \frac{\kappa}{m} \\ \dot{m} &= \phi(g_o + (g_c - g_o)u(\kappa) - g_c m). \end{aligned} \tag{A15}$$

with $u(\kappa) = 0$ when $\kappa \leq \widetilde{\kappa}_0$, $u(\kappa) = 1$ when $\kappa \geq \widetilde{\kappa}_1$ and $u(\kappa) = u^*(\kappa) \in (0,1)$ for $\widetilde{\kappa}_1 > \kappa > \widetilde{\kappa}_0$. In particular in the last regime where $u(\kappa) = u^*(\kappa)$, after substitution of (A14) the system can be rewritten as:

$$\begin{aligned} \dot{\kappa} &= (\rho + \phi g_c)\kappa - \frac{\lambda - \theta}{1-\theta} \left[\frac{g_c - g_o}{\omega_o - \omega_c} \right]^{\frac{1-\lambda}{\theta-\lambda}} \phi^{\frac{1-\theta}{\theta-\lambda}} \frac{\kappa^{\frac{1-\lambda}{\theta-\lambda}}}{m} - (1-\lambda) \frac{\omega_o(g_c - g_o) + g_o(\omega_o - \omega_c)}{\omega_o - \omega_c} \frac{\kappa}{m} \\ \dot{m} &= \phi\left(\frac{\omega_o(g_c - g_o) + (\omega_o - \omega_c) g_o}{\omega_o - \omega_c} - \frac{1}{1-\theta} \left(\frac{g_c - g_o}{\omega_o - \omega_c} \right)^{\frac{1-\lambda}{\theta-\lambda}} \phi^{\frac{1-\theta}{\theta-\lambda}} \kappa^{\frac{1-\theta}{\theta-\lambda}} - g_c m \right) \end{aligned} \tag{A16}$$

More specifically, in a phase diagram in the (m, κ) space, the $\dot{m} = 0$ locus is given as shown in Figure 7, with the curve linking $m = g_o/g_c$ (which applies when $\kappa \leq \widetilde{\kappa}_0$ and thus $u(\kappa) = 0$) to $m = 1$ (which applies when $\kappa \geq \widetilde{\kappa}_1$ and thus $u(\kappa) = 1$) given from (A15) as $m = g_o + (g_c - g_o)u(\kappa)/g_c$, which defines an increasing relationship. Clearly, $\dot{m} > 0$ above this curve and $\dot{m} < 0$ below this curve.

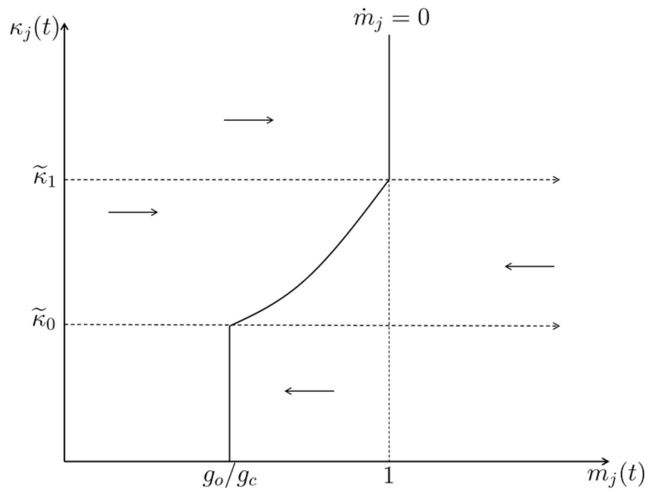

Figure 7: Construction of phase diagram for Proposition 7. $\dot{m} = 0$ locus.

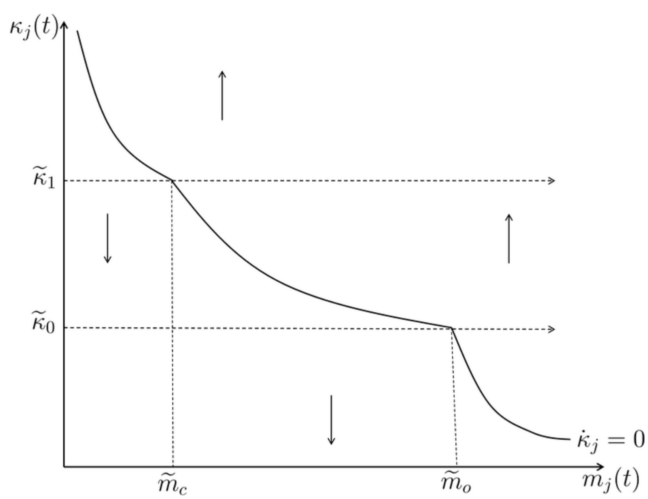

Figure 8: Construction of phase diagram for Proposition 7. $\dot{\kappa} = 0$ locus.

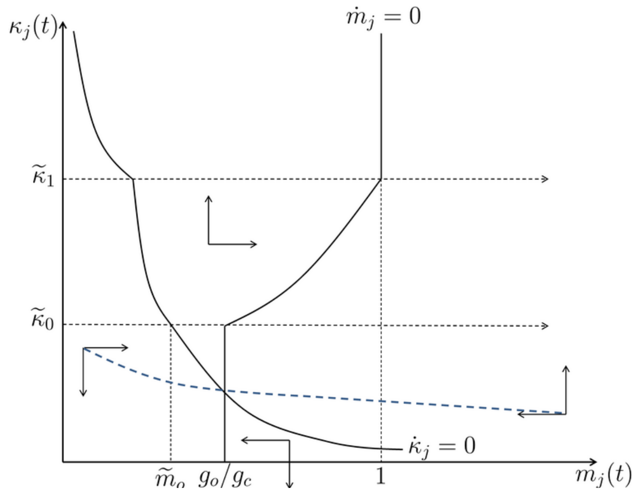

Figure 9: Phase diagram for Proposition 7. Case $g_o/g_c > \widetilde{m}_o > \widetilde{m}_c$.

The $\dot{\kappa} = 0$ locus in the (m, κ) space is derived similarly from (A15), and thus shown in Figure 8, with $\dot{\kappa} > 0$ above this curve and $\dot{\kappa} < 0$ below this curve.

Once we put these two curves together, equilibrium dynamics are determined by the point of intersection. Figure 9 corresponds to the case where $g_o/g_c > \widetilde{m}_o > \widetilde{m}_c$. The two curves for $\dot{m} = 0$ and $\dot{\kappa} = 0$ intersect where $m = g_o/g_c$. The laws of motion of (m, κ) we have just derived imply the existence of a unique stable arm as shown in the figure. The dynamics in part 1 of the proposition that follow from this figure. In particular, asymptotically (for m close enough to g_o/g_c) the follower necessarily chooses $u = 0$, and this is preceded by regions in which $u \in (0,1)$ and $u = 1$.

Figure 10 corresponds to the case where $g_o/g_c < \widetilde{m}_o$ and $\widetilde{m}_c < 1$, where the intersection of the curves for $\dot{m} = 0$ and $\dot{\kappa} = 0$ takes place where $\dot{m} = 0$ is downward sloping. The unique stable arm's location then follows again straightforwardly from the laws of motion derived above, and these dynamics established the results in part 2.

Finally, Figure 11 corresponds case where $\widetilde{m}_o > \widetilde{m}_c > 1$, leading to the intersection of the terms for $\dot{m} = 0$ and $\dot{\kappa} = 0$ at $m = 1$. The shape of the unique stable arm now implies that $u = 1$ throughout.

We also note that when $\theta = \lambda$, $\widetilde{\kappa}_0 = \widetilde{\kappa}_1$, and the phase diagrams simplify to Figures 1, 2 and 3 in the text.

To complete the proof, we show that the Mangasarian sufficiency condition holds, so that the dynamics characterized here give the unique global optimal. Note that

$$
\begin{aligned}
\frac{\partial^2 H}{\partial u^2} &= (\theta - \lambda)\left[\omega_c - \omega_o\right]^2 \left[(1-\theta)\omega(u)\right]^{\frac{\theta-\lambda}{1-\theta}-1} m^{\frac{1-\lambda}{\phi}} \\
\frac{\partial^2 H}{\partial u \partial m} &= \frac{1-\lambda}{\phi}\left[\omega_c - \omega_o\right]\left[(1-\theta)\omega(u)\right]^{\frac{\theta-\lambda}{1-\theta}} m^{\frac{1-\lambda}{\phi}-1} \\
\frac{\partial^2 H}{\partial m^2} &= \frac{1-\lambda-\phi}{\phi^2}\left[(1-\theta)\omega(u)\right]^{\frac{1-\lambda}{1-\theta}} m^{\frac{1-\lambda}{\phi}-2}
\end{aligned}
$$

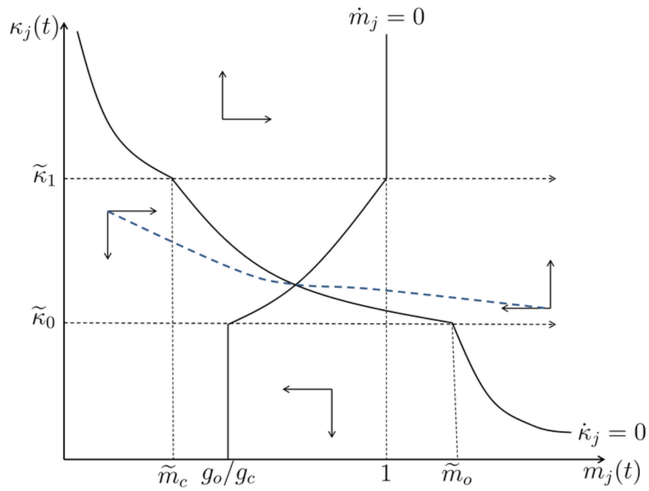

Figure 10: Phase diagram for Proposition 7. Case $g_o/g_c < \widetilde{m}_o$ and $\widetilde{m}_c < 1$.

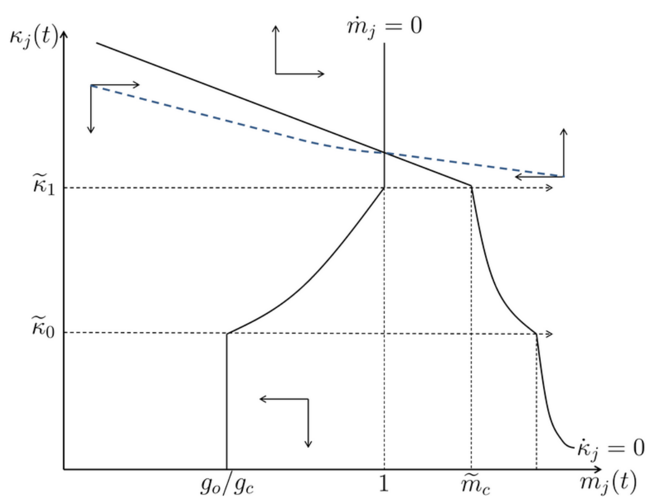

Figure 11: Phase diagram for Proposition 7. Case $\widetilde{m}_o > \widetilde{m}_c > 1$.

where recall that $\frac{\omega_c}{\omega_o} \lesseqgtr 1$ when $\theta \lesseqgtr 1$. Mangasarian sufficiency condition, the joint concavity of H, is equivalent to $\frac{\partial^2 H}{\partial u^2} \leq 0$, $\frac{\partial^2 H}{\partial n^2} \leq 0$ and $\frac{\partial^2 H}{\partial u^2}\frac{\partial^2 H}{\partial n^2} - \left(\frac{\partial^2 H}{\partial u \partial n}\right)^2 \geq 0$. The first of conditions are satisfied. Some algebra establishes that the third one is also satisfied when $\phi > \frac{(1-\theta)(1-\lambda)}{(\lambda-\theta)}$. ■

Proof of Proposition 7. We will prove that under the hypotheses of the proposition, there does not exist a symmetric equilibrium.

Suppose first that all countries choose a cuddly reward structure for all $t \geq 0$. Then the world economy converges to a Balanced Growth Path (BGP) where every country has the same level of income, $N_j(t)/(1-\beta) = N(t)/(1-\beta)$, and grows at the same rate, which from (6) is equal to $\dot{N}(t)/N(t) = g_o$. The time t welfare of country j in this equilibrium can be written as

$$\mathcal{W}_j^o(t) = \int_t^\infty e^{-\delta(\tau - t)} \omega_o \left(\frac{N_j(\tau)}{N(\tau)}\right)^{1-\theta} N(\tau)^{1-\theta} d\tau,$$

which implies that for any $\epsilon > 0$, there exists T_1 such that for all $t > T_1$, we are close enough to the BGP equilibrium in the sense that $1 - \epsilon < \frac{N_j(t)}{N(t)} < 1 + \epsilon$, $\dot{N}/N < g_o + \epsilon$, and

$$\mathcal{W}_j^o(t) < \frac{\omega_o N(t)^{1-\theta}(1+\epsilon)^{1-\theta}}{\rho - (1-\theta)(g_o + \epsilon)}$$

Consider now a deviation of one country k to a cutthroat reward structure at all times $t > T_1$. Denote by $\widehat{N}_j(t)$, the new growth path of country j and by $\widehat{N}(t)$ the growth path to the world technology frontier. The world economy converges again to a new BGP with growth rate $\widehat{g}$. This BGP growth rate can be written as

$$\widehat{g} = \frac{1}{J^{\frac{\sigma}{\sigma-1}\phi}} \left[(J-1) g_o^{\frac{1}{\phi}\frac{\sigma-1}{\sigma}} + g_c^{\frac{1}{\phi}\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}\phi} > g_o.$$

After this deviation, we have $\widehat{N}_k(t) > N_k(t)$ and $\widehat{N}_k(t) > N(t)$ for all $t > T_1$. Then for $\epsilon_1 > 0$, there exists $T_1' > T_1$ such that for all $t > T_1'$, $\dot{\widehat{N}}_k/\widehat{N}_k \geq \widehat{g} - \epsilon_1$, and welfare of country k satisfies

$$\begin{aligned}
\mathcal{W}_k^c(T_1) &= \int_{T_1}^\infty e^{-\rho(t-T_1)} \omega_c \widehat{N}_k(t)^{1-\theta} dt \\
&= \int_{T_1}^{T_1'} e^{-\rho(t-T_1)} \omega_c \widehat{N}_k(t)^{1-\theta} dt + e^{-\rho(T_1'-T_1)} \int_{T_1'}^\infty e^{-\rho(t-T_1')} \omega_c \widehat{N}_k(t)^{1-\theta} dt \\
&> e^{-\rho(T_1'-T_1)} \omega_c \frac{\widehat{N}_k(T_1')^{1-\theta}}{\rho - (1-\theta)(\widehat{g} - \epsilon_1)}.
\end{aligned}$$

Now using the fact that $\widehat{N}_k(T_1') \geq N_k(T_1') \geq e^{g_o(T_1'-T_1)} N_k(T_1)$, a sufficient condition for the deviation for country k to be profitable is

$$\begin{aligned}
e^{-(\rho-(1-\theta)g_o)(T_1'-T_1)} \omega_c \frac{N_k(T_1)^{1-\theta}}{\rho - (1-\theta)(\widehat{g} - \epsilon_1)} &> \omega_o \frac{N_k(T_1)^{1-\theta}(1+\epsilon)^{1-\theta}}{\rho - (1-\theta)(g_o + \epsilon)} \\
&> \mathcal{W}_k^o(T_1) = \int_{T_1}^\infty e^{-(\rho-(1-\theta)g_o)(t-T_1)} \omega_o N_k(t)^{1-\theta} dt.
\end{aligned}$$

Rearranging terms, this can be written as

$$\left(\frac{\omega_c}{\omega_o}\right)^{\frac{1}{1-\theta}} > (1+\epsilon)\, e^{\frac{\rho-(1-\theta)g_o}{1-\theta}(T_1'-T_1)} \left(\frac{\rho-(1-\theta)(\widehat{g}-\epsilon_1)}{\rho-(1-\theta)(g_0+\epsilon)}\right)^{\frac{1}{1-\theta}}. \tag{A17}$$

Next suppose that all countries adopt a cutthroat reward structure for all $t \geq 0$. In this case, the world economy converges to a BGP where every country has the same level of income and grows at the same rate, which from (6) is equal to $\dot{N}(t)/N_j(t) = g_c$. With a similar reasoning, for $\epsilon > 0$, there exists T_2 such that for all j and $t > T_2$, $1-\epsilon < N_j(t)/N(t) < 1+\epsilon$ and $\dot{N}/N < g_c + \epsilon$. Thus

$$\mathcal{W}_j^c(t) < \frac{\omega_c N(t)^{1-\theta}(1+\epsilon)^{1-\theta}}{\rho-(1-\theta)(g_c+\epsilon)}$$

Consider now a deviation of one country k to a cuddly reward structure at all time $t > T_2$ while all other countries $j \neq k$ stay with cutthroat reward structures throughout. Denote the path of technology of country j after this deviation by $\widetilde{N}_j(t)$, and the path of world technology frontier by $\widetilde{N}(t)$. Clearly, $\dot{\widetilde{N}}(t)/\widetilde{N}(t) = \widetilde{g} < g_c$, and moreover $\widetilde{N}_k(t) \leq N_k(t)$ for all $t > T_2$. Let us also note that

$$\widetilde{g} = \frac{\dot{\widetilde{N}}(t)}{\widetilde{N}(t)} = \frac{1}{J^{\frac{\sigma}{\sigma-1}\phi}}\left[(J-1)g_c^{\frac{1}{\phi}\frac{\sigma-1}{\sigma}} + g_o^{\frac{1}{\phi}\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}\phi} > g_o.$$

Now, again fixing $\epsilon_2 > 0$, there exists $T_2' > T_2$ such that for all $t > T_2'$, $\dot{\widetilde{N}}_k/\widetilde{N}_k \geq \widetilde{g} - \epsilon_2$, and the welfare of country k satisfies

$$\begin{aligned}
\mathcal{W}_k^o(T_2) &= \int_{T_2}^{\infty} e^{-\rho(t-T)}\omega_o \widetilde{N}_k(t)^{1-\theta}dt \\
&= \int_{T_2}^{T_2'} e^{-\rho(t-T_2)}\omega_o \widetilde{N}_k(t)^{1-\theta}dt + e^{-\rho(T_2'-T_2)}\int_{T_2'}^{\infty} e^{-\rho(t-T"_2)}\omega_o \widetilde{N}_k(t)^{1-\theta}dt \\
&> \omega_o N_k(T_2)^{1-\theta}\int_{T_2}^{T_2'} e^{-\rho(t-T)}e^{(1-\theta)g_o(t-T)}dt + e^{-\rho(T_2'-T_2)}\omega_o \frac{N_k(T_2)^{1-\theta}e^{(1-\theta)g_0(T_2'-T_2)}}{\rho-(1-\theta)(\widetilde{g}-\epsilon_2)} \\
&> \omega_o N_k(T_2)^{1-\theta}\frac{1-e^{-(\rho-(1-\theta)g_o)(T_2'-T_2)}}{\rho-(1-\theta)\,g_o} + e^{-(\rho-(1-\theta)g_o)(T_2'-T_2)}\omega_o\frac{N_k(T_2)^{1-\theta}}{\rho-(1-\theta)(\widetilde{g}-\epsilon_2)},
\end{aligned}$$

where the second line uses the fact $\widetilde{N}_k(t) > N_k(T_2)e^{g_0(t-T_2)}$. Then a sufficient condition for the deviation to the cuddly reward structure for country k to be profitable is

$$e^{-(\rho-(1-\theta)g_o)(T_2'-T_2)}\omega_o\frac{N_k(T_2)^{1-\theta}}{\rho-(1-\theta)(\widetilde{g}-\epsilon_2)} > \omega_c\frac{N(T_2)^{1-\theta}(1+\epsilon)^{1-\theta}}{\rho-(1-\theta)(g_c+\epsilon)}.$$

Since $N_k(T_2) > N(T_2)(1-\epsilon)$, this sufficient condition can be rewritten as

$$\frac{1-\epsilon}{1+\epsilon}\left(\frac{\rho-(1-\theta)(g_c+\epsilon))}{\rho-(1-\theta)(\widetilde{g}-\epsilon_2)}\right)^{\frac{1}{1-\theta}} e^{-\frac{\rho-(1-\theta)g_o}{1-\theta}(T_2'-T_2)} > \left(\frac{\omega_c}{\omega_o}\right)^{\frac{1}{1-\theta}}. \tag{A18}$$

Thus combining (A17) and (A18), we obtain that the following is a sufficient condition for an asymmetric equilibrium not to exist after some time $T = \max\{T_1, T_2\}$:

$$\frac{1-\epsilon}{1+\epsilon}\left(\frac{\rho-(1-\theta)(g_c+\epsilon))}{\rho-(1-\theta)(\widetilde{g}-\epsilon_2)}\right)^{\frac{1}{1-\theta}} e^{-\frac{\rho-(1-\theta)g_o}{1-\theta}(T_2'-T_2)} > \left(\frac{\omega_c}{\omega_o}\right)^{\frac{1}{1-\theta}} > (1+\epsilon)e^{\frac{\rho-(1-\theta)g_o}{1-\theta}(T_1'-T_1)}\left(\frac{\rho-(1-\theta)(\widehat{g}-\epsilon_1)}{\rho-(1-\theta)(g_o+\epsilon)}\right)^{\frac{1}{1-\theta}} \tag{A19}$$

Now note that as $\sigma \uparrow 0$ in (6), $\widehat{g} \longrightarrow g_c$ and $\widetilde{g} \longrightarrow g_c$. Therefore, for $\epsilon' > 0$, there exists $\bar{\sigma} < 0$ such that for $\sigma > \bar{\sigma}$, $\widehat{g} - \epsilon' < g_c$ and $\widetilde{g} - \epsilon' < g_c$. Thus choosing ϵ, ϵ_1, ϵ_2, and ϵ' sufficiently small, the following is also a sufficient condition:

$$e^{-\frac{\rho-(1-\theta)g_o}{1-\theta}(T_2'-T_2)} > \left(\frac{\omega_c}{\omega_o}\right)^{\frac{1}{1-\theta}} > e^{\frac{\rho-(1-\theta)g_o}{1-\theta}(T_1'-T_1)}\left(\frac{\rho-(1-\theta)g_c}{\rho-(1-\theta)g_o}\right)^{\frac{1}{1-\theta}}. \tag{A20}$$

Consider first the case $\theta < 1$. Choosing ρ sufficiently close to $(1-\theta)\,g_c$ and defining $\bar{T} \equiv \max\{T_1' - T_1, T_2' - T_2\}$, a further sufficient condition is obtained as

$$e^{-(g_c-g_o)\bar{T}} > \left(\frac{\omega_c}{\omega_o}\right)^{\frac{1}{1-\theta}} > e^{(g_c-g_o)\bar{T}}\left(\frac{\rho-(1-\theta)g_c}{\rho-(1-\theta)g_o}\right)^{\frac{1}{1-\theta}}. \tag{A21}$$

For given choices of ϵ and ϵ_1, $\bar{T}$ is fixed. Hence there exists $\bar{\rho} > (1-\theta)\,g_c$ such that for $(1-\theta)\,g_c < \rho < \bar{\rho}$, the right-hand side term inequality is close to zero and the left-hand term is given by some positive number. Next recall that

$$\left(\frac{\omega_c}{\omega_o}\right)^{\frac{1}{1-\theta}} = \frac{\left[q_1 A^{1-\theta} + (1-q_1)\right]^{\frac{1}{1-\theta}}(1-\gamma)}{q_1 A + (1-q_1)}.$$

Denote $\widetilde{\gamma} = 1 - \left(\frac{q_0}{q_1}\right)^{1/(1-\theta)}$, then $\left(\frac{\omega_c}{\omega_o}\right)^{\frac{1}{1-\theta}} \rightarrow \frac{q_0^{1/(1-\theta)}}{q_1}$ as $\gamma \rightarrow \widetilde{\gamma}$. This in turn $\rightarrow 0$ as $\theta \rightarrow 1$. Hence there exists $\bar{\theta} < 1$ and $\bar{\gamma} < 1$ such that for $\theta > \bar{\theta}$ and $\gamma > \bar{\gamma}$, $\left(\frac{\omega_c}{\omega_o}\right)^{\frac{1}{1-\theta}}$ is sandwiched between these two terms, ensuring that (A21) is satisfied and a symmetric equilibrium does not exist.

When $\theta > 1$, (A20) can be rewritten as

$$e^{\frac{\rho+(\theta-1)g_o}{\theta-1}(T_2'-T_2)} > \left(\frac{\omega_o}{\omega_c}\right)^{\frac{1}{\theta-1}} > e^{-\frac{\rho+(\theta-1)g_o}{\theta-1}(T_1'-T_1)}\left(\frac{\rho+(\theta-1)g_o}{\rho+(\theta-1)g_c}\right)^{\frac{1}{\theta-1}}.$$

Again defining $\bar{T} \equiv \max\{T_1' - T_1, T_2' - T_2\}$ and taking $\rho < \bar{\rho}$ for $\bar{\rho}$ small enough, a further sufficient condition writes as

$$e^{g_o\bar{T}} > \left(\frac{\omega_o}{\omega_c}\right)^{\frac{1}{\theta-1}} > e^{-g_o\bar{T}}\left(\frac{g_o}{g_c}\right)^{\frac{1}{\theta-1}}.$$

Now the first inequality is satisfied as $e^{g_o\bar{T}} > 1 > \left(\frac{\omega_o}{\omega_c}\right)^{\frac{1}{\theta-1}}$ as $\theta > 1$ and $\omega_c < \omega_o < 0$. The second inequality is satisfied when $\frac{\omega_o}{\omega_c} > \frac{g_o}{g_c}$, ensuring therefore again that (A21) is satisfied and a symmetric equilibrium does not exist.

Finally, when these conditions are satisfied, a similar analysis to that in the proof of Proposition 2 implies that the equilibrium will take the form where after some T, subset of countries choose a cuddly reward structure and the remaindered choose a cutthroat reward structure. ■

Proof of Proposition 8. The first part follows by noting that when Assumption 3 holds, one of the two countries will asymptotically use cutthroat incentives. The second part follows immediately from the observation that the domestic political constraint prevents country j from adopting cutthroat incentives, and thus eliminates one of these two equilibria. The welfare conclusion follows directly from Proposition 4. ■

References

Acemoglu, Daron (2003) "Cross-Country Inequality Trends," *Economic Journal*, 113, F121-149.

Acemoglu, Daron (2009) *Introduction to Modern Economic Growth*, Princeton University Press.

Acemoglu, Daron, Philippe Aghion and Fabrizio Zilibotti (2006) "Distance to Frontier, Selection, and Economic Growth," *Journal of the European Economic Association*, MIT Press, vol. 4(1), pages 37-74, 03.

Acemoglu, Daron, Ufuk Akcigit and Murat Alp Celik (2014) "Young, Restless and Creative: Openness to Disruption and Creativity of Innovations" NBER working paper No. 19894.

Acemoglu, Daron and Joshua Linn (2004) "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Sector" *Quarterly Journal of Economics*, 119:3, 1049-1090.

Acemoglu, Daron and Jorn-Steffen Pischke (1998) "Why Do Firms Train? Theory and Evidence" *Quarterly Journal of Economics*, 113:1, 78-118.

Acemoglu, Daron and Jaume Ventura (2001) "The World Income Distribution" *Quarterly Journal of Economics*, 117:2, 659-694.

Akkermans, Dirk, Carolina Castaldi, and Bart Los (2009) "Do 'liberal market economies' really innovate more radically than 'coordinated market economies'? Hall and Soskice reconsidered," *Research Policy*, 38, 181-91.

Alesina, Alberto, Edward L. Glaeser and Bruce Sacerdote (2001) "Why Doesn't the United States Have a European-Style Welfare State?" *Brookings Paper on Economics Activity*, Fall 2001, 187-278.

Alesina, Alberto, Edward L. Glaeser and Bruce Sacerdote (2005) "Work and Leisure in the U.S. and Europe: Why So Different?" *NBER Macroeconomic Annual*, 1-64.

Alesina, Alberto and with Eliana La Ferrara (2005) "Preferences for Redistribution in the Land of Opportunities," *Journal of Public Economics*, 89, 897-931.

Algan, Yann, Pierre Cahuc and Marc Sangnier (2011) "Efficient and Inefficient Welfare States," https://sites.google.com/site/pierrecahuc/ unpublished_papers.

Aoki Masahiko (2001) *Towards a Compact of Institutional Analysis*, Cambridge, MIT Press.

Atkinson, Anthony B., Thomas Piketty and Emmanuel Saez (2007) "Top Incomes in the Long Run of History" *Journal of Economic Literature*, 49:1, 3-71.

Baldwin, Peter (1992) *The Politics of Social Solidarity: Class Bases of the European Welfare State, 1875-1975*, New York: Cambridge University Press.

Bénabou, Roland J.M. (2000) "Unequal Societies: Income Distribution and the Social Contract," *American Economic Review*, 90, 96-129.

Bénabou, Roland J.M. (2006) "Inequality, Technology, and the Social Contract," in Philippe Aghion and Steven N. Durlauf eds. *Handbook of Economic Growth*, Amsterdam: North-Holland, 1595-1638.

Bénabou, Roland J.M. and Efe Ok (2001) "Social Mobility and the Demand for Redistribution: The POUM Hypothesis," *Quarterly Journal of Economics*, 116:2, 447-487.

Bénabou, Roland J.M. and Jean Tirole (2006) "Belief in a Just World and Redistributive Politics," *Quarterly Journal of Economics*, 121:2, 699-746.

Bhosle, Monali and Rajesh Balakrishnan (2007) "Drug Reimportation Practices in the United States," *Therapeutics and Clinical Risk Management*, 3(1), 41-46.

Blanchard, Olivier (2004) "The Economic Future of Europe," *Journal of Economic Perspectives*, 18:4, 3-26.

Bottazzi, Laura and Giovanni Peri (2003) "Innovation and spillovers in regions: Evidence from European patent data," *European Economic Review*, 47:4, 687-710.

Brezis, Elise S., Paul R. Krugman and Daniel Tsiddon (1993) "Leapfrogging in International Competition: A Theory of Cycles in National Technological," *American Economic Review*, 83(5), 1211-1219.

Cameron, David R. (1978) "The Expansion of the Public Economy: A Comparative Analysis," *American Political Science Review*, 72, 1243–61.

Cardoso, Fernando H. and Enzi Faletto Enzo (1979) *Dependency and Development in Latin America*, Berkeley: University of California Press.

Chatterjee, Arpita (2010) "Why Do Similar Countries Choose Different Policies? Endogenous Comparative Advantage and Welfare Gains" Princeton University mimeo.

Choudhury, Niteesh and Allan Destsky (2005) "A Perspective on US Drug Reimportation," *The Journal of the American Medical Association*, 293(3), 358-362.

Coe, David and Elhanan Helpman (1995) "International R&D spillovers," *European Economic Review,* 39(5), 859-887.

Crouch, Colin (2009) "Typologies of Capitalism," in Bob Hancké ed. *Debating Varieties of Capitalism: A Reader*, New York: Oxford University Press.

Danzon, Patricia, (1997) "Price Discrimination for Pharmaceuticals: Welfare Effects in the US and the EU, International *Journal of the Economics of Business*, 4(3), 301-322.

Danzon, Patricia and Adrian Towse (2003) "Differential Pricing for Pharmaceuticals:

Reconciling Access, R&D and Patents" *International Journal of Health Care Finance and Economics*, 3, 183-205.

Danzon, Patricia and Richard Wang (2005) "The Impact of Price Regulation on the Launch Delay of New Drugs: Evidence from Twenty-five Major Markets in the 1990s" *Health Economics*, 14(3), 269-292.

DiMasi, Joseph, Ronald Hansen, Henry Grabowski (2003) "The Price of Innovation: New Estimates of Drug Development Costs" *Journal of Health Economics*, 22, 151-185.

Epstein, Larry G. and Stanley E. Zin (1989) "Substitution, Risk Aversion and Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57:4, 937-969.

Esping-Anderson, Gösta (1990) *The Three Worlds of Welfare Capitalism*, Princeton: Princeton University Press.

Finkelstein, Amy (2004) "Static and Dynamic Effects of Health Care Policy: Evidence from the Vaccine Industry" *Quarterly Journal of Economics*, 119:2, 527-564.

Friedman, Gerald (2010) "Labor Unions in the United States," http://eh.net/encyclopedia/article/friedman.unions.us

Gerschenkron, Alexander (1962) *Economic backwardness in historical perspective, a book of essays*, Belknap Press of Harvard University Press, Cambridge, Massachusetts.

Giacoto, Carmelo, Rexford Santerre and John Vernon (2005) "Drug Prices and Research and Development Investment Behavior in the Pharmaceutical Industry," *Journal of Law and Economics*, 48(1): 195-214.

Golec, Joseph and John Vernon (2006) "European Pharmaceutical Price Regulation, Firm Profitability, and R&D Spending," NBER Working Paper, No. 12676

Griffith, Rachel, Stephen Redding and John Van Reenen (2003) "R&D and Absorptive Capacity: Theory and Empirical Evidence," *Scandinavian Journal of Economics*, Wiley Blackwell, vol. 105(1), pages 99-118, 03.

Gourevitch, Peter (1986) *Politics in Hard Times*, Ithaca: Cornell University Press.

Hall, Peter and David Soskice (2001) *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*, Oxford University Press, USA.

Hölmstrom, Bengt (1979) "Moral Hazard and Observability," *The Bell Journal of Economics*, Vol. 10, No. 1 , pp. 74-91.

Howitt, Peter (2000) "Endogenous Growth and Cross-Country Income Differences," *American Economic Review*, 90:4, 829-846.

Keller, Wolfgang (2001) "Knowledge Spillovers at the World's Technology Frontier," CEPR Discussion Papers 2815.

Kerr, William (2008) "Ethnic Scientific Communities and International Technology Diffusion," *Review of Economics and Statistics*, 90:3, 518-537.

Kneller, Robert (2010) "The Importance of New Companies for Drug Discovery: Origins of a Decade of New Drugs" *Nature Reviews Drug Discovery*, 9(11), 867-882.

Krueger, Dirk and Krishna Kumar (2004) "US-Europe differences in technology-driven growth: quantifying the role of education," *Journal of Monetary Economics*, 51:1, 161-190.

Krugman, Paul and Anthony Venables (1996) "Integration, specialization, and adjustment," *European Economic Review*, 40, 959-967.

Landier, Agustin (2005) "Entrepreneurship and the Stigma of Failure," mimeo.

Levchenko, Andrei (2007) "Institutional Quality and International Trade" *Review of Economic Studies*, 74, 791-819.

Lipset, Seymour M. (1959) "Some Social Requisites of Democracy: Economic Development and Political Legitimacy," *American Political Science Review,* 53, 69-105.

Lundberg, Urban and Klas Åmark (2001) "Social Rights and Social Security: The Swedish Welfare State, 1900–2000," Scandinavian *Journal of History*, 26:3, 157-176.

Mansfield, Edwin (1986) "Patents and Innovation: An Empirical Study," *Management Science*, 32(2), 173-181.

Marmor, Theodore R. and Jacob S. Hacker (2005) "Medicare Reform and Social Insurance: The Clashes of 2003 and Their Potential Fallout," *Yale Journal of Health Policy, Law, and Ethics*, 5(1), 475-490.

Matsuyama, Kiminori (2002) "Explaining Diversity: Symmetry-Breaking in Complementarity Games," *American Economic Review,* 92, 241-246.

Matsuyama, Kiminori (2004) "Financial Market Globalization, Symmetry-Breaking, and Endogenous Inequality of Nations," *Econometrica,* 72, 853-884.

Matsuyama, Kiminori (2005) "Symmetry-Breaking," in L. Blume and S. Durlauf, eds., the New Palgrave Dictionary of Economics, 2nd Edition, Palgrave Macmillan.

Matsuyama, Kiminori (2007) "Aggregate Implications of Credit Market Imperfections," in D. Acemoglu, K. Rogoff, and M. Woodford, eds., *NBER Macroeconomics Annual 2007.*

Matsuyama, Kiminori (2013) "Endogenous Ranking and Equilibrium Lorenz Curve Across (ex ante) Identical Countries," *Econometrica,* 81, 2009-2031.

Milesi-Ferretti, Gian Maria, Roberto Perotti and Massimo Rostagno (2002) "Electoral Systems and Public Spending," *Quarterly Journal of Economics*,

Moene, Karl Ove and Michael Wallerstein (1997) "Pay Inequality," *Journal of Labor Economics*, 15(3), 403-430.

Moffit, Robert (2013) "Medicare Drugs: Why Congress Should Reject Government Price Fixing" Heritage Foundation Issue Brief #3880, Available online: http://www.heritage.org/research/reports/2013/03/medicare-drugs-why-congress-should-reject-government-price-fixing

Nelson, Richard and Edmond Phelps (1966) "Investment in Humans, Technological Diffusion, and Economic Growth," *American Economic Review*, 56:1/2 69-75.

Persson, Torsten and Guido Tabellini (2003) *The Economic Effect of Constitutions*, Cambridge: MIT Press.

Quon, Bradley, Rafael Firszt, and Mark Eisenberg (2005) "A Comparison of Brand-

name Drug Prices between Canadian-Based Internet Pharmacies and Major US Drug Chain Pharmacies," *Annals of Internal Medicine*, 143(6), 397-403.

Rodrik, Dani (2008) "Second-Best Institutions," *American Economic Review*, 98(2), 100-104.

Romer, Paul (1990) "Endogenous Technological Change," *Journal of Political Economy*, 98:5, pp. S71-S102.

Saint-Paul, Gilles and Thierry Verdier (1993) "Education, democracy and growth," *Journal of Development Economics*, 42, 399-407.

Scherer, Frederic M. (1993) "Pricing, Profits, and Technological Progress in the Pharmaceutical Industry," *Journal of Economic Perspectives*, 7(3), 97-115.

Scherer, Frederic M. (2001) "The Link Between Gross Profitability and Pharmaceutical R&D Spending," *Health Affairs*, 20(5), 216-220 .

Scherer, Frederic M. (2004) "The Pharmaceutical Industry: Prices and Progress," *New England Journal of Medicine*, 351(9), 927-932.

Schor, Juliet (1993) *The Overworked American: The Unexpected Decline of Leisure*, Basic Books.

Taylor, Mark Z. (2004) "Empirical Evidence Against Varieties of Capitalism's Theory of Technological Innovation," *International Organization*, 58, 601-631.

Wallerstein, Immanuel (1974-2011) *The Modern World System*, 4 Volumes, New York: Academic Press.