

## MIT Open Access Articles

*Asymmetric Growth and Institutions in an Interdependent World*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Acemoglu, Daron, James A. Robinson, and Thierry Verdier. "Asymmetric Growth and Institutions in an Interdependent World." *Journal of Political Economy* 125, no. 5 (October 2017): 1245–1305.

**As Published:** <http://dx.doi.org/10.1086/693038>

**Publisher:** University of Chicago Press

**Persistent URL:** <http://hdl.handle.net/1721.1/118645>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of Use:** Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



# Asymmetric Growth and Institutions in an Interdependent World

---

Daron Acemoglu

*Massachusetts Institute of Technology*

James A. Robinson

*University of Chicago*

Thierry Verdier

*Paris School of Economics—Ecole des Ponts Paris-Tech and PUC—Rio*

We present a model of technologically interconnected countries that benefit and potentially contribute to advances in the world technology frontier. Greater inequality between successful and unsuccessful entrepreneurs increases entrepreneurial effort and a country's contribution to that frontier. Under plausible assumptions, the world equilibrium is asymmetric, involving different economic institutions and technology levels for different countries. Some countries become technology leaders and opt for a type of “cutthroat” capitalism with greater inequality and innovations, while others free ride on the cutthroat incentives of the leaders and choose a more “cuddly” form of capitalism with greater social insurance for entrepreneurs.

We thank Adi Dasgupta and 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 the 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.

Electronically published August 30, 2017

[*Journal of Political Economy*, 2017, vol. 125, no. 5]

© 2017 by The University of Chicago. All rights reserved. 0022-3808/2017/12505-0005\$10.00

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*, 2006, 42)

## I. 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, are at the root of its innovative economy, technological leadership, and high level of per capita income (see Acemoglu, Robinson, and Verdier [2012] for evidence).<sup>1</sup> To its critics, the inequality and the economic uncertainty they create more than offset any efficiency gains that the US economic system may create (and

<sup>1</sup> For example, the top 1 percent of earners in the United States account for about 23.5 percent of national income, while the same number is 10 percent in Sweden, 11.1 percent in Spain, and 12.7 percent in Germany; when capital gains are excluded, it is 9.25 percent in France, 8.3 percent in Finland, 6.1 percent in Denmark, 8.5 percent in Norway, and 9.9 percent in Italy (Atkinson, Piketty, and Saez 2011).

many also doubt that there are such major efficiency gains). 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.

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. Nevertheless, an 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 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 continental Europe.

The basic idea we propose is simple and echoes Hayek’s quote with which we started. 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.<sup>2</sup> We combine this with the idea that innovations require incentives for workers and entrepreneurs (Acemoglu, Akcigit, and Celik 2014). The well-known incentive-insurance trade-off (e.g., Hölmstrom 1979) implies that a society strongly encouraging innovation will have greater inequality (especially among entrepreneurs). 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. The reason is that innovation by followers will create only a “level effect” since the world technology

<sup>2</sup> 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), Keller (2001), Bottazzi and Peri (2003), and Griffith, Redding, and Van Reenen (2003) 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 knowledge spillovers.

frontier is already being pushed forward by advanced economies, while innovation by the technology leader(s), given its impact on the world technology frontier, has a “growth effect,” motivating it (them) to choose high-powered incentives for innovation. This logic underlies the asymmetric nature of world equilibrium. Yet 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, for example, in the form of greater equality among entrepreneurs. As a result, 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.<sup>3</sup>

The bulk of our paper formalizes these ideas using a simple 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, chap. 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” because there are more ideas at the frontier that they have not yet incorporated into their technology. 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 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. In this case, when both the growth benefits from high-powered incentives for innovative success and the gains from insurance are sufficiently high, the leader will adopt a “cutthroat” reward structure while other countries free ride on the leader’s innovations and choose a more egalitarian, “cuddly,” reward structure, at least once they reach a certain level of income and technology.<sup>4</sup>

<sup>3</sup> This observation highlights the parallel between the logic of the dynamic equilibrium in our model and the asymmetric equilibria in some symmetric static games such as chicken as well as games of private provision of public goods (e.g., Bergstrom, Blume, and Varian 1986; Cornes and Hartley 2007); where when one player chooses a certain action (e.g., a high level of contribution), other players may be induced to choose a different action (e.g., to free ride by making a low contribution).

<sup>4</sup> In particular, 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

In the long run, all countries grow at the same rate, but those with cuddly reward structures are poorer.

Our model therefore challenges the idea implicit in much of the debate on US versus European (or Scandinavian) capitalism. Under the assumptions of our model, we cannot 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.<sup>5</sup> The reason is that 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 structures by other countries. 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 and, as already noted, may even yield higher welfare than the cutthroat leader.

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.

Our results extend 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 the technology 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 the asymmetry of 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 similar technology levels, the social democratic party in country 1 may prevent

---

1959), our model implies that some follower countries may first adopt cutthroat reward structures for rapid convergence and then start building more egalitarian institutions once they have approached the world technology frontier sufficiently and reached a certain level of income.

<sup>5</sup> We should emphasize that the relevant aspect of the cutthroat reward structure here is greater reward for innovation and thus greater inequality among entrepreneurs, not greater overall inequality or a lack of safety net at the bottom of the income distribution.

cutthroat capitalism in that country, inducing a unique equilibrium in which country 2 adopts the cutthroat reward structure, while country 1 enjoys greater 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 country 1 commits to a cuddly reward structure and country 2 adopts a cutthroat reward structure in response, unsuccessful entrepreneurs in country 2 are made worse off.

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 (e.g., 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—or 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. 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 VI 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 VI, 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 VIII, interdependence across countries may not be purely technological. For instance, 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 III introduces the economic environment. Section IV presents the main results of the paper. Section V presents two important generalizations of these results. Section VI shows how domestic political economy constraints can be advantageous for a country because they prevent it from adopting a cutthroat reward structure. Section VII provides case study evidence that illustrates the main mechanism proposed in this paper, focusing primarily on innovation and technological spillovers in the pharmaceutical industry. Section VIII presents conclusions, and the proofs of the main results are provided in the Appendix.

## II. 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, for example, Hall and Soskice (2001), which itself builds on earlier intellectual traditions offering taxonomies of different types of capitalism (Crouch 2009) or welfare states (Esping-Anderson 1990). A related argument is developed by Aoki (2001), who also notes the possibility that international interconnections might be at the root of this type of institutional diversity. As mentioned above, Hall and Soskice (2001) argue that CMEs and LMEs innovate in different ways and in different sectors: LMEs are good at “radical innovation” characteristic of particular sectors, such as software development, biotechnology, and semiconductors, while CMEs are good at “incremental innovation” in sectors such as machine tools, consumer durables, and specialized transport equipment. This literature has not considered that growth in a CME might critically depend 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 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 United States 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 United States 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 United States and France. Bénabou and Tirole (2006) develop a model in which self-fulfilling beliefs about justice and fairness can lead to divergent re-



distributive policies across countries. Bénabou (2000, 2006) establishes the possibility of multiple political economy 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; Moene and Wallerstein 1997). Bénabou (2006, 1598), in particular, emphasizes the implications of this multiple equilibria for technology, asking “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?”<sup>6</sup> None of this work, however, focuses on asymmetric equilibria resulting from cross-country institutional interactions.

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 et al. (2006) (see also Krueger and Kumar 2004), though without endogenizing different institutional equilibria as best responses to each other.

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.

### III. 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, closely following chapter 18 of Acemoglu (2009). The second introduces moral hazard on the part of entrepreneurs, thus linking entrepreneurial innovative activ-

<sup>6</sup> Another, and somewhat more distant, branch of the literature has emphasized the role of differences in political systems. These include how electoral systems with proportional representation, widely used in continental Europe, may lead to greater redistribution (Alesina, Glaeser, and Sacerdote 2001; Milesi-Ferretti, Perotti, and Rostagno 2002; Persson and Tabellini 2003); how federalism may lower redistribution (Cameron 1978); how the greater ethnic heterogeneity of the United States may reduce the demand for redistribution (Alesina et al. 2001); how greater social mobility in the United States may mute the desire for redistributive taxation (Bénabou and Ok 2001; Alesina and La Ferrara 2005); or the possibility of greater redistribution in northern Europe because of higher levels of social capital and trust (Algan, Cahuc, and Sangnier 2011).

ity 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 utility.

### A. Economic Environment

Consider an infinite-horizon economy consisting of  $J$  countries, indexed by  $j = 1, 2, \dots, J$ . Each country is inhabited by nonoverlapping generations of agents who live for a period of length  $\Delta t$ , work, produce, consume, and then die. A continuum of agents, with measure normalized to one, 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  $\Delta t \rightarrow 0$ , represented as a continuous-time model (see Sec. IV.H for a discussion of this and other modeling assumptions).

The aggregate production function at time  $t$  in country  $j$  is

$$Y_j(t) = \frac{1}{1 - \beta} \left[ \int_0^{N_j(t)} x_j(v, t)^{1-\beta} dv \right] L_j^\beta, \quad (1)$$

where  $L_j$  is labor input, and  $N_j(t)$  denotes the number of intermediates (or blueprints for producing them) 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,  $x_j(v, t)$  is the total amount of intermediate  $v$  used in country  $j$  at time  $t$ . Crucially, blueprints for producing these intermediates, captured by  $N_j(t)$ , live on, and the increase in the range of these blueprints will be the source of economic growth.

Each blueprint in economy  $j$  is owned by an “entrepreneur,” who sells the corresponding intermediate at the profit-maximizing price  $p_j(v, t)$  within the country (there is no international trade). This monopolist entrepreneur can produce each unit of the intermediate at a marginal cost of  $\psi$  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 intermediate. 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.

The utility function of entrepreneur  $i$  takes the form

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

where  $\theta \geq 0$  (and  $\theta \neq 1$ ) is the coefficient of relative risk aversion. This form of the utility function ensures balanced growth.<sup>7</sup>

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  $j = 1, \dots, J$ .

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

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

new intermediates, where  $N(t)$  is an index of the world technology frontier, which will 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 the frontier is more advanced, an innovation will lead to more rapid progress, and the parameter  $\phi$  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

$$\dot{N}_j(t) = \{q_1 e_j(t) + q_0 [1 - e_j(t)]\} \eta N(t)^\phi N_j(t)^{1-\phi}. \quad (3)$$

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 produced monopolistically as well.<sup>8</sup>

The world technology frontier is assumed to be given by

$$N(t) = G(N_1(t), \dots, N_J(t)), \quad (4)$$

<sup>7</sup> When  $\theta = 1$ , the utility function (2) should be set to  $\ln C_{j,i}(t) + \ln[1 - \gamma e_{j,i}(t)]$ . Note also that we do not include the  $-1$  sometimes included in the numerator of this class of utility functions in order to simplify notation, and hence, the limit  $\theta \rightarrow 1$  does not converge to log preferences. All of our results apply to this log case also, but in what follows we do not discuss this case separately to save space.

<sup>8</sup> 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 change just the value of  $B$  in (10) below.

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

$$G(N_1(t), \dots, N_j(t)) = \max\{N_1(t), \dots, N_j(t)\}, \quad (5)$$

which implies that the world technology frontier is represented by the technology level of the most advanced country, the *technology leader*, and all other countries benefit from the advances of this technology leader. The second is a more general convex aggregator

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

with  $\sigma < 0$ . The term  $J^{\sigma/(\sigma-1)}$  in the denominator ensures that the convex aggregator is homogeneous of degree one 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 we show in Section V that our general results are robust when we use (6) with  $\sigma$  sufficiently close to zero.

### B. Reward Structures

Entrepreneurial effort levels will depend on the reward structure in each country, which determines the relative rewards to successful entrepreneurship. 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  $t$  (where  $\tilde{R}^u(t)$  and  $\tilde{R}^s(t)$  also include the expected rents that entrepreneurs make because of existing ideas being randomly allocated to them). 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 Sec. VI). Limiting the social planner to choosing only the sequence of reward structures is done for simplicity and without any consequence.<sup>9</sup>

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

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

where  $\rho$  is the discount rate that the social planner applies to future generations, and  $\{C_{j,i}(t)[1 - \gamma e_{j,i}(t)]\}^{1-\theta}/(1 - \theta)$  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,  $\theta$ , 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) yields 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. Defining

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

we can represent these more general preferences as

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

Here  $\theta$  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  $\theta$  and  $\lambda$  except 1. These more general preferences are useful because they enable us to separate the intragenerational relative risk aversion,  $\theta$ , and the willingness of the social planner to substitute consumption across cohorts, given by  $\lambda$ , enabling us to show that it is  $\theta$  that regulates entrepreneurs' incentives, while  $\lambda$  matters only in conjunction

<sup>9</sup> If, in addition, we allowed the social planner to set prices that remove the monopoly markup, this would change only the value of the constant  $B$  in (10) below, with no impact on any of our results.

with the future rate of economic growth in the social planner’s calculations. Moreover, the separation of these two distinct concepts will lead to richer institutional dynamics as we will see in Section V.

**IV. 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 in which preferences are given by (8), and the world technology frontier is as in (6) in Section III.

*A. 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

$$\begin{aligned} & \frac{1}{1 - \theta} [q_1 R^s(t)^{1-\theta} + (1 - q_1)R^u(t)^{1-\theta}](1 - \gamma)^{1-\theta} \\ & \geq \frac{1}{1 - \theta} [q_0 R^s(t)^{1-\theta} + (1 - q_0)R^u(t)^{1-\theta}], \end{aligned}$$

where recall that  $R^s(t)$  is the income and thus the consumption of an entrepreneur conditional on successful innovation, and  $R^u(t)$  is the income level when unsuccessful. 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

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

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.

ASSUMPTION 1.

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

More specifically, this assumption guarantees that when  $\theta < 1$ ,

$$1 + \frac{1 - (1-\gamma)^{1-\theta}}{q_1(1-\gamma)^{1-\theta} - q_0}$$

in (9) is greater than one and is raised to a positive power, while when  $\theta > 1$ , it is less than one and is raised to a negative power. Thus in both cases, we have  $A > 1$ , implying that high effort will be forthcoming only if entrepreneurial success is rewarded relative to failure.

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 intermediates provided in the Appendix, this implies

$$q_1 R^s(t) + (1 - q_1) R^u(t) = B N_j(t),$$

where

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

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

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

The alternative to a reward structure that encourages effort is one that forgoes effort and provides full consumption insurance, that is, the same level of income to all entrepreneurs of  $R^o(t)$ , regardless of whether they are successful or not. The same resource constraint then implies

$$R^o(t) = B N_j(t). \quad (12)$$

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

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

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

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

where

$$\omega_c \equiv \frac{[q_1 A^{1-\theta} + (1 - q_1)](1 - \gamma)^{1-\theta} B^{1-\theta}}{[q_1 A + (1 - q_1)]^{1-\theta} (1 - \theta)} \quad \text{and} \quad \omega_o \equiv \frac{B^{1-\theta}}{1 - \theta}, \quad (14)$$

where  $A$  was defined in (9). It can be verified that  $\omega_c < \omega_o$ , though when  $\theta > 1$ ,  $\omega_c < \omega_o < 0$ . Moreover, we have that (i)  $\omega_c$ , and thus  $(\omega_c/\omega_o)^{1/(1-\theta)}$ , is decreasing in  $A$  (since a higher  $A$  translates into greater consumption variability); (ii)  $A$  is increasing in  $\gamma$  (and thus  $(\omega_c/\omega_o)^{1/(1-\theta)}$  is decreasing in  $\gamma$ ), which compensates for the higher cost of effort (see lemma A2 in the Appendix); and (iii)  $A$  is nonmonotone in  $\theta$  (because a higher coefficient of relative risk aversion also reduces the disutility of effort).

Two more observations are useful. First, the key expressions  $\omega_c$  and  $\omega_o$  depend on agents' coefficient of risk aversion,  $\theta$ . This will continue to be the case when we adopt the more general preferences in (8) for the social planner (rather than depending on both  $\theta$  and the social planner's inequality aversion parameter,  $\lambda$ ). Second, without loss of any generality, we focus on cutthroat and cuddly incentives for entrepreneurs (and thus on  $\omega_c$  and  $\omega_o$ ), since if the social planner wishes to induce effort, she cannot do better than a cutthroat reward structure; and if she wishes to provide perfect insurance, she has to choose a cuddly reward structure. Therefore, the most general set of reward structures 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 set of reward structures we focus on in the rest of our analysis.

### B. World Equilibrium Given Reward Structures

We 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 just the technology leader (with [5]). In addition, differences in reward structures determine the relative income of each country in the long run.

**PROPOSITION 1.** Suppose that the reward structure for each country is asymptotically time invariant (i.e., for each  $j$ ,  $R_j^s(t)/R_j^u(t) \rightarrow r_j$ ). Then starting from any initial condition  $(N_1(0), \dots, N_J(0))$ , the world economy converges to a unique stationary distribution  $(n_1^*, \dots, n_J^*)$ , with  $n_j(t) \equiv N_j(t)/N(t)$  and  $\dot{N}(t)/N(t) = g^*$ . In addition,  $(n_1^*, \dots, n_J^*)$  and  $g^*$  are functions of  $(r_1, \dots, r_J)$ . Moreover, with the max specification of the world technology frontier, (5),  $g^*$  is a function of only the most innovative country's reward structure,  $r_e$ .

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—provided that the reward structures do not fluctuate asymptotically. 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 pull them to the same growth rate as those that provide greater inducements to innovation. The proposition also shows that in the special case in which (5) applies, the world growth rate,  $g^*$ , will be determined only by innovation, and thus the reward structure, in the technologically most advanced country.

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.

### C. *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 that provide cutthroat incentives to a fraction of the entrepreneurs. The Markovian restriction implies that strategies at time  $t$  are conditioned only on payoff-relevant variables, which are the vector of technology levels,  $N_1(t), \dots, N_J(t)$ .

Let us define  $u_j(t) \in [0, 1]$  as the fraction of entrepreneurs receiving cutthroat incentives. Clearly,  $u_j(t) = 0$  corresponds to a fully cuddly reward structure and  $u_j(t) = 1$  corresponds to a fully cutthroat reward structure throughout.

From equation (13), average utility in country  $j$  at time  $t$  is  $\omega(u_j(t))N_j(t)^{1-\theta}$ , where

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

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

$$g_o \equiv q_o \eta \quad \text{and} \quad g_c \equiv q_i \eta > g_o \tag{18}$$

corresponding, respectively, to the growth rates from fully cuddly and fully cutthroat reward structures. The fact that  $g_c > g_o$  reiterates that cutthroat incentives generate more successful innovations than cuddly incentives.

#### D. *Asymmetric World Equilibria*

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  $\ell$  that is the technology leader, that is, a single  $\ell$  for which  $N_\ell(0) = \max\{N_1(0), \dots, N_j(0)\}$ . 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 is the standard condition ensuring that the cutthroat growth rate,  $g_c$ , is not so high as to lead to infinite discounted utility for the country social planners.

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, since  $\omega_c < \omega_o$  (reflecting the risk-sharing benefits of cuddly reward structures), this assumption requires that the cutthroat

growth rate  $g_j$  is sufficiently higher than the cuddly growth rate  $g_\ell$  to compensate for the loss of consumption insurance under cutthroat incentives. This assumption also implies that the technology leader will prefer cutthroat to cuddly incentives.

Finally, we start by restricting attention to a specific equilibrium selection rule, which imposes that the same country,  $\ell$ , remains the technology leader throughout. This is stated in the next assumption.

ASSUMPTION 4.  $N_\ell(t) = \max\{N_1(t), \dots, N_j(t)\}$  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 may be 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 IV.G.

To build toward our main result, suppose that country  $\ell$  adopts a cutthroat reward structure at all times, that is,  $u_\ell(t) = 1$  for all  $t$ . The problem of follower country  $j$ 's social planner can then be written as

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

With a change of variable  $m_j \equiv (N_j/N_\ell)^\phi \leq 1$ , this maximization problem can be expressed as

$$\begin{aligned} \mathcal{W}_j(m_j(t)) &= N_\ell(t) \max_{u_j(\cdot) \in [0,1]} \int_t^\infty e^{-[\rho - (1-\theta)g_\ell](\tau-t)} \\ &\quad \times \omega(u(\tau)) m_j(\tau)^{(1-\theta)/\phi} d\tau, \\ \dot{m}_j(\tau) &= \phi[g(u(\tau)) - g_\ell m_j(\tau)]. \end{aligned} \tag{19}$$

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,  $\ell$ . The Markov perfect equilibrium, which we are interested in, corresponds to the situation in which all countries use “closed-loop” strategies. However, under assumption 4, which imposes that the same country,  $\ell$ , remains the leader (and that it will adopt a cutthroat reward structure in view of assumption 3), the open- and closed-loop solutions coincide. Hence we can characterize all equilibria by deriving the solution to (19).

The solution to (19) can be characterized by setting up the current-value Hamiltonian, which, suppressing the country index  $j$  to simplify notation, takes the form

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

where  $\mu(t)$  is the current-value costate variable and  $u(t)$  is the control variable summarizing the reward structure. The maximum principle yields the candidate bang-bang solution

$$u(t) \begin{cases} = 1 & \text{if } \Psi(t) < 0 \\ \in [0, 1] & \text{if } \Psi(t) = 0 \\ = 0 & \text{if } \Psi(t) > 0, \end{cases} \tag{20}$$

where  $\Psi(t)$  is a switching function that summarizes the social return from cuddly incentives:

$$\Psi(t) \equiv (\omega_o - \omega_c)m(t)^{(1-\theta)/\phi} - \mu(t)\phi[g_c - g_o]. \tag{21}$$

In addition, the differential equations for the laws of motion of transformed relative technology and the costate variable,

$$\begin{aligned} \dot{m}(t) &= \phi[g(u(t)) - g_c m(t)] \quad \text{with } m(0) > 0 \text{ given,} \\ \dot{\mu}(t) &= [\rho - (1 - \theta)g_c + \phi g_c]\mu(t) - \frac{1 - \theta}{\phi} m(t)^{[(1-\theta)/\phi]-1} \omega(u(t)), \end{aligned} \tag{22}$$

and the transversality condition,

$$\lim_{t \rightarrow \infty} e^{-[\rho - (1 - \theta)g_c]t} \mu(t) = 0, \tag{23}$$

have to hold.

This dynamical system can be analyzed using phase diagrams as we do in figures 1–3. For this purpose, we define and use  $\kappa(t) \equiv \mu(t)m(t)^{(\theta-1)/\phi}$  (instead of  $\mu(t)$ ), which ensures that the behavior in the phase diagram does not depend on whether  $\theta$  is greater than or less than one. Substituting this transformation into (21), the choice of reward structure, (20) simplifies to

$$u(\kappa) \begin{cases} = 1 & \text{if } \kappa > \tilde{\kappa} \\ \in [0, 1] & \text{if } \kappa = \tilde{\kappa} \\ = 0 & \text{if } \kappa < \tilde{\kappa}, \end{cases} \tag{24}$$

where we have suppressed the dependence of  $\kappa$  on time, and

$$\tilde{\kappa} \equiv \frac{\omega_o - \omega_c}{\phi(g_c - g_o)}$$

is the value  $\tilde{\kappa}$  that makes the social planner indifferent between cuddly and cutthroat incentives.

In the phase diagrams, we place  $n_j(t)^\phi$  on the horizontal axis and  $\kappa_j(t)$  on the vertical axis (now using country subscripts for emphasis). Inspection of (22) implies that the locus for  $\dot{\kappa}_j(t) = 0$  is downward sloping and intersects  $\tilde{\kappa}$  at  $n(t)^\phi = \tilde{m}$ , defined by (25) in proposition 2. To understand the shape of the locus for  $\dot{n}_j(t) = 0$ , recall that when  $\kappa$  is above  $\tilde{\kappa}$ , the social planner will choose cutthroat incentives, and when it is below this value, she will choose cuddly incentives. Consequently, above  $\tilde{\kappa}$ , country  $j$  tends to the world technology frontier and the locus for  $\dot{n}_j(t) = 0$  is vertical at one, while below  $\tilde{\kappa}$ , it approaches  $(g_o/g_c)^{1/\phi}$ , so that the locus for  $\dot{n}_j(t) = 0$  is vertical at  $g_o/g_c$  (and is horizontal in between at  $\tilde{\kappa}$ ).<sup>10</sup>

Figure 1 applies when  $\tilde{m} < g_o/g_c$  and world equilibrium is asymmetric as emphasized in the introduction (this corresponds to part 1 of proposition 2 presented next). Recall that the technology leader, country  $\ell$ , always chooses cutthroat rewards,  $u_\ell(t) = 1$ . Moreover, in this configuration, the curves for  $\dot{n}_j(t) = 0$  and  $\dot{\kappa}_j(t) = 0$  intersect below the horizontal line corresponding to  $\tilde{\kappa}$ , which ensures that as  $t \rightarrow \infty$ , all followers will choose cuddly rewards,  $u_j(t) = 0$  for all  $j \neq \ell$  and will therefore approach an income level equal to a fraction  $g_o/g_c$  of the leader's. In addition, there is another noteworthy threshold,  $\tilde{m}$ , given by the intersection (if any) of the stable arm with the horizontal line at  $\tilde{\kappa}$ . When  $n_j(0) < \tilde{m}^{1/\phi}$ , we have  $\kappa > \tilde{\kappa}$ , and thus follower  $j$  starts with  $u_j(t) = 1$  and then switches to cuddly incentives,  $u_j(t) = 0$  at  $n_j(t) = \tilde{m}^{1/\phi}$ .

Figure 2 corresponds to the case in which  $1 > \tilde{m} > g_o/g_c$  (part 2 of proposition 2). In this case, the curves for  $\dot{n}_j(t) = 0$  and  $\dot{\kappa}_j(t) = 0$  intersect at  $\kappa = \tilde{\kappa}$ , which implies that in the long run followers will adopt mixed reward structures. 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—and in the long run, level of income per capita—between those implied by a fully cuddly and a fully cutthroat reward structure.

Finally, figure 3 is for the case in which  $\tilde{m} > 1$  (part 3 of the proposition). Now the curves for  $\dot{n}_j(t) = 0$  and  $\dot{\kappa}_j(t) = 0$  intersect where  $\kappa > \tilde{\kappa}$ , and thus followers also adopt cutthroat incentives and there is “institutional” technology and income per capita convergence. This configuration thus contrasts with the patterns in the previous two figures, where countries maintain their different institutions in the long run, and as a result, their levels of income per capita do not converge.

The next proposition summarizes these results more formally.

<sup>10</sup> The laws of motion in these phase diagrams follow readily from (22) and are derived in the Appendix as a special case of the phase diagrams introduced in the proof of proposition 6.

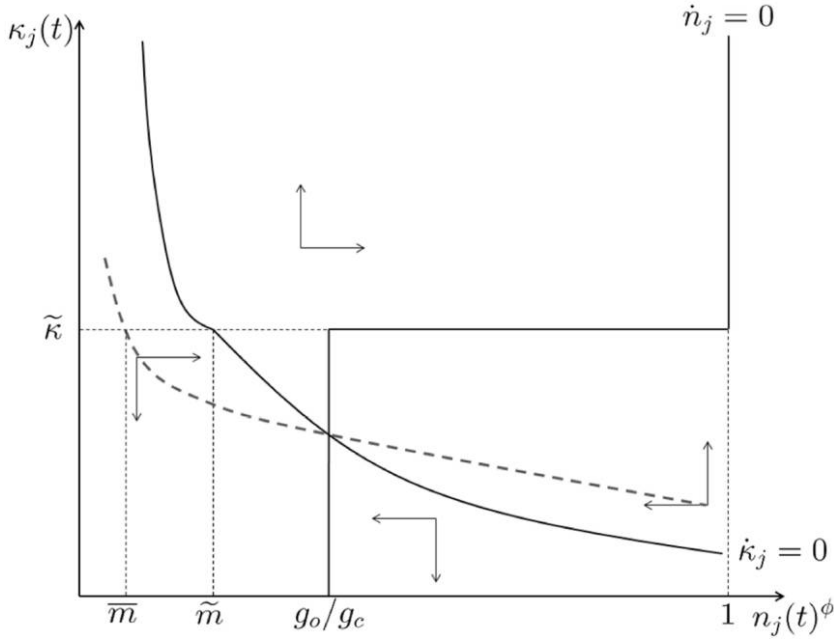


FIG. 1.—Phase diagram: part 1 of proposition 2

PROPOSITION 2. Suppose that country social planners maximize (7), the world technology frontier is given by (5), and assumptions 1–4 hold. Let

$$\tilde{m} \equiv \frac{(1 - \theta)g_o}{\rho + \phi g_c} \left[ \frac{(g_c - g_o)/g_o}{(\omega_o - \omega_c)/\omega_o} + 1 \right], \tag{25}$$

where  $\omega_o, \omega_c, g_o$  and  $g_c$  are given by (14) and (18). Then world equilibrium is characterized as follows. The leader country  $\ell$  always chooses cutthroat rewards, that is,  $u_\ell(t) = 1$  for all  $t$ . For each follower  $j \neq \ell$ , we have the following:

1. If

$$\tilde{m} < \frac{g_o}{g_c}, \tag{26}$$

there exists  $\bar{m} < g_o/g_c$  such that, for  $n_j(0) < \bar{m}^{1/\phi}$ , the reward structure of country  $j$  is cutthroat (i.e.,  $u_j(t) = 1$ ) for all  $t < T < \infty$ , where  $n_j(T) = \bar{m}^{1/\phi}$ , and cuddly (i.e.,  $u_j(t) = 0$ ) for all  $t \geq T$ ; for  $n_j(0) \geq \bar{m}^{1/\phi}$ , 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  $n_j(t) \rightarrow (g_o/g_c)^{1/\phi}$ . (This case is illustrated in fig. 1.)

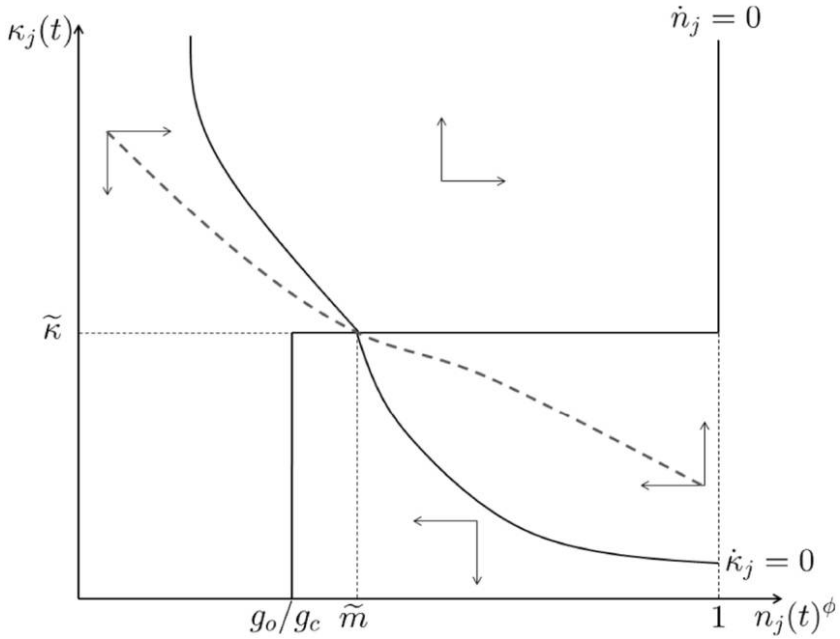


FIG. 2.—Phase diagram: part 2 of proposition 2

2. If

$$\frac{g_o}{g_c} < \tilde{m} < 1, \tag{27}$$

then for  $n_j(0) < \tilde{m}^{1/\phi}$ , the reward structure of country  $j$  is cutthroat (i.e.,  $u_j(t) = 1$ ) for all  $t < T < \infty$ , where  $n_j(T) = \tilde{m}^{1/\phi}$ , and then at  $t = T$ , the country adopts a “mixed” reward structure and stays at  $n_j(t) = \tilde{m}^{1/\phi}$  (i.e.,  $u_j(t) = u_j^* \in (0, 1)$ ) for all  $t \geq T$ ; for  $n_j(0) > \tilde{m}^{1/\phi}$ , the reward structure of country  $j$  is cuddly (i.e.,  $u_j(t) = 0$ ) for all  $t < T$ , and then at  $t = T$  when  $n_j(T) = \tilde{m}^{1/\phi}$ , the country adopts a mixed reward structure and stays at  $n_j(t) = \tilde{m}^{1/\phi}$  (i.e.,  $u_j(t) = u_j^* \in (0, 1)$ ) for all  $t \geq T$ . (This case is illustrated in fig. 2.)

3. If

$$\tilde{m} > 1, \tag{28}$$

then the reward structure of country  $j$  is cutthroat for all  $t$ ; that is,  $u_j(t) = 1$  for all  $t$ . (This case is illustrated in fig. 3.)

The proof of this proposition, like the proofs of all of the remaining results in this paper, is provided in the Appendix. Here, we provide a discussion and interpretation of the results.

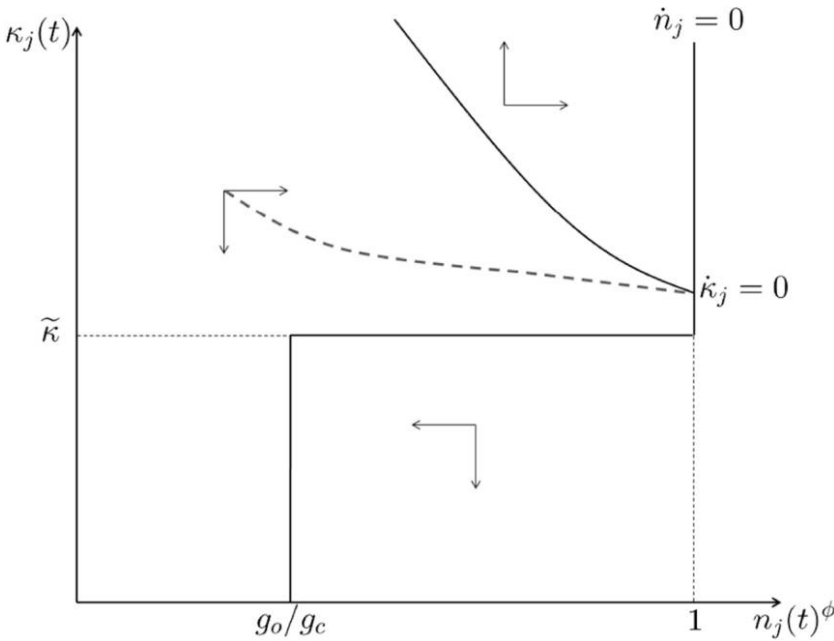


FIG. 3.—Phase diagram: part 3 of proposition 2

First, as already noted, the most noteworthy result in this proposition is the asymmetric nature of the equilibrium in part 1 (and also in part 2). This asymmetry 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,  $\omega_c/[\rho - (1 - \theta)g_c]$  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 structure, if country  $l$  were also to do so, the world economy would grow only at the rate  $g_o < g_c$ , yielding a discounted value of  $\omega_o/[\rho - (1 - \theta)g_o]$  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  $l$  to its entrepreneurs have a growth effect on the world economy (and thus on itself).

In contrast, 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 (26) then ensures that in the limit this level effect is more than compen-



sated for by the better risk sharing offered 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 nonexistence of a symmetric equilibrium).

It is also worth noting that, without assumption 4, there could be other equilibria, with some other country playing the role of the leader  $\ell$  and choosing a cutthroat reward structure. 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  $\ell$ ). We discuss conditions for uniqueness without assumption 4 below.

As already noted,  $\tilde{m}^{1/\phi}$ , defined in (25), is the threshold value of the position of a follower country  $j$  relative to the world technology frontier, that is,  $N_j(t)/N(t)$ , such that the follower is just indifferent between cuddly and cutthroat incentives (and above this threshold prefers cuddly incentives). To obtain the intuition for the expression for  $\tilde{m}$ , let us start with the first term in (25). When the discount rate  $\rho$  is greater and the growth spillover from the leader,  $\phi g_c$ , is larger (relative to the growth rate of followers,  $g_o$ ), the followers benefit more from free riding on the leader's cutthroat incentives and the threshold  $\tilde{m}$  above which cuddly incentives will be adopted is smaller. The second term,  $[(g_c - g_o)/g_o]/[(\omega_o - \omega_c)/\omega_o]$ , on the other hand, can be identified as the proportional gain in growth from cutthroat incentives relative to the proportional loss in utility from cutthroat incentives. The lower this term is, the less valuable are the growth benefits of cutthroat incentives relative to their consumption insurance losses, and thus the smaller is the threshold  $\tilde{m}$ .

For reasonable parameter values,  $\tilde{m}$  is in the range corresponding to asymmetric world equilibria (in particular, part 1 of proposition 2). To illustrate this point, let us adopt relatively standard values for the discount rate, the coefficient of relative risk aversion, and the labor share,  $\rho = 0.01$ ,  $\theta = 2$ , and  $\beta = 0.6$ . To choose the entrepreneurial success probability conditional on high effort,  $q_1$ , and the composite parameter  $A$  given by (9), suppose that, as is often assumed, the right tail of the income distribution (corresponding to income inequality among entrepreneurs) is approximately Pareto. The shape parameter of the Pareto distribution,  $\nu = \mathbb{E}[y|y \geq \bar{y}]/\bar{y}$ , provides a measure of inequality at the top of the income distribution—here corresponding to inequality among entrepreneurs. In our model under cutthroat incentives, this number is

$$\frac{A}{Aq_1 + 1 - q_1}.$$

We take the United States, which has the highest Pareto shape parameter in the database of Atkinson et al. (2011) of 2.5, as the technology leader.

We also approximate  $q_1$  as the population fraction of “successful entrepreneurs” from the Global Entrepreneurship Monitoring Survey (2015), which we define as those who have started a business paying wages, excluding those who report to have chosen entrepreneurship by necessity. This gives  $q_1 = 0.09$  and, combined with the previous expression, implies  $A = 2.95$ . We set  $g_c = 0.021$  to match the average US growth rate between 1970 and 2005, the period on which we focus. Combined with  $q_1 = 0.09$ , this growth rate implies  $\eta = 0.229$ .

We estimate the two remaining parameters  $\phi$  and  $q_0$  from cross-country growth dynamics. We first solve the differential equation (16) under the assumption that each follower is choosing cuddly incentives ( $u_j(t) = 0$ ), which gives

$$n_j(t) = \left\{ \frac{g_o}{g_c} + \left[ n_j(0)^\phi - \frac{g_o}{g_c} \right] e^{-\phi g_c t} \right\}^{1/\phi}.$$

We add a multiplicative error term to this expression,  $e^{\xi_j(t)}$ , and since variations in  $n_j(t)$  are the sole source of variations in GDP, we approximate  $n_j(t)$  with country  $j$ 's GDP per capita relative to the United States in year  $t$ ,  $y_j(t)$ . Taking logs, we end up with the estimating equation

$$\ln y_j(t) = \frac{1}{\phi} \ln \left\{ \frac{g_o}{g_c} + \left[ y_j(0)^\phi - \frac{g_o}{g_c} \right] e^{-\phi g_c t} \right\} + \xi_j(t).$$

We use data for 20 OECD countries from the Penn World Tables and take the base year to be 1970 and the end year, year  $t$ , to be 2005. Our sample excludes the United States, which is used for normalization, and Norway, because of the role of oil in its GDP. Estimating this equation with nonlinear least squares (and imposing  $g_c = 0.021$ ), we obtain  $\phi = 0.51$  (standard error = 0.16) and  $g_o/g_c = 0.82$  (standard error = 0.053). This latter estimate also yields  $q_0 = 0.076$ . (Finally, this combination of parameters also implies a value of  $\gamma = 0.012$ , which we find reasonable as the proportional cost of additional effort.)

These numbers imply  $\bar{m} = 0.64$ , which is less than  $g_o/g_c = 0.82$ , putting us comfortably in the region of asymmetric equilibria as in part 1 of proposition 2.

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 beneficial and this is achieved best by adopting cutthroat incentives. However, after  $n_j(t)$  exceeds  $\bar{m}^{1/\phi}$ , 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 stable position relative to the leader given by  $(g_o/g_c)^{1/\phi}$ .

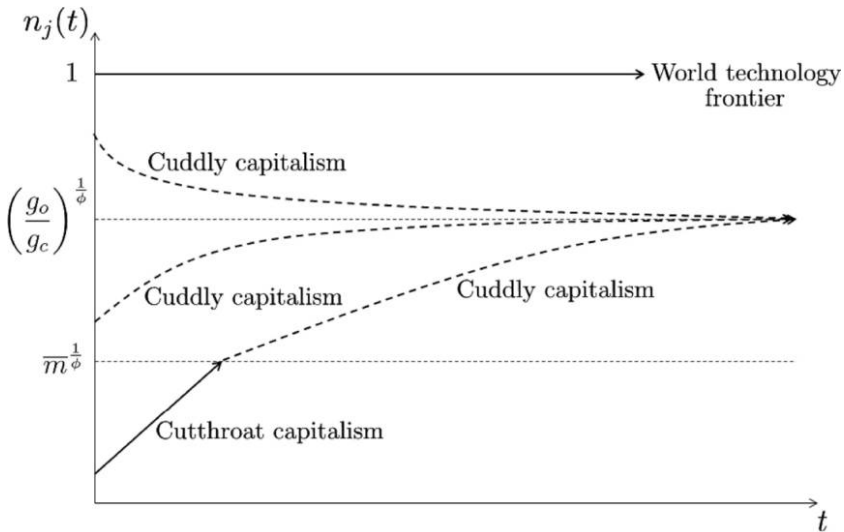


FIG. 4.—Growth dynamics: part 1 of proposition 2

The dynamic path for a follower country starting below  $\bar{m}^{1/\phi}$  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 (e.g., Deyo 1993), but then started building stronger social insurance arrangements. 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 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 toward more democratic institutions, which went hand in hand with the development of better social insurance in these countries).

Figure 5 illustrates the second part of the proposition, where the long-run reward structure of followers is a mixed one with some fraction of its entrepreneurs receiving cutthroat incentives while the rest receive cuddly incentives. This fraction is chosen such that the proportional gap between this country and the leader remains exactly at  $\tilde{m}^{1/\phi}$ . The growth rate of this economy at any point in time is greater than the one that would have followed from fully cuddly incentives. As a result, this country becomes richer in the long run than one that chooses a fully cuddly reward structure. (We do not show part 3, which involves all countries choosing cutthroat incentives, in a separate diagram to save space.)

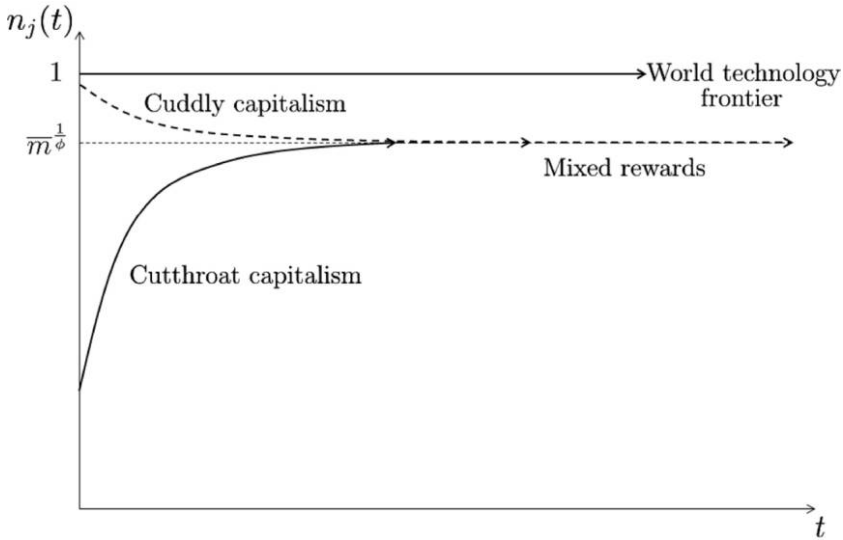


FIG. 5.—Growth dynamics: part 2 of proposition 2

*E. When Is the Equilibrium Asymmetric?*

The next proposition shows how the range in which asymmetric equilibria exist, that is, where  $\hat{m} < 1$ , is affected by changes in several parameters of the model.

**PROPOSITION 3.** Suppose assumptions 1–4 hold. Then, the world equilibrium is more likely to be asymmetric (meaning that the condition  $\hat{m} < 1$  corresponding to part 1 or part 2 of proposition 2 is more likely to be satisfied) when  $\gamma$  is large (the cost of innovation is large) and when  $\phi$  is large (spillovers are large). Also for  $\theta < 1$  and  $\gamma$  sufficiently small, world equilibrium is more likely to be asymmetric when  $\theta$  is higher (risk aversion is large).

These results are intuitive. When  $\gamma$  is large, cutthroat incentives become less attractive, though assumption 3 still ensures that the technology leader, country  $\ell$ , prefers cutthroat incentives. When  $\phi$  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  $\theta$  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  $\theta$  also determines the willingness of the social planner to trade off consumption between cohorts, this result requires us to focus on the case in which  $\theta < 1$  (we separate these two notions in Sec. V).

### *F. Welfare*

We next look at the welfare implications of the asymmetric world equilibrium. We show that, even though the technology leader, country  $\ell$ , starts out ahead of others and chooses a “growth-maximizing” strategy, discounted utility (using the social planner’s discount rate) is higher in a follower country choosing cuddly incentives than the leader choosing cutthroat incentives, provided that the follower is not too far behind the leader to start with. This result, stated in the next proposition, captures the central economic force of our model: followers are able to both choose an egalitarian reward structure providing perfect insurance to their entrepreneurs and benefit from the rapid growth of technology driven by the technology leader, country  $\ell$ , because they are able to free ride on the cutthroat reward structure in country  $\ell$ , which is advancing the world technology frontier. In contrast, country  $\ell$ , as the technology leader, must bear the cost of high risk for its entrepreneurs. The fact that followers prefer to choose the cuddly reward structure implies that, all else equal, country  $\ell$  would have also liked to do so but cannot because it realizes that if it did, the growth rate of the world technology frontier would slow down, while followers know that the world technology frontier is being advanced by country  $\ell$  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  $n_j(0) > 1 - \delta$ , discounted utility (at time  $t = 0$ ) in country  $j$  is higher than discounted utility in country  $\ell$ .

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, such as the Scandinavian countries, even though the former may be technologically more advanced and richer, discounted utility 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 induced by US cutthroat incentives, and if the United States also switched to cuddly reward structures, the world growth rate would slow down.

### *G. Uniqueness without the Selection Rule*

The analysis so far has proceeded under assumption 4, which imposed the selection rule that country  $\ell$  remains the technology leader. Without

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  $\tilde{m} < 1$ ). Yet not only are there two mirror-image equilibria, one in which the first country is the technology leader and another 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.<sup>11</sup>

Nevertheless, in this subsection we show that if the technology leader, again denoted by  $\ell$ , is sufficiently ahead of the followers, the asymmetric equilibrium characterized in proposition 2 is unique. The reasoning is simple but important: if followers are significantly behind country  $\ell$ , then any profile of reward structures that involves country  $\ell$  losing the technology leadership will involve the world technology frontier, and thus country  $\ell$ , growing at the slower rate  $g_o$  rather than  $g_c$  for an extended period of time. Then under assumption 3, country  $\ell$  would in fact prefer to adopt a cutthroat reward structure. If, in addition, the asymmetric equilibrium in which country  $\ell$  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  $\ell$ , then the unique equilibrium will involve  $\ell$  choosing a cutthroat reward structure and 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 present sufficient conditions for the uniqueness of the equilibrium in which the initial technology leader always remains the leader.

**PROPOSITION 5.** Suppose that all of the hypotheses of proposition 2 hold and  $\tilde{m} < g_o/g_c$  (so that we are in part 1 of proposition 2). Then there exist  $\phi \in (0, 1)$  and  $\underline{n} \in (0, 1)$  (where the exact expression for  $\underline{n}$  is given in the proof) such that if  $\phi < \phi$  and  $n_j(0) < \underline{n}$  for all  $j \neq \ell$ , the equilibrium is unique and involves country  $\ell$  adopting a cutthroat reward structure throughout and all other countries choosing a cuddly reward structure

<sup>11</sup> Equilibria in which there are switches in technology leadership are 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.

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).

Returning to the illustrative parameter values used above, we can compute  $\underline{n}$  as 0.60 (which also confirms that  $\phi = 0.51$  in this case is less than  $\underline{\phi}$ ). This shows that for plausible, though certainly nontrivial, values of the technology gaps between leader and follower countries, the asymmetric world equilibrium is unique without the selection assumption (assumption 4).

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, the largest country may wish to be the technology leader by choosing cutthroat incentives because technology leadership by a smaller country would reduce the growth rate of the world technology frontier. This economic force would also resolve the multiplicity problem without assumption 4.

#### *H. 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 given only short-term incentives, nothing in the analysis changes. However, with long-lived agents, optimal contracts for moral hazard 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 that 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 im-



plications of partially uninsured income risk for risk taking is an interesting area for research and the exact implications of risk sharing for innovative activities need 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 along the equilibrium path, and 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 as well. Though we believe that 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, are important, some of these other linkages may lead to asymmetric world equilibria as well. For instance, when countries trade, there are complementary reasons for them 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 reward 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.<sup>12</sup>

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. Abstracting from these considerations has enabled us to delineate the key force leading to an asymmetric world equilibrium. In Section VI, we show how domestic political constraints, modeled in a simple manner, interact

<sup>12</sup> Relatedly, see Acemoglu and Ventura (2002) 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. See also the recent paper by Guimaraes and Sheedy (2017), which proposes a related mechanism for institutional specialization based on the differential value of rule of law in an international trade equilibrium depending on whether other countries do or do not have rule of law.



with the forces we have highlighted so far and play the role of selecting which country will be a technology leader with cutthroat incentives.

## V. 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 our results hold with the general convex form of the world technology frontier given in (6).

### A. General Preferences

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

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  $\lambda$ , since it regulates how future increases in utility are valued today.

ASSUMPTION 3'.

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

The structure of world equilibrium in this setting parallels proposition 2 and features three regimes. In the first regime, which is similar to part 1 of proposition 2, followers distant enough from the leader adopt a cutthroat reward structure, while followers close enough to the leader switch to cuddly incentives. The second regime, which is similar to part 2 of proposition 2, is also asymmetric but features a mixed reward structure in the long run, where only a fraction of entrepreneurs receive cuddly incentives. The third regime, which is analogous to part 3 of proposition 2, once again involves institutional and technological convergence across all countries in the long run.

Despite these similarities, there are noteworthy differences as well. First, the roles of the coefficient of relative risk aversion,  $\theta$ , and the willingness of the social planner to substitute consumption across cohorts,  $\lambda$ , are now more clearly delineated:  $\theta$  is relevant whenever the decision concerns individual entrepreneurs' risk-taking behavior, while  $\lambda$  regulates the social planner's intertemporal calculus. One consequence of this separation is

that, in contrast to the last part of proposition 3, an increase in  $\theta$  always makes an asymmetric equilibrium more likely when  $\gamma$  is small, not just when  $\theta < 1$ .<sup>13</sup> A more distinctive consequence of the separation between  $\theta$  and  $\lambda$  is that the institutional dynamics are now richer and feature mixed rewards that emerge and then disappear along the transition path (as opposed to proposition 2, where mixed rewards emerge only at the end of a transition path). The reason is that, instead of a single threshold  $\tilde{m}$ , there are now two critical thresholds,  $\tilde{m}_c$  and  $\tilde{m}_o > \tilde{m}_c$ ; and when  $g_o/g_c > \tilde{m}_o > \tilde{m}_c$ , a follower country that starts sufficiently behind the leader will first adopt cutthroat incentives, then switch to mixed rewards, and then finally permanently transition to a cuddly reward structure. To obtain the intuition for these richer dynamics, note that

$$\tilde{m}_o - \tilde{m}_c = (\lambda - \theta) \frac{g_c - g_o}{\rho + \phi g_c}.$$

This expression implies that the gap between  $\tilde{m}_o$  and  $\tilde{m}_c$  reflects the difference between the innovation rate under cutthroat and cuddly incentives as well as the gap between aversion to intergenerational inequality,  $\lambda$ , and aversion to intragenerational risks,  $\theta$ . Then in the intermediate region in the phase diagram between  $\tilde{m}_o$  and  $\tilde{m}_c$  (see fig. A5 in the Appendix), rather than immediately switching from the earlier phase of cutthroat incentives to cuddly incentives, the social planner prefers to smooth intertemporal inequality at the expense of maintaining some degree of intragenerational inequality.

Finally, when  $g_o/g_c < \tilde{m}_o$  and  $\tilde{m}_c < 1$ , the equilibrium involves mixed rewards in the long run as in part 2 of proposition 2. It can also be verified that when  $\theta = \lambda$ ,  $\tilde{m}_o = \tilde{m}_c = \tilde{m}$ , so that we recover proposition 2 exactly.

The next proposition summarizes this characterization more formally.

**PROPOSITION 6.** Suppose that country social planners maximize (8), the world technology frontier is given by (5), and assumptions 1, 2', 3', and 4 hold,  $\lambda > \theta$ , and  $\phi > (1 - \theta)(1 - \lambda)/(\lambda - \theta)$ . Let

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

and

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

where  $\omega_o$ ,  $\omega_c$ ,  $g_o$ , and  $g_c$  are given by (14) and (18). Then the world equilibrium is characterized as follows. The leader country  $\ell$  always chooses

<sup>13</sup> A proof is available on request.

cutthroat rewards, that is,  $u_\ell(t) = 1$  for all  $t$ . For each follower  $j \neq \ell$ , we have the following:

1. If

$$\frac{g_o}{g_c} > \tilde{m}_o > \tilde{m}_c, \tag{29}$$

there exist  $\underline{m}$ ,  $\underline{m}$ ,  $T$ , and  $T' \geq 0$  such that for  $n_j(0) < \underline{m}^{1/\phi}$ , 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  $n_j(T) = \underline{m}^{1/\phi}$  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  $n_j(t) \rightarrow (g_o/g_c)^{1/\phi}$ .

2. If

$$\frac{g_o}{g_c} < \tilde{m}_o \quad \text{and} \quad \tilde{m}_c < 1, \tag{30}$$

then there exist  $\underline{m}$  and  $\underline{m}$  such that for  $\underline{m}^{1/\phi} < n_j(0) < \underline{m}^{1/\phi}$ , the reward structure of country  $j$  is mixed (i.e.,  $u_j(t) \in (0, 1)$ ) for all  $t$ , and  $(n_j(t), u_j(t)) \rightarrow (m^*, u^*)$ . If  $n_j(0) < \underline{m}^{1/\phi}$ , 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

$$\tilde{m}_o > \tilde{m}_c > 1, \tag{31}$$

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  $t$ ).

*B. 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,” that is, 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 are 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  $\bar{\sigma} < 0$ ,  $\bar{\rho} > 0$ ,  $\bar{\theta} < 1$ , and  $\bar{\gamma} < 1$  such that when  $\sigma \in (\bar{\sigma}, 0)$ ,  $\rho \leq \bar{\rho}$ ,  $\theta > \bar{\theta}$ , 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  $\sigma$  has to be above some  $\bar{\sigma} < 0$  is also intuitive: if  $\sigma$  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 and thus removes the economic rationale for an asymmetric equilibrium in which the contributions of the more advanced economies to world technology enable the rest to choose cuddly incentives. Finally, the condition  $\theta > \bar{\theta}$  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).

## VI. Equilibrium under Domestic Political Constraints

In this section, we discuss how domestic political constraints can influence the world equilibrium. For simplicity, consider a world economy consisting of two countries,  $j$  and  $j'$ , starting out with the same technology level, that is,  $N_j(0) = N_{j'}(0)$ , and suppose that the world technology frontier is given by (5). Let us also relax assumption 4 so that asymmetric equilibria in which countries that initially start out behind later become the technology leader are possible.

Suppose that there are domestic political constraints in country  $j$ , for example, imposed by a social democratic party or a labor movement, lim-

iting inequality and preventing the ratio of rewards between successful and unsuccessful entrepreneurs. In particular, suppose that this ratio in country  $j$  cannot exceed some amount  $\zeta$ . 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  $\zeta < A$ , so that domestic political constraints make it impossible for country  $j$  to adopt a cutthroat strategy regardless of the strategy of country  $j'$ . Consequently, of the asymmetric equilibria, the ones in which country  $j$  adopts a cutthroat reward structure at some point disappear, and the unique equilibrium becomes the one in which country  $j'$  adopts the cutthroat strategy and country  $j$  chooses an egalitarian structure forever. However, from proposition 2 above, this implies that country  $j$  will now have higher discounted utility than country  $j'$ .

This simple example thus illustrates how domestic political constraints, which restrict the amount of inequality in society, can act as a strategic commitment device and create an advantage in world equilibrium. This example can be generalized straightforwardly. For instance, instead of having the two countries start out with identical technologies, the same result applies provided that their technology levels are not too dissimilar (so that either country becoming the technology leader is an equilibrium; see Acemoglu et al. 2012).

An interesting implication that follows from this example is that country  $j$ , which has a stronger social democratic party, benefits in terms of discounted utility 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 United States to Scandinavian economies, the latter clearly have a history of a 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 (e.g., Baldwin 1992; Friedman 2010).

## VII. Case Study Evidence from the Pharmaceutical Industry

In this section we discuss the industrial organization of the global pharmaceutical sector, one of the most innovation-intensive parts of the economy, and one of the high-tech sectors in which the United States appears to have a global lead.<sup>14</sup> We seek to establish four main claims,

<sup>14</sup> Besides pharmaceuticals, important high-tech sectors in which the United States appears to be the world technology leader include aerospace, communication equipment, and computer machinery. See Sciences and Engineering Indicators Reports (2002, 2012), chap. 6, “Industry, Technology and the Global Marketplace” at <https://wayback.archive-it.org/5902/20150818104216/http://www.nsf.gov/statistics/seind02/pdf/c06.pdf> and

which are consistent with the major assumptions and predictions of our model: (i) there are large cross-national spillovers in pharmaceutical research and development; (ii) there are large, persistent differences between the United States and other OECD countries in drug prices/markups, which result mostly from drug price controls in the latter set of countries and which, similarly to our distinction between cuddly and cutthroat reward structures, translate into different rewards to pharmaceutical innovations: (iii) “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” United States; and (iv) politicians in the United States and 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.

Innovation 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, Hansen, and Grabowski 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 also the direction of innovation in the sector strongly responds to profit incentives (Acemoglu and Linn 2004; Finkelstein 2004). R&D investments create cross-national spillovers because once a new drug is discovered, no marginal R&D costs are involved in bringing it to market in other countries (Danzon and Towse 2003). For this reason, the pharmaceutical industry is dominated by multinational companies, which market new drugs globally, either directly or indirectly through licensing. Because trade secrecy is not a viable option, pharmaceutical innovations are highly susceptible to free riding.

Consumers pay very different costs across countries for similar drugs, mainly as a result of government policy. Golec and Vernon (2006) show, using drug price indices, that between 1986 and 2004, inflation-adjusted 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 United States and Canada is a particularly stark illustration of this phenomenon. Quon, Firszt, and Eisenberg (2005) find that brand-name medications are approximately 24 percent cheaper on the websites of Canadian pharmaceutical retailers than on the websites of US drug chain pharmacies. These higher drug prices in the United States are widely argued to be a result of the congressional ban on the federal government negotiating drug prices with companies, whereas all other OECD countries have adopted drug price controls (Scherer 2004). Thus

---

<https://www.nsf.gov/statistics/seind12/c6/c6s2.htm>. Though the US share of world production in some of these sectors shows slight declines over the last two decades, this is starting from a very high base (aerospace) or is partly a consequence of rising Chinese production (computers).

cross-border price disparities are preserved via laws that prohibit the reimportation of prescription medicines from other countries, though there is a large illegal cross-border drug reimportation trade (Bhosle and Balakrishnan 2007).

The trade-off between “the affordability of drugs and technological progress” is widely acknowledged in the pharmaceutical industry (Scherer 1993, 2004). In our model, countries at the world technology frontier disproportionately contribute to the incentives for innovation while countries behind the frontier can free ride on these incentives and enjoy “cuddlier” domestic institutions. Although, given the multinational nature of pharmaceutical companies, the issue is difficult to assess empirically, there is *prima facie* evidence that US firms invest more in new drug development, presumably in part because they are more profitable.<sup>15</sup> Specifically, in 2006, the most recent year for which data are available, R&D divided by sales for the US pharmaceutical industry stood at 21.8 percent, while the same number was 10.4 percent for German firms, 8.7 percent for French firms, 1.5 percent for Italian firms, 6.2 percent for Spanish firms, 10 percent for Dutch firms, 11.1 percent for Swedish firms, and 23.1 percent for British firms (OECD 2010).<sup>16</sup> Golec and Vernon (2006) confirm that the US pharmaceutical industry is more profitable, with the profitability of US firms between 1993 and 2004 standing at 17.1 percent, while that of EU-based firms averaged 12.2 percent.

Various papers in the literature indeed link profitability and drug prices to innovation intensity. Danzon, Wang, and Wang (2005) analyze the launch of 85 new drugs in 25 major markets and find that even though new drugs reach all major markets, they are launched earlier in countries with higher expected drug prices, notably in the United States. Scherer (2001) finds that within the United States, R&D outlays closely track profits over time, in terms of both long-term trends and short-term deviations from trend, a finding echoed by Giaccotto, Santerre, and Vernon (2005). As a result, 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, where all major global drugs seek approval, were patented by US-based firms and research entities.

Finally, there is evidence that policy makers in the United States and European countries are aware of this discriminatory pricing practice but

<sup>15</sup> The degree of multinationalization of the pharmaceutical sector derives in particular from the fact that European firms also benefit from the more profitable US market. For example, Gambardella, Orsenigo, and Pammolli (2000) report that in 1999, US firms accounted for 60 percent of the North American market and 26.1 percent of the European market, while European firms had a market share of 24 percent in North America and 45.7 percent in Europe.

<sup>16</sup> The numbers reported by the OECD are surprisingly volatile, however, and in some years, somewhat implausibly, the US pharmaceuticals' R&D to sales ratio is reported to be as low as 8 percent.



view it as their best response to preserve it. This is seen clearly in the efforts in both the United States and Canada to prevent the equalization of drug prices via cross-border trade. Though more than 70 percent of Americans above 50 report that they would buy drugs from Canada if it were allowed (Choudhry and Detsky 2005), prescription drug reimportation is illegal in the United States (Bhosle and Balakrishnan 2007). The Canadian Pharmacists' Association also supports the ban of the exportation of drugs from Canada to the United States. Meanwhile, proposals in the United States that the federal government negotiate drug prices with pharmaceutical companies have consistently been defeated on the grounds that this would stifle innovation and competition (Moffitt 2013). In fact, while in other OECD countries governments utilize their bargaining power as managers of national health care schemes to negotiate and control drug prices, the United States did not implement drug price controls or negotiation as part of the Affordable Care Act. Though it is difficult to separate absence of price controls in the United States 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 United States than politicians in other OECD countries.

Overall, consistent with our model's emphasis, in the pharmaceutical sector, where global free riding on innovation is commonplace, there are major differences in innovation incentives, ranging from more "cutthroat" ones in the United States to more "cuddly" ones in many European markets, often supported by price control policies. Echoing our notion of a stable, asymmetric institutional equilibrium, policy makers in both cuddly and cutthroat systems appear to understand the economic forces that make this discriminatory pricing arrangement stable.

### VIII. Conclusion

In this paper, we have taken a first step toward 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 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, and thus greater inequality at the top of the income distribution, 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 in-



novation 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. This configuration is an equilibrium all the same, 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.

There are several research directions suggested by our paper. First, our work painted a stark picture of a world consisting of technological leaders operating under cutthroat incentives and followers adopting either cuddly or mixed reward systems. The reality is clearly more complex, with both considerable sectoral heterogeneity and major differences in innovation and productivity across European countries, with the Scandinavian countries exhibiting greater levels of R&D intensity and innovativeness than much of the rest of Europe. A fruitful direction would be to incorporate other dimensions of country and sectoral productivity in a model of interdependent world equilibrium. Second, and relatedly, as we have already noted, there are reasons why better risk sharing may also encourage entrepreneurship by insuring entrepreneurs against the worst risks. Future research could further investigate the optimal mix of risk sharing and incentives in a world equilibrium and whether the interplay of these two forces accounts for why some cuddly capitalists, like the Scandinavian ones, may have the lead in some sectors, while cutthroat incentives may remain necessary for technological leadership in the most high-tech sectors. Third, we have also ignored the importance of a safety net and redistribution in the supply of potential entrepreneurs (e.g., by generating a sufficient number of high human capital agents in the economy). Incorporating these considerations would also open the way to a richer analysis of the optimal design of redistribution and incentives for innovation in a world equilibrium. Fourth, another promising research direction might be to develop the theoretical ideas toward asymmetric world equilibria resulting from trade linkages pointed out in Section IV.H. Specifically, when different sectors benefit differentially from high-powered

incentives and risk sharing, and reward structures cannot be differentially calibrated for sectoral needs but must be chosen at the country level, specialization induced by trade might trigger institutional divergence. This channel not only is theoretically distinct but also makes different empirical predictions, for example, concerning the prerequisites of asymmetric institutional equilibria and the feedbacks from different institutions to sectoral productivity and investment.

Finally, we should emphasize that 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. Some possible avenues for empirical work include investigation of whether more high-powered incentives for entrepreneurs are associated with higher rates of innovation, greater numbers of high-impact patents, higher productivity, and greater export intensity in the most R&D-intensive industries (relative to other industries), and study of the related but distinct sectoral predictions of the aforementioned trade version of interdependent world equilibrium, where institutional choices depend on sectoral specialization.

## 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 intermediates with elasticity  $1/\beta$ , and thus each monopolist will charge a constant monopoly price of  $\psi/(1 - \beta)$ , where recall that  $\psi$  is the marginal cost in terms of the final good of producing any of the intermediates (given its blueprint, which is either 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  $p_j(v, t) = 1$  and  $x_j(v, t) = L_j = 1$  for all  $j, v$ , and  $t$ . This gives that total expenditure on intermediates 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). QED

*Proof of Proposition 2*

In the text, we presented a diagrammatic argument for proposition 2. Here, we provide a full derivation, also establishing uniqueness. For this purpose, we differentiate (21) and combine it with (22) to obtain

$$\dot{\Psi}(t) = [\rho - (1 - \theta)g_c + \phi g_c]\Psi(t) + (\omega_o - \omega_c)(\rho + \phi g_c)m(t)^{(1-\theta)/\phi-1}[\tilde{m} - m(t)]. \tag{A1}$$

Integrating (A1), we have

$$\Psi(t) = (\omega_o - \omega_c)(\rho + \phi g_c) \int_t^\infty e^{-[\rho - (1-\theta)g_c + \phi g_c](\tau-t)} m(\tau)^{(1-\theta)/\phi-1} [m(\tau) - \tilde{m}] d\tau. \tag{A2}$$

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

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

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  $1 > \tilde{m} > g_o/g_c$ . Then

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

with  $u^*$  given such that  $\tilde{m} = g(u^*)/g_c$  satisfies (20) in the text, that is,  $\Psi(t) = 0$ . (Observe that this equation is identical to [24] since  $m(t) = \tilde{m}$  if and only if  $\kappa(t) = \tilde{\kappa}$ , but there is no reason in this proof to adopt this second change of variable, which was useful for the diagrammatic analysis.) Moreover, when  $m(t) = \tilde{m}$ ,  $\dot{m}(t) = 0$ , and thus (20) holds at all dates. The resulting path defines an MRAP and is thus the unique global maximum, establishing part 2 of the proposition.

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

Finally, consider part 1, where  $\tilde{m} < g_o/g_c$ . If  $m(0) \geq g_o/g_c$ , then (22) implies that  $m(t) \geq g_o/g_c > \tilde{m}$  for all  $t$ . Hence (A2) implies that  $\Psi(t) > 0$  for all  $t$ , and thus  $u(t) = 0$  for all  $t$  (which also implies from [22] that  $m(t)$  is monotonically decreasing toward  $g_o/g_c$ ). This again defines an MRAP, yielding the desired result.

The only remaining case is the one in which  $\tilde{m} < g_o/g_c$  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  $m(t) = \tilde{m}$  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 prop-

osition, that is,  $u(t) = 1$  for all  $t \leq T$  and  $u(t) = 0$  for all  $t > T$  for some  $T$  (or conversely for  $n(t) \geq \bar{m}^{1/\phi}$  for some  $\bar{m}$ ).

LEMMA A1. When  $\tilde{m} < g_o/g_c$ , 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  $\tilde{m} < g_o/g_c$ . Therefore, there exists at least some interval in which  $u(t) = 0$ . Take  $[T_1, T_2]$  to be the first such interval.

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

Note first that  $\Psi(t)$  defined by (A2) 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

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

Once again because  $\tilde{m} < g_o/g_c$ ,  $u(t) = 1$  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

$$m(t) = m(T_2)e^{-\phi g_c(t-T_2)} + 1 - e^{-\phi g_c(t-T_2)} \quad \text{for } t \in [T_2, T_3]. \tag{A3}$$

Combining this with (A1), we have

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

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  $\tilde{m} < g_o/g_c < m(T_2) < m(T_1)$ , 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  $\tilde{m} < m(T_1) < m(T_2)$ , then  $\dot{\Psi}(T_1) < 0$  and  $\dot{\Psi}(T_2) < 0$ , leading to a contradiction.
2. If  $m(T_1) < m(T_2) < \tilde{m} < g_o/g_c$ , then  $\dot{\Psi}(T_1) > 0$  and  $\dot{\Psi}(T_2) > 0$ , yielding another contradiction.
3. If  $m(T_1) < \tilde{m} < m(T_2) < g_o/g_c$ , 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 (A3) 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.

QED

This lemma thus implies that in the case in which  $\tilde{m} < g_o/g_c$ , the equilibrium will involve  $u(t) = 1$  for all  $t \leq T$  and  $u(t) = 0$  for all  $t > T$ . Then from (22) evaluated with  $u(t) = 1$ , we define  $m(T) = \bar{m}$  to complete the proof of the proposition. QED

*Proof of Proposition 3*

The condition that  $\tilde{m} < 1$  can be written as

$$\tilde{m} = \frac{1 - \theta}{1 - \Omega} \frac{g_c - \Omega g_o}{\rho + \phi g_c} < 1, \tag{A4}$$

where

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

with  $A$  given as in equation (9). First note that  $\Omega$  does not depend on  $g_o$ , and thus a higher  $g_o$  reduces the left-hand side of (A4) and makes this inequality more likely to hold, establishing the second claim. Next  $\Omega$  also does not depend on  $\phi$ , and thus a higher  $\phi$  makes (A4) more likely to hold as well, establishing the second claim.

To prove the first claim, observe that

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

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

LEMMA A2. Suppose that assumption 1 holds. Then

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

*Proof.* Straightforward differentiation gives

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

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

and

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

QED

We next turn to the third claim. We have

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

$$\frac{\partial \Omega / \partial \theta}{\Omega} = -\log(1 - \gamma) - \frac{q_1 A^{1-\theta} \log A}{q_1 A^{1-\theta} + (1 - q_1)} - \log[q_1 A + (1 - q_1)] \\ + (1 - \theta) \frac{\partial A}{\partial \theta} \frac{q_1(1 - q_1)(A^{-\theta} - 1)}{[q_1 A^{1-\theta} + (1 - q_1)][q_1 A + (1 - q_1)]},$$

and

$$(1 - \theta) \frac{\partial A / \partial \theta}{A} = \frac{(q_1 - q_0)(1 - \gamma)^{1-\theta} \log(1 - \gamma)}{[(1 - q_0) - (1 - q_1)(1 - \gamma)^{1-\theta}][q_1(1 - \gamma)^{1-\theta} - q_0]} \\ + \frac{1}{1 - \theta} \{ \log[(1 - q_0) - (1 - q_1)(1 - \gamma)^{1-\theta}] \\ - \log[q_1(1 - \gamma)^{1-\theta} - q_0] \}.$$

Consider next a first-order Taylor expansion of  $A$  around  $\gamma = 0$ , which gives  $\log A \approx \gamma / (q_1 - q_0)$  and

$$(1 - \theta) \frac{1}{A} \frac{\partial A}{\partial \theta} \approx \text{constant} \cdot \gamma^2.$$

Therefore, ignoring second-order terms in  $\gamma$ , we have that around

$$\gamma = 0,$$

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

Therefore, there exists a value  $\bar{\gamma} > 0$  such that for  $\gamma < \bar{\gamma}$ ,  $(\partial \Omega / \partial \theta) / \Omega < 0$  and thus  $\partial \bar{m} / \partial \theta < 0$  when  $\theta < 1$ , establishing the third claim. QED

*Proof of Proposition 4*

Consider the case in which  $n_j(0) = 1$ . 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 discounted utility as country  $\ell$ , and thus at  $n_j(0) = 1$ , country  $j$  has strictly higher discounted utility than country  $\ell$ . Next by continuity, this is also true for  $n_j(0) > 1 - \delta$  for  $\delta$  sufficiently small and positive. QED

*Proof of Proposition 5*

To simplify notation, let us focus on the case with two countries,  $\ell = 1$  and  $j = 2$ . Suppose that there exists another equilibrium than the one characterized in prop-

osition 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 country 1 adopts  $u_1(t) = 0$  for some (positive measure) interval. 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 discounted utility if  $u_1(t) = 0$  for some interval.

Next consider the case in which 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

$$N_1(T) = N_2(T),$$

where

$$N_1(t) = N_1(0)e^{g_c t},$$

$$N_2(t) = \left\{ N_2(0)^\phi + \frac{g_c}{g_o} (e^{\phi g_o t} - 1) [N_1(0)]^\phi \right\}^{1/\phi},$$

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

$$T = \frac{1}{\phi g_o} \ln \left[ \frac{\frac{g_c}{g_o} - n(0)^\phi}{\frac{g_c}{g_o} - 1} \right]. \tag{A5}$$

Next consider the discounted utility of country  $\ell = 1$  at time  $t = 0$  in an equilibrium in which there is this most rapid switch of leadership from it to the lower country. This can then be written as

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

where  $\mathcal{W}_1(T)$  is the continuation utility of country 1 after  $T$  when the technology leadership shifts to country 2. Clearly  $\mathcal{W}_1(T) \leq e^{(1 - \theta)g_c T} \omega_o / [\rho - (1 - \theta)g_c]$ , in view of the fact that consumption has grown at the rate  $g_o$  until date  $T$  and can grow at most at the rate  $g_c$  thereafter, and  $\omega_o$  corresponds to the highest flow utility (resulting from cuddly incentives). So  $\mathcal{W}_1(0)$  can be upper-bounded by

$$\mathcal{W}_1^o(0) = \{1 - e^{-[\rho - (1 - \theta)g_c]T}\} \frac{\omega_o}{\rho - (1 - \theta)g_o} + e_1^{-[\rho - (1 - \theta)g_c]T} \frac{\omega_o}{\rho - (1 - \theta)g_c}.$$

In contrast, choosing cutthroat incentives from the beginning gives country 1 discounted utility

$$\mathcal{W}_1^e(0) = \frac{\omega_e}{\rho - (1 - \theta)g_c}.$$

The condition that  $\mathcal{W}_1^e(0) > \mathcal{W}_1^o(0)$  would be sufficient to rule out equilibria with the most rapid switch of leadership. In view of the previous two expressions, this is equivalent to

$$\begin{aligned} \frac{\omega_c}{\rho - (1 - \theta)g_c} &> \left\{ 1 - e^{-[\rho - (1 - \theta)g_c]T} \right\} \frac{\omega_o}{\rho - (1 - \theta)g_o} \\ &+ e_1^{-[\rho - (1 - \theta)g_c]T} \frac{\omega_o}{\rho - (1 - \theta)g_c}, \\ \frac{\omega_c}{\rho - (1 - \theta)g_c} &> \frac{\omega_o}{\rho - (1 - \theta)g_o} \\ &\times \left[ \frac{\rho - (1 - \theta)(e^{-[\rho - (1 - \theta)g_c]T} g_o + \{1 - e^{-[\rho - (1 - \theta)g_c]T}\} g_c)}{\rho - (1 - \theta)g_c} \right], \end{aligned}$$

or rearranging (and noting that  $(1 - \theta)\omega_o$  is always positive), it is equivalent to

$$T > T^* \equiv -\frac{1}{\rho - (1 - \theta)g_o} \ln G, \tag{A6}$$

where

$$G \equiv \frac{[\rho - (1 - \theta)g_o]\omega_c - [\rho - (1 - \theta)g_c]\omega_o}{(g_c - g_o)(1 - \theta)\omega_o}.$$

Assumption 3 ensures that  $G > 0$ . Moreover, it can be verified that  $G < 1$ , which ensures  $T^* > 0$ . To see this, suppose  $G \geq 1$ , which implies

$$[\rho - (1 - \theta)g_o]\omega_c - [\rho - (1 - \theta)g_c]\omega_o \geq (g_c - g_o)(1 - \theta)\omega_o,$$

or after canceling common terms from both sides and rearranging,

$$[\rho - (1 - \theta)g_o](\omega_c - \omega_o) > 0,$$

which is impossible.

Now combining (A6) with (A5) and rearranging, we have that

$$n(0) < \underline{n} \equiv \left\{ \frac{g_c - (g_c - g_o)G^{-\phi g_o/[\rho - (1 - \theta)g_c]}}{g_o} \right\}^{1/\phi} \tag{A7}$$

is thus sufficient for  $\mathcal{W}_1^c(0) > \mathcal{W}_1^n(0)$ . To ensure that  $\underline{n} > 0$ , it is sufficient to have  $\phi \in (0, \phi')$ , where

$$\phi' \equiv -\frac{\ln[g_c/(g_c - g_o)]/\ln G}{[\rho - (1 - \theta)g_o]/g_o}.$$

This establishes that for  $\phi < \phi'$  and  $n(0) < \underline{n}$ , there is no equilibrium in which country 1 chooses  $u_1(t) = 0$  for some time interval  $[0, T)$  until it loses the technology leadership.

We next show that there exist no equilibria in which country 1 chooses  $u_1(t) = 1$  for some time interval  $[0, T'')$  and then reverts to  $u_1(t) = 0$  at  $t = T''$  (and may or may not revert to  $u_1(t) = 1$  again). (The proof for the case in which country 1 first chooses  $u_1(t) = 0$  for some time interval  $[0, T')$  and then  $u_1(t) = 1$  over  $[T', T'')$  and then reverts to  $u_1(t) = 0$  at  $t = T''$ , with possibly further switches in the future, is also analogous and is omitted.)

This can be established by the following reasoning. First, suppose that  $u_2(t) = 0$  for all  $t \in [0, T'')$ . Then from (16),  $n(T'') \leq (g_o/g_c)^{1/\phi}$ . From (A7), for  $\phi$  sufficiently small, say  $\phi < \phi''$  for some  $\phi'' > 0$ ,  $\underline{n}$  is greater than  $(g_o/g_c)^{1/\phi}$ , and thus



$n(T'') \leq (g_o/g_c)^{1/\phi} < \underline{n}$ , so that the same argument as above can be applied: either there will be no switch in technology leadership after  $T''$ , in which case  $u_1(t) = 0$  (for a positive measure interval) following  $T''$  is suboptimal from proposition 2, or there will be a switch in technology leadership; but in that case we have that  $\mathcal{W}_1^i(T'') > \mathcal{W}_1^o(T'')$ , yielding a contradiction.

Second, suppose that  $u_2(t) = 1$  during some subinterval of  $[0, T'')$ , say  $[\hat{T}, T'') \subset [0, T'')$  (if there are many such intervals, take the first). Once again we need to focus on only the case in which there will be a switch in technology leadership, say at some date  $\tilde{T}$ . Suppose first that  $n(T''') < \underline{n}$ , but in that case with the same argument as above,  $\mathcal{W}_1^i(T''') > \mathcal{W}_1^o(T''')$ , yielding a contradiction. Next suppose that  $\underline{n} \leq n(T''') < 1$ . Let the discounted utility of country 2 when  $u_1(t) = 1$  for  $t \geq T'''$  and  $u_2(t) = 1$  for  $t \geq \tilde{T} > T'''$  be  $\mathcal{W}_2(T'''|\mathbf{u}_1, \mathbf{u}_2)$  (where  $\mathbf{u}_1$  and  $\mathbf{u}_2$  respectively denote  $u_1(t) = 1$  for  $t \geq T'''$  and  $u_2(t) = 1$  for  $t \geq \tilde{T} > T'''$ ), and let its discounted utility when we keep its strategy at  $\mathbf{u}_2$  but change  $u_1(t)$  to zero for some part of the time after  $T'''$ , represented by the strategy  $\mathbf{u}'_1$ , be  $\mathcal{W}_2^{\text{equilibrium switch}}(T'''|\mathbf{u}'_1, \mathbf{u}_2)$  (where the reference to “equilibrium switch” is due to the fact that the switch and technology leadership will take this form with some suitably chosen  $\mathbf{u}'_1$  that ensures such a switch). Since country 2 always benefits from cutthroat incentives in country 1, and since we are holding  $u_2$  constant and reducing  $u_1$ , we have that for any  $(\mathbf{u}_1, \mathbf{u}_2)$  as specified above,

$$\mathcal{W}_2^{\text{equilibrium switch}}(T'''|\mathbf{u}'_1, \mathbf{u}_2) \leq \mathcal{W}_2(T'''|\mathbf{u}_1, \mathbf{u}_2).$$

But moreover, from proposition 2, when  $u_1(t) = 1$  for  $t \geq T'''$ , the utility-maximizing strategy for country 2 is  $u_2(t) = 0$  for  $t \geq \hat{T} \geq T'''$  (where  $\hat{T} = T'''$  when  $\bar{m}^{1/\phi} \leq n(T''')$  and  $\hat{T} > T'''$  otherwise as specified in proposition 2), yielding discounted utility  $\mathcal{W}_2^o(T''')$ . From proposition 2 for any such  $(\mathbf{u}_1, \mathbf{u}_2)$ , we have

$$\mathcal{W}_2(T'''|\mathbf{u}_1, \mathbf{u}_2) < \mathcal{W}_2^o(T'''),$$

thus establishing that for any  $(\mathbf{u}'_1, \mathbf{u}_2)$  involving an equilibrium switch in leadership, we have

$$\mathcal{W}_2^{\text{equilibrium switch}}(T'''|\mathbf{u}'_1, \mathbf{u}_2) < \mathcal{W}_2^o(T'''),$$

so that it would be a profitable deviation for country 2 to switch to  $u_2(t) = 0$  and induce country 1 to remain the technology leader choosing  $u_1(t) = 1$ . Setting  $\underline{\phi} = \min\{\phi', \phi''\}$  completes the proof of the proposition. QED

*Proof of Proposition 6*

First consider country  $\ell$ , 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, that is,  $N(t) = N_\ell(t)$ . Then (16) implies that

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

Then assumption 3 implies that country  $\ell$  always prefers fully cutthroat incentives, that is,  $u_\ell(t) = 1$ .

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

$$\begin{aligned} & \max_{u(t)} \frac{1}{1-\lambda} \int_0^\infty e^{-\rho t} [(1-\theta)\omega(u(t))]^{(1-\lambda)/(1-\theta)} n(t)^{1-\lambda} N_\ell(t)^{1-\lambda} dt \\ & \text{such that } \frac{\dot{n}(t)}{n(t)} = g(u(t))n(t)^{-\phi} - g_c, \\ & \dot{N}_\ell(t) = g_c N_\ell(t), \\ & n(0) = N(0)/N_\ell(0) \text{ given.} \end{aligned} \tag{A8}$$

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 (A8) can equivalently be written as

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

The current-value Hamiltonian for this problem can be written as

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

Consider the candidate solution given by the maximum principle, that is, as a solution to the following equations:

$$\begin{aligned} \frac{\partial H}{\partial u} &= \Psi(t) \equiv (\omega_c - \omega_o)[(1-\theta)\omega(u(t))]^{(\theta-\lambda)/(1-\theta)} m(t)^{(1-\lambda)/\phi} + \mu(t)\phi(g_c - g_o) \\ &= 0 \quad \text{for } 0 \leq u(t) \leq 1, \\ \dot{m}(t) &= \phi[g(u(t)) - g_c m(t)], \\ \dot{\mu}(t) &= [\rho - (1-\lambda)g_c + \phi g_c]\mu(t) - \frac{1}{\phi} [(1-\theta)\omega(u(t))]^{(1-\lambda)/(1-\theta)} m(t)^{[(1-\lambda)/\phi]-1}, \end{aligned} \tag{A9}$$

together with the transversality condition, which takes the form

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

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

We next introduce the same change of variable as in the text,  $\kappa(t) \equiv \mu(t)m(t)^{(\lambda-1)/\phi}$ , so that

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

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

$$u(\kappa) = \begin{cases} 1 & \text{if } \kappa \geq \tilde{\kappa}_1 \\ u^*(\kappa) & \text{if } \tilde{\kappa}_1 > \kappa > \tilde{\kappa}_0 \\ 0 & \text{if } \kappa \leq \tilde{\kappa}_0, \end{cases} \tag{A10}$$

where

$$\begin{aligned} \tilde{\kappa}_1 &\equiv \frac{(\omega_c - \omega_o)[(1 - \theta)\omega_c]^{(\theta-\lambda)/(1-\theta)}}{\phi(g_c - g_o)}, \\ \tilde{\kappa}_0 &\equiv \frac{(\omega_c - \omega_o)[(1 - \theta)\omega_o]^{(\theta-\lambda)/(1-\theta)}}{\phi(g_c - g_o)}, \\ u^*(\kappa) &\equiv \frac{\omega_o}{\omega_o - \omega_c} - \frac{1}{1 - \theta} \frac{1}{\omega_o - \omega_c} \left[ \frac{\phi(g_c - g_o)}{\omega_o - \omega_c} \kappa \right]^{(1-\theta)/(\theta-\lambda)}. \end{aligned} \tag{A11}$$

Notice also that this derivation slightly generalizes the derivation of (24) in the text.

Then substituting for (A10) into the differential equations for  $\dot{m}$  and  $\dot{\kappa}$  and using the definition of  $\kappa$ , we obtain the law of motion of the system consisting of  $(m, \kappa)$  as

$$\begin{aligned} \dot{\kappa} &= [\rho - (1 - \lambda)g_c + \phi g_c] \kappa - \frac{1}{\phi} [(1 - \theta)\omega u(\kappa)]^{(1-\lambda)/(1-\theta)} \frac{1}{m} \\ &\quad - (1 - \lambda)[g_o + (g_c - g_o)u(\kappa)] \frac{\kappa}{m}, \\ \dot{m} &= \phi[g_o + (g_c - g_o)u(\kappa) - g_c m], \end{aligned} \tag{A12}$$

with  $u(\kappa) = 0$  when  $\kappa \leq \tilde{\kappa}_0$ ,  $u(\kappa) = 1$  when  $\kappa \geq \tilde{\kappa}_1$ , and  $u(\kappa) = u^*(\kappa) \in (0, 1)$  for  $\tilde{\kappa}_1 > \kappa > \tilde{\kappa}_0$ . In particular in the last regime where  $u(\kappa) = u^*(\kappa)$ , after substitution of (A11) the system can be rewritten as

$$\begin{aligned} \dot{\kappa} &= (\rho + \phi g_c) \kappa - \frac{\lambda - \theta}{1 - \theta} \left( \frac{g_c - g_o}{\omega_o - \omega_c} \right)^{(1-\lambda)/(\theta-\lambda)} \phi^{(1-\theta)/(\theta-\lambda)} \frac{\kappa^{(1-\lambda)/(\theta-\lambda)}}{m} \\ &\quad - (1 - \lambda) \frac{\omega_o(g_c - g_o) + g_o(\omega_o - \omega_c)}{\omega_o - \omega_c} \frac{\kappa}{m}, \\ \dot{m} &= \phi \left[ \frac{\omega_o(g_c - g_o) + (\omega_o - \omega_c)g_o}{\omega_o - \omega_c} - \frac{1}{1 - \theta} \left( \frac{g_c - g_o}{\omega_o - \omega_c} \right)^{(1-\lambda)/(\theta-\lambda)} \right. \\ &\quad \left. \times \phi^{(1-\theta)/(\theta-\lambda)} \kappa^{(1-\theta)/(\theta-\lambda)} - g_c m \right]. \end{aligned} \tag{A13}$$

More specifically, in a phase diagram in the  $(m, \kappa)$  space, the  $\dot{m} = 0$  locus is given as shown in figure A1, with the curve linking  $m = g_o/g_c$  (which applies when  $\kappa \leq \tilde{\kappa}_0$  and thus  $u(\kappa) = 0$ ) to  $m = 1$  (which applies when  $\kappa \geq \tilde{\kappa}_1$  and thus  $u(\kappa) = 1$ )

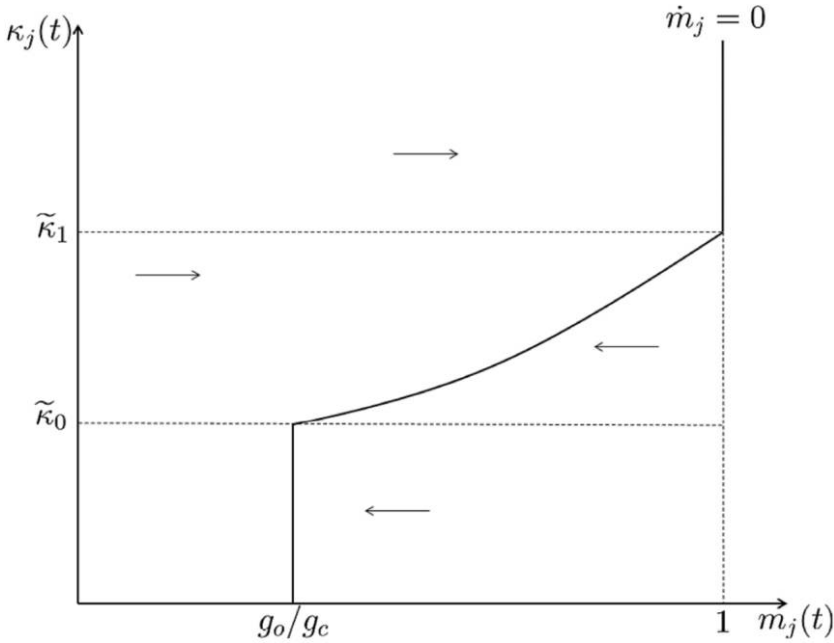


FIG. A1.—Construction of phase diagram for proposition 6:  $\dot{m} = 0$  locus

given from (A12) 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.

The  $\dot{\kappa} = 0$  locus in the  $(m, \kappa)$  space is derived similarly from (A12), and thus shown in figure A2, 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 A3 corresponds to the case in which  $g_o/g_c > \tilde{m}_o > \tilde{m}_c$ . The two curves for  $\dot{m} = 0$  and  $\dot{\kappa} = 0$  intersect where  $m = g_o/g_c$ . The laws of motion of  $(m, \kappa)$  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 follows 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 A4 corresponds to the case in which  $g_o/g_c < \tilde{m}_o$  and  $\tilde{m}_c < 1$ , 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. These cases together establish the claims in part 2.

Finally, figure A5 corresponds to the case in which  $\tilde{m}_o > \tilde{m}_c > 1$ , 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  $\theta = \lambda$ ,  $\tilde{\kappa}_0 = \tilde{\kappa}_1$ , and the phase diagrams simplify to figures 1, 2, and 3 in the text.

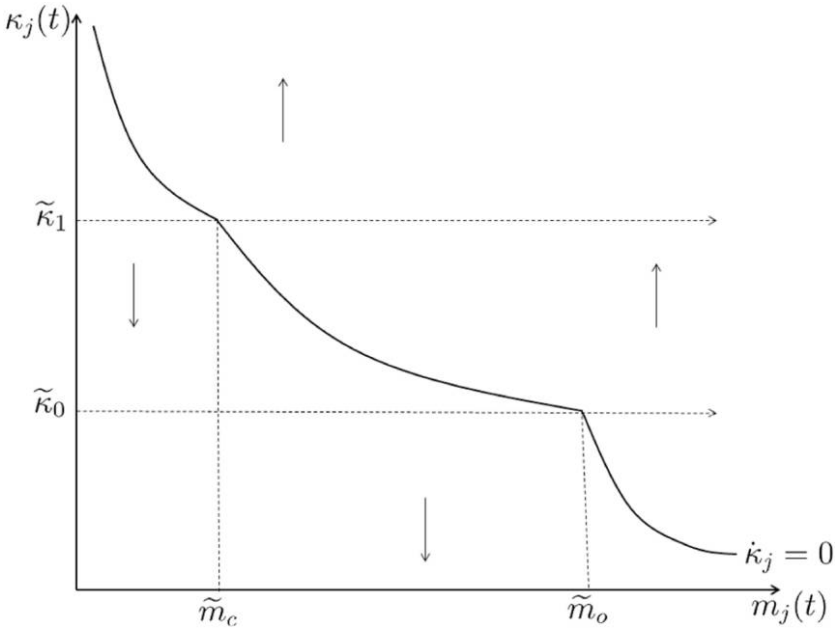


FIG. A2.—Construction of phase diagram for proposition 6:  $\dot{\kappa} = 0$  locus

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

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

where recall that  $\omega_c/\omega_o \leq 1$  when  $0 \leq 1$ . The Mangasarian sufficiency condition, the joint concavity of  $H$ , is equivalent to  $\partial^2 H/\partial u^2 \leq 0$ ,  $\partial^2 H/\partial n^2 \leq 0$ , 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 the conditions are satisfied. Some algebra establishes that the third one is also satisfied provided that  $\phi > [(1 - \theta)(1 - \lambda)]/(\lambda - \theta)$ . QED

*Proof of Proposition 7*

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

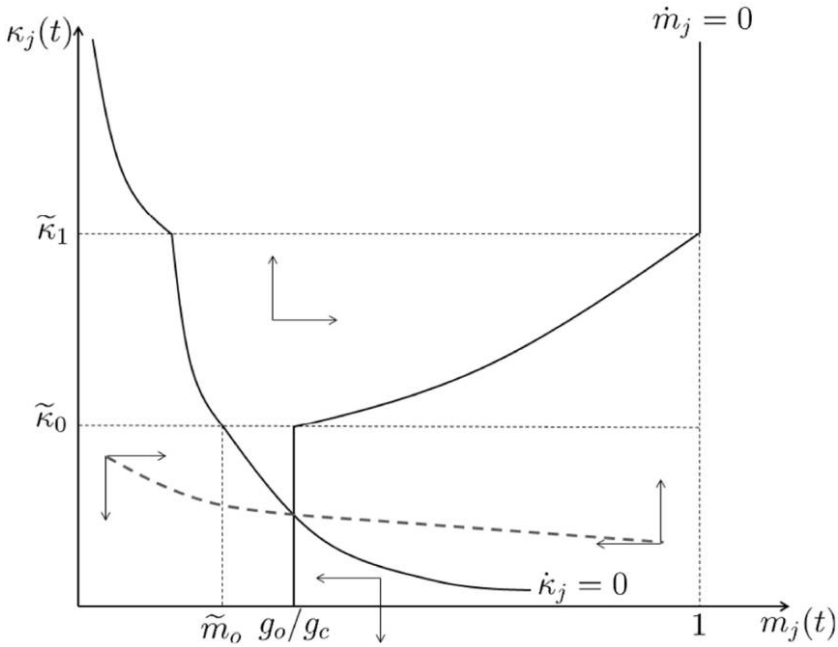


FIG. A3.—Phase diagram for proposition 6: case  $g_o/g_c > \tilde{m}_o > \tilde{m}_c$

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  $t$  utility of country  $j$  in this equilibrium can be written as

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

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

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

Consider now a deviation of one country  $k$  to the cutthroat reward structure at all times  $t > T_1$ . Denote by  $\hat{N}_j(t)$  the new growth path of country  $j$  and by  $\hat{N}(t)$  the

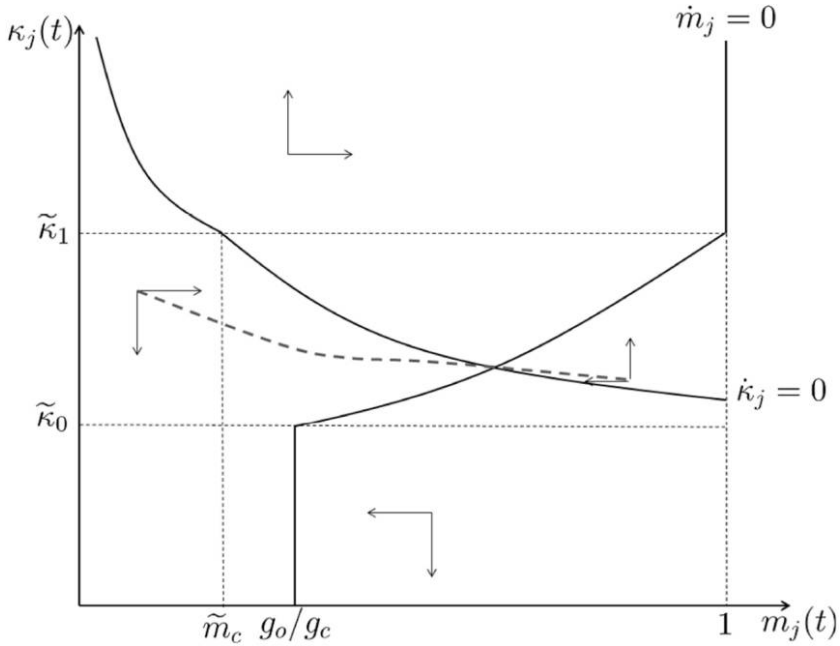


FIG. A4.—Phase diagram for proposition 6: case  $g_o/g_c < \tilde{m}_o$  and  $\tilde{m}_c < 1$

growth path to the world technology frontier. The world economy converges again to a new BGP with growth rate  $\hat{g}$ . This BGP growth rate can be written as

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

After this deviation, we have  $\hat{N}_k(t) > N_k(t)$  and  $\hat{N}_k(t) > \hat{N}(t)$  for all  $t > T_1$ . Then for  $\epsilon_1 > 0$ , there exists  $T'_1 > T_1$  such that for all  $t > T'_1$ ,  $\hat{N}_k/\hat{N}_k \geq \hat{g} - \epsilon_1$ , and discounted utility of country  $k$  satisfies

$$\begin{aligned} \mathcal{W}_k^c(T_1) &= \int_{T_1}^{\infty} e^{-\rho(t-T_1)} \omega_c \hat{N}_k(t)^{1-\theta} dt \\ &= \int_{T_1}^{T'_1} e^{-\rho(t-T_1)} \omega_c \hat{N}_k(t)^{1-\theta} dt + e^{-\rho(T'_1-T_1)} \int_{T'_1}^{\infty} e^{-\rho(t-T'_1)} \omega_c \hat{N}_k(t)^{1-\theta} dt \\ &> e^{-\rho(T'_1-T_1)} \omega_c \frac{\hat{N}_k(T'_1)^{1-\theta}}{\rho - (1-\theta)(\hat{g} - \epsilon_1)}. \end{aligned}$$

Now using the fact that  $\hat{N}_k(T'_1) \geq N_k(T'_1) \geq e^{g_o(T'_1-T_1)} N_k(T_1)$ , a sufficient condition for the deviation for country  $k$  to be profitable is

$$\begin{aligned} e^{-[\rho-(1-\theta)g_o](T'_1-T_1)} \omega_c \frac{N_k(T_1)^{1-\theta}}{\rho - (1-\theta)(\hat{g} - \epsilon_1)} &> \omega_o \frac{N_k(T_1)^{1-\theta} (1 + \epsilon)^{1-\theta}}{\rho - (1-\theta)(g_o + \epsilon)} \\ &> \mathcal{W}_k^o(T_1) = \int_{T_1}^{\infty} e^{-[\rho-(1-\theta)g_o](t-T_1)} \omega_o N_k(t)^{1-\theta} dt. \end{aligned}$$

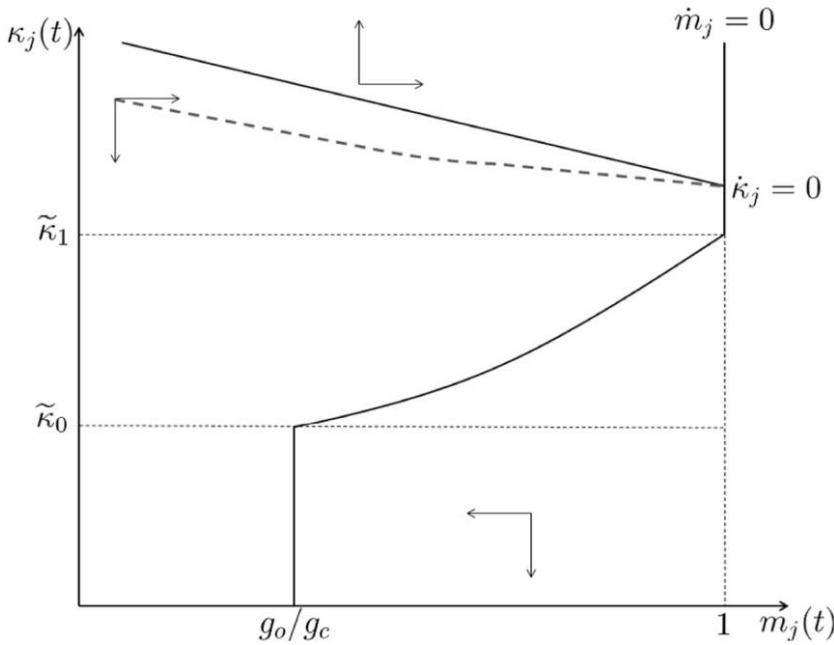


FIG. A5.—Phase diagram for proposition 6: case  $\tilde{m}_0 > \tilde{m}_c > 1$

Rearranging terms, this can be written as

$$\left(\frac{\omega_c}{\omega_o}\right)^{1/(1-\theta)} > (1 + \epsilon)e^{[\rho - (1-\theta)g_c]/(1-\theta)(T_1 - T_1)} \left[\frac{\rho - (1-\theta)(\hat{g} - \epsilon_1)}{\rho - (1-\theta)(g_0 + \epsilon)}\right]^{1/(1-\theta)}. \tag{A14}$$

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  $\tilde{N}_j(t)$  and the path of world technology frontier by  $\tilde{N}(t)$ . Clearly,  $\dot{\tilde{N}}(t)/\tilde{N}(t) = \tilde{g} < g_c$ , and moreover  $\tilde{N}_k(t) \leq N_k(t)$  for all  $t > T_2$ . Let us also note that

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

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



$$\begin{aligned}
 W_k^o(T_2) &= \int_{T_2}^{\infty} e^{-\rho(t-T)} \omega_o \tilde{N}_k(t)^{1-\theta} dt \\
 &= \int_{T_2}^{T_2'} e^{-\rho(t-T_2)} \omega_o \tilde{N}_k(t)^{1-\theta} dt + e^{-\rho(T_2'-T_2)} \int_{T_2'}^{\infty} e^{-\rho(t-T_2')} \omega_o \tilde{N}_k(t)^{1-\theta} dt \\
 &> \omega_o N_k(T_2)^{1-\theta} \int_{T_2}^{T_2'} e^{-\rho(t-T)} e^{(1-\theta)g_o(t-T)} dt + e^{-\rho(T_2'-T_2)} \omega_o \frac{N_k(T_2)^{1-\theta} e^{(1-\theta)g_o(T_2'-T_2)}}{\rho - (1-\theta)(\tilde{g} - \epsilon_2)} \\
 &> \omega_o N_k(T_2)^{1-\theta} \frac{1 - e^{-[\rho - (1-\theta)g_o](T_2'-T_2)}}{\rho - (1-\theta)g_o} \\
 &\quad + e^{-[\rho - (1-\theta)g_o](T_2'-T_2)} \omega_o \frac{N_k(T_2)^{1-\theta}}{\rho - (1-\theta)(\tilde{g} - \epsilon_2)},
 \end{aligned}$$

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

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

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

$$\frac{1 - \epsilon}{1 + \epsilon} \left[ \frac{\rho - (1-\theta)(g_c + \epsilon)}{\rho - (1-\theta)(\tilde{g} - \epsilon_2)} \right]^{1/(1-\theta)} \exp \left[ -\frac{\rho - (1-\theta)g_o}{1-\theta} (T_2' - T_2) \right] > \left( \frac{\omega_c}{\omega_o} \right)^{1/(1-\theta)}. \tag{A15}$$

Thus combining (A14) and (A15), we obtain that the following is a sufficient condition for an asymmetric equilibrium not to exist after some time  $T = \max\{T_1, T_2\}$ :

$$\begin{aligned}
 &\frac{1 - \epsilon}{1 + \epsilon} \left[ \frac{\rho - (1-\theta)(g_c + \epsilon)}{\rho - (1-\theta)(\tilde{g} - \epsilon_2)} \right]^{1/(1-\theta)} \exp \left[ -\frac{\rho - (1-\theta)g_o}{1-\theta} (T_2' - T_2) \right] > \left( \frac{\omega_c}{\omega_o} \right)^{1/(1-\theta)} \\
 &> (1 + \epsilon) \exp \left[ \frac{\rho - (1-\theta)g_o}{1-\theta} (T_1' - T_1) \right] \left[ \frac{\rho - (1-\theta)(\hat{g} - \epsilon_1)}{\rho - (1-\theta)(g_o + \epsilon)} \right]^{1/(1-\theta)}.
 \end{aligned} \tag{A16}$$

Now note that as  $\sigma \uparrow 0$  in (6),  $\hat{g} \rightarrow g_c$  and  $\tilde{g} \rightarrow g_c$ . Therefore, for  $\epsilon' > 0$ , there exists  $\bar{\sigma} < 0$  such that for  $\sigma > \bar{\sigma}$ ,  $\hat{g} - \epsilon' < g_c$  and  $\tilde{g} - \epsilon' < g_c$ . Thus choosing  $\epsilon, \epsilon_1, \epsilon_2$ , and  $\epsilon'$  sufficiently small, the following is also a sufficient condition:

$$\begin{aligned}
 &\exp \left[ -\frac{\rho - (1-\theta)g_o}{1-\theta} (T_2' - T_2) \right] > \left( \frac{\omega_c}{\omega_o} \right)^{1/(1-\theta)} \\
 &> \exp \left[ \frac{\rho - (1-\theta)g_o}{1-\theta} (T_1' - T_1) \right] \left[ \frac{\rho - (1-\theta)g_c}{\rho - (1-\theta)g_o} \right]^{1/(1-\theta)}.
 \end{aligned} \tag{A17}$$

Consider first the case  $\theta < 1$ . Choosing  $\rho$  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

$$e^{-(g-g_c)\bar{T}} > \left(\frac{\omega_c}{\omega_o}\right)^{1/(1-\theta)} > e^{(g-g_c)\bar{T}} \left[\frac{\rho - (1 - \theta)g_c}{\rho - (1 - \theta)g_o}\right]^{1/(1-\theta)}. \tag{A18}$$

For given choices of  $\epsilon$  and  $\epsilon_1$ ,  $\bar{T}$  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

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

Denote  $\tilde{\gamma} = 1 - (q_0/q_1)^{1/(1-\theta)}$ ; then

$$\left(\frac{\omega_c}{\omega_o}\right)^{1/(1-\theta)} \rightarrow \frac{q_0^{1/(1-\theta)}}{q_1}$$

as  $\gamma \rightarrow \tilde{\gamma}$ . This in turn approaches zero as  $\theta \rightarrow 1$ . Hence there exists  $\bar{\theta} < 1$  and  $\tilde{\gamma} < 1$  such that for  $\theta > \bar{\theta}$  and  $\gamma > \tilde{\gamma}$ ,  $(\omega_c/\omega_o)^{1/(1-\theta)}$  is sandwiched between these two terms, ensuring that (A18) is satisfied and a symmetric equilibrium does not exist.

When  $\theta > 1$ , (A17) can be rewritten as

$$\begin{aligned} & \exp\left[\frac{\rho + (\theta - 1)g_o}{\theta - 1}(T'_2 - T_2)\right] > \left(\frac{\omega_o}{\omega_c}\right)^{1/(\theta-1)} \\ & > \exp\left[-\frac{\rho + (\theta - 1)g_o}{\theta - 1}(T'_1 - T_1)\right] \left[\frac{\rho + (\theta - 1)g_o}{\rho + (\theta - 1)g_c}\right]^{1/(\theta-1)}. \end{aligned}$$

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

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

Now the first inequality is satisfied as  $e^{g\bar{T}} > 1 > (\omega_o/\omega_c)^{1/(\theta-1)}$  as  $\theta > 1$  and  $\omega_c < \omega_o < 0$ . The second inequality is satisfied when  $\omega_o/\omega_c > g_o/g_c$ , ensuring therefore again that (A18) is satisfied and a symmetric equilibrium does not exist.

Finally, when these conditions are satisfied, an analysis similar to that in the proof of proposition 2 implies that the equilibrium will take the form in which after some  $T$ , a subset of countries choose a cuddly reward structure and the remainder choose a cutthroat reward structure. QED

**References**

Acemoglu, Daron. 2003. "Cross-Country Inequality Trends." *Econ. J.* 113 (February): 121–49.  
 ———. 2009. *Introduction to Modern Economic Growth*. Princeton, NJ: Princeton Univ. Press.

- Acemoglu, Daron, Philippe Aghion, and Fabrizio Zilibotti. 2006. "Distance to Frontier, Selection, and Economic Growth." *J. European Econ. Assoc.* 4 (March): 37–74.
- Acemoglu, Daron, Ufuk Akcigit, and Murat Alp Celik. 2014. "Young, Restless and Creative: Openness to Disruption and Creativity of Innovations." Working Paper no. 19894 (February), NBER, Cambridge, MA.
- Acemoglu, Daron, and Joshua Linn. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Sector." *Q.J.E.* 119 (August): 1049–90.
- Acemoglu, Daron, and Jorn-Steffen Pischke. 1998. "Why Do Firms Train? Theory and Evidence." *Q.J.E.* 113 (February): 78–118.
- Acemoglu, Daron, James A. Robinson, and Thierry Verdier. 2012. "Can't We All Be More like Scandinavians? Asymmetric Growth and Institutions in an Interdependent World." Working paper no. 18441 (October), NBER, Cambridge, MA.
- Acemoglu, Daron, and Jaume Ventura. 2002. "The World Income Distribution." *Q.J.E.* 117 (May): 659–94.
- Alesina, Alberto, Edward L. Glaeser, and Bruce Sacerdote. 2001. "Why Doesn't the United States Have a European-Style Welfare State?" *Brookings Paper Econ. Activity* (Fall): 187–278.
- Alesina, Alberto, and Eliana La Ferrara. 2005. "Preferences for Redistribution in the Land of Opportunities." *J. Public Econ.* 89 (June): 897–931.
- Algan, Yann, Pierre Cahuc, and Marc Sangnier. 2011. "Efficient and Inefficient Welfare States." IZA Discussion Paper no. 5445 (January), Inst. Study Labor, Bonn.
- Aoki, Masahiko. 2001. *Towards a Comparative Institutional Analysis*. Cambridge, MA: MIT Press.
- Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. "Top Incomes in the Long Run of History." *J. Econ. Literature* 49 (March): 3–71.
- Baldwin, Peter. 1992. *The Politics of Social Solidarity: Class Bases of the European Welfare State, 1875–1975*. New York: Cambridge Univ. Press.
- Bénabou, Roland J. M. 2000. "Unequal Societies: Income Distribution and the Social Contract." *A.E.R.* 90 (March): 96–129.
- . 2006. "Inequality, Technology, and the Social Contract." In *Handbook of Economic Growth*, edited by Philippe Aghion and Steven N. Durlauf, 1595–1638. Amsterdam: North-Holland.
- Bénabou, Roland J. M., and Efe Ok. 2001. "Social Mobility and the Demand for Redistribution: The POUM Hypothesis." *Q.J.E.* 116 (May): 447–87.
- Bénabou, Roland J. M., and Jean Tirole. 2006. "Belief in a Just World and Redistributive Politics." *Q.J.E.* 121 (May): 699–746.
- Bergstrom, Ted, Lawrence Blume, and Hal Varian. 1986. "On the Private Provision of Public Goods." *J. Public Econ.* 29 (1): 25–49.
- Bhosle, Monali, and Rajesh Balakrishnan. 2007. "Drug Reimportation Practices in the United States." *Therapeutics and Clinical Risk Management* 3 (March): 41–46.
- Bottazzi, Laura, and Giovanni Peri. 2003. "Innovation and Spillovers in Regions: Evidence from European Patent Data." *European Econ. Rev.* 47 (August): 687–710.
- Brezis, Elise S., Paul R. Krugman, and Daniel Tsiddon. 1993. "Leapfrogging in International Competition: A Theory of Cycles in National Technological Leadership." *A.E.R.* 83 (December): 1211–19.
- Cameron, David R. 1978. "The Expansion of the Public Economy: A Comparative Analysis." *American Polit. Sci. Rev.* 72 (December): 1243–61.
- Chatterjee, Arpita. 2010. "Why Do Similar Countries Choose Different Policies? Endogenous Comparative Advantage and Welfare Gains." Manuscript, Princeton Univ.

- Choudhry, Niteesh, and Allan Detsky. 2005. "A Perspective on US Drug Reimportation." *J. American Medical Assoc.* 293 (January 19): 358–62.
- Coe, David, and Elhanan Helpman. 1995. "International R&D Spillovers." *European Econ. Rev.* 39 (May): 859–87.
- Cornes, Richard, and Roger Hartley. 2007. "Aggregative Public Good Games." *J. Public Econ. Theory* 9 (2): 201–19.
- Crouch, Colin. 2009. "Typologies of Capitalism." In *Debating Varieties of Capitalism: A Reader*, edited by Bob Hancké. New York: Oxford Univ. Press.
- Danzon, Patricia. 1997. "Price Discrimination for Pharmaceuticals: Welfare Effects in the US and the EU." *Internat. J. Econ. Bus.* 4 (3): 301–22.
- Danzon, Patricia, and Adrian Towse. 2003. "Differential Pricing for Pharmaceuticals: Reconciling Access, R&D and Patents." *Internat. J. Health Care Finance and Econ.* 3 (September): 183–205.
- Danzon, Patricia, Richard Wang, and Liang Wang. 2005. "The Impact of Price Regulation on the Launch Delay of New Drugs: Evidence from Twenty-Five Major Markets in the 1990s." *Health Econ.* 14 (March): 269–92.
- Deyo, Frederick C. 1993 *Beneath the Miracle: Labor Subordination in the New Asian Industrialism*. Berkeley: Univ. California Press.
- DiMasi, Joseph, Ronald Hansen, and Henry Grabowski. 2003. "The Price of Innovation: New Estimates of Drug Development Costs." *J. Health Econ.* 22 (March): 151–85.
- Epstein, Larry G., and Stanley E. Zin. 1989. "Substitution, Risk Aversion and Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57 (July): 937–69.
- Esping-Anderson, Gösta. 1990. *The Three Worlds of Welfare Capitalism*. Princeton, NJ: Princeton Univ. Press.
- Finkelstein, Amy. 2004. "Static and Dynamic Effects of Health Care Policy: Evidence from the Vaccine Industry." *Q.J.E.* 119 (May): 527–64.
- Friedman, Gerald. 2010. "Labor Unions in the United States." <http://eh.net/encyclopedia/labor-unions-in-the-united-states/>.
- Gambardella, Alfonso, Luigi Orsenigo, and Fabio Pammolli. 2000. "Global Competition in Pharmaceuticals: A European Perspective." Report, Enterprise Directorate-General of the European Comm., Brussels.
- Gerschenkron, Alexander. 1962. *Economic Backwardness in Historical Perspective: A Book of Essays*. Cambridge, MA: Belknap Press/Harvard Univ. Press.
- Giaccotto, Carmelo, Rexford Santerre, and John Vernon. 2005. "Drug Prices and Research and Development Investment Behavior in the Pharmaceutical Industry." *J. Law and Econ.* 48 (April): 195–214.
- Global Entrepreneurship Monitor. 2015. "Global Entrepreneurship Monitor Report 2015/2016." <http://www.gemconsortium.org/report>.
- Golec, Joseph, and John Vernon. 2006. "European Pharmaceutical Price Regulation, Firm Profitability, and R&D Spending." Working Paper no. 12676 (November), NBER, Cambridge, MA.
- Griffith, Rachel, Stephen Redding, and John Van Reenen. 2003. "R&D and Absorptive Capacity: Theory and Empirical Evidence." *Scandinavian J. Econ.* 105 (March): 99–118.
- Guimaraes, Bernardo, and Kevin D. Sheedy. 2017. "Political Specialization." Discussion Paper no. 11880, Centre Econ. Policy Res., London.
- Hall, Peter, and David Soskice. 2001. *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. New York: Oxford Univ. Press.
- Hayek, Friedrich von. 2006. *The Constitution of Liberty*. Classics Edition. London: Routledge.

- Hölmstrom, Bengt. 1979. "Moral Hazard and Observability." *Bell J. Econ.* 10 (April): 74–91.
- Howitt, Peter. 2000. "Endogenous Growth and Cross-Country Income Differences." *A.E.R.* 90 (September): 829–46.
- Keller, Wolfgang. 2001. "Knowledge Spillovers at the World's Technology Frontier." Discussion Paper no. 2815, Centre Econ. Policy Res., London.
- Kneller, Robert. 2010. "The Importance of New Companies for Drug Discovery: Origins of a Decade of New Drugs." *Nature Revs. Drug Discovery* 9 (November): 867–82.
- Krueger, Dirk, and Krishna Kumar. 2004. "US-Europe Differences in Technology-Driven Growth: Quantifying the Role of Education." *J. Monetary Econ.* 51 (January): 161–90.
- Krugman, Paul, and Anthony Venables. 1996. "Integration, Specialization, and Adjustment." *European Econ. Rev.* 40 (April): 959–67.
- Landier, Agustin. 2005. "Entrepreneurship and the Stigma of Failure." Manuscript, Soc. Sci. Res. Network. <https://ssrn.com/abstract=850446>.
- Levchenko, Andrei. 2007. "Institutional Quality and International Trade." *Rev. Econ. Studies* 74 (July): 791–819.
- Lipset, Seymour M. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Polit. Sci. Rev.* 53 (March): 69–105.
- Marmor, Theodore R., and Jacob S. Hacker. 2005. "Medicare Reform and Social Insurance: The Clashes of 2003 and Their Potential Fallout." *Yale J. Health Policy, Law, and Ethics* 5 (Winter): 475–90.
- Matsuyama, Kiminori. 2002. "Explaining Diversity: Symmetry-Breaking in Complementarity Games." *A.E.R.* 92 (May): 241–46.
- . 2004. "Financial Market Globalization, Symmetry-Breaking, and Endogenous Inequality of Nations." *Econometrica* 72 (May): 853–84.
- . 2005. "Symmetry-Breaking." In *The New Palgrave Dictionary of Economics*, 2nd ed., edited by L. Blume and S. Durlauf. London: Palgrave Macmillan.
- . 2007. "Aggregate Implications of Credit Market Imperfections." In *NBER Macroeconomics Annual 2007*, edited by D. Acemoglu, K. Rogoff, and M. Woodford. Chicago: Univ. Chicago Press.
- . 2013. "Endogenous Ranking and Equilibrium Lorenz Curve across (Ex Ante) Identical Countries." *Econometrica* 81 (September): 2009–31.
- Milesi-Ferretti, Gian Maria, Roberto Perotti, and Massimo Rostagno. 2002. "Electoral Systems and Public Spending." *Q.J.E.* 117 (May): 609–57.
- Moene, Karl Ove, and Michael Wallerstein. 1997. "Pay Inequality." *J. Labor Econ.* 15 (July): 403–30.
- Moffit, Robert. 2013. "Medicare Drugs: Why Congress Should Reject Government Price Fixing." Heritage Foundation Issue Brief no. 3880. <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." *A.E.R.* 56 (March): 69–75.
- OECD. 2010. STAN: OECD Structural Analysis Statistics. Database. Paris: OECD.
- Persson, Torsten, and Guido Tabellini. 2003. *The Economic Effects of Constitutions*. Munich Lectures in Economics. Cambridge, MA: 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." *Ann. Internal Medicine* 143 (September): 397–403.
- Romer, Paul. 1990. "Endogenous Technological Change." *J.P.E.* 98 (September): 71–102.

- Saint-Paul, Gilles, and Thierry Verdier. 1993. "Education, Democracy and Growth." *J. Development Econ.* 42 (December): 399–407.
- Scherer, Frederic M. 1993. "Pricing, Profits, and Technological Progress in the Pharmaceutical Industry." *J. Econ. Perspectives* 7 (Summer): 97–115.
- . 2001. "The Link between Gross Profitability and Pharmaceutical R&D Spending." *Health Affairs* 20 (September): 216–20.
- . 2004. "The Pharmaceutical Industry: Prices and Progress." *New England J. Medicine* 351 (August): 927–32.
- Spence, Michael, and David Starrett. 1975. "Most Rapid Approach Paths in Accumulation Problems." *Internat. Econ. Rev.* 16 (June): 388–403.