

NBER WORKING PAPER SERIES

WHY NOT A POLITICAL COASE THEOREM?
SOCIAL CONFLICT, COMMITMENT AND POLITICS

Daron Acemoglu

Working Paper 9377

<http://www.nber.org/papers/w9377>

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

I thank Raphael Auer for outstanding research assistance and seminar participants at the World Bank Appropriate Institutions conference and at the International Society for New Institutional Economics conference, and Simeon Djankov, Simon Johnson, James Robinson and two anonymous referees for comments. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2002 by Daron Acemoglu. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Why Not a Political Coase Theorem? Social Conflict, Commitment and Politics
Daron Acemoglu
NBER Working Paper No. 9377
December 2002
JEL No. H2, N10, N40, O1

ABSTRACT

Do societies choose inefficient policies and institutions, in contrast to what would be suggested by a reasoning extending the Coase Theorem to politics? Do societies choose inefficient policies and institutions because of differences in the beliefs and ideologies of their peoples or leaders? Or are inefficiencies in politics and economics the outcome of social and distributional conflicts? This paper discusses these various approaches to political economy, and develops the argument that there are strong empirical and theoretical grounds for believing that inefficient policies and institutions are prevalent, and that they are chosen because they serve the interests of politicians or social groups holding political power, at the expense of the society at large. At the center of the theoretical case are the commitment problems inherent in politics: parties holding political power cannot make commitments to bind their future actions because there is no outside agency with the coercive capacity to enforce such arrangements.

Daron Acemoglu
Department of Economics
Massachusetts Institute of Technology
50 Memorial Drive
Cambridge, MA 02142
and NBER
daron@mit.edu

1 Introduction

There is increasing interest in the effects of government policies, economic, political and legal institutions, and more broadly, of the organization of society on economic outcomes. There is also a sense among economists and social scientists that differences in policies and institutions are a first-order determinant of the very different economic and social fortunes of various countries, though we are far from a consensus on this point. Despite the range of important contributions on these topics (see, for example, the surveys in Austen-Smith and Banks, 1999, and Persson and Tabellini, 2000), we still do not know how to organize our thinking on the determinants of policies and institutions.¹ Why do certain societies choose different policies, different institutions, radically different ways of organizing their lives? This paper is an attempt to provide a simple taxonomy, and argue for the relevance of a particular approach.

The taxonomy consists of grouping various approaches into three categories:

1. Political Coase Theorem (PCT): the Coase Theorem maintains that when property rights are well-defined and there are no “transaction costs”, economic agents will “contract” to achieve an efficient (output- or surplus-maximizing) outcome, irrespective of who has the property rights on particular assets (Coase, 1960, Stigler, 1966). An extension of this reasoning to the political sphere would suggest that political and economic transactions create a strong tendency towards policies and institutions that achieve the best outcomes given the varying needs and requirements of societies, *irrespective of who, or which social group, has political power*. According to this approach, policy and institutional differences are *not* the major determinant of the differences in economic outcomes, since societies choose, at least approximately, the “right” policies and institutions given their conditions.
2. Theories of Belief Differences (Modified Political Coase Theorem) societies may

¹There is an important distinction between policies and institutions. We can loosely think of policies as choices made within a given political and social structure, e.g., the tax rate and fiscal policies. In contrast, institutions can be thought as determinants of the political and social structure that are more durable, and as such, constrain future choices and policies, e.g., whether the society is democratic or not, the nature of the legal constraints on the government, or the extent of private property rights enforcement. Although institutions are often predetermined at the point in time when certain policy choices are made, they are also chosen by the society. For example, governments and citizens decide what legal code will apply, and how stringently it will be enforced. The focus here is on why both inefficient policies and inefficient institutions are chosen and continue to be chosen. So for most of this paper I will not make much of a distinction between policies and institutions.

choose different policies, with very different implications, because they—or their leaders—disagree about what would be good for the society. According to this approach, there is sufficient uncertainty about the right policies and institutions that well-meaning political actors differ about what is good for their own people. Societies where the leaders or the electorate turn out to be right *ex post* are those that prosper. The important point is that, just as with the PCT, there are strong forces preventing the implementation of policies that are *known* to be bad for the society at large, hence the label Modified Political Coase Theorem.

3. Theories of Social Conflict: societies choose different policies, some very disastrous for their citizens, because those decisions are made by politicians or politically powerful social groups that are interested in maximizing their own payoffs, *not* aggregate output or social welfare. This category includes both theories where internal conflict within the society leads to inefficient choices and those where inefficient institutions and policies are imposed on societies from the outside (e.g., by colonial powers).²

At some level, the major divide is between Theories of Social Conflict, which emphasize the prevalence of systematically inefficient government policies and institutional arrangements, versus the first two approaches, which stress the presence of social forces that rule out these types of inefficiencies. The rest of the paper argues that the PCT, in its simple form, or in its modified version built on belief differences, is not an appropriate framework for thinking about policy and institutional differences across countries.

²This taxonomy implicitly classifies a lot of interesting theories that combine features from the three groups, while still giving a prominent role to social conflict, in the category of Theories of Social Conflict. Most important, in many theories featuring social conflict, societies will make different choices because of differences in their economic conditions, but generally there will not be strong forces towards efficient outcomes in any of these societies.

A number of interesting interactions are also worth mentioning briefly. For example, certain groups may attempt, or manage, to convince others that their most-preferred policies also benefit the society at large (see Coate and Morris, 1995, for a model with this flavor). Another interesting interaction arises when some societies choose different policies initially because of differences in beliefs, but then these policies create or strengthen their own political constituencies, supporting the continued implementation of these inefficient policies (see Acemoglu, Aghion and Zilibotti, 2002, for a model with this flavor).

Another set of approaches, especially popular among sociologists and political scientists, is also worth mentioning here. These also maintain that many policies and institutions are inefficient, but contrary to Theories of Social Conflict, they do not attempt to explain these inefficiencies by the economic or social objectives of competing groups. Instead, institutions and policies are presumed to arise as unintended consequences of other interactions. A salient example is Tilly's (1990) work, and its extension by Herbst (2000), which stresses the importance of the emergence of the nation-state for economic development, but whether the nation-state emerges or not is a consequence of other unrelated factors, for example population density or the frequency of wars.

Existing evidence suggests that societies often choose inefficient policies and institutions, and in most cases they do this *not* because of differences in beliefs, but because of severe misalignments in the economic interests of politically decisive actors and the rest of the society. So Theories of Social Conflict appear to provide the right starting point for an analysis of policy and institutional differences.

A major challenge for Theories of Social Conflict, however, is to pinpoint what specific “transaction costs” would systematically prevent the Political Coase Theorem from applying. In other words, why do politicians and powerful social groups not make a deal with the rest of the society to choose the policies and institutions that maximize output (or social welfare), and then redistribute part of the gains to themselves? Put even more strongly; why do powerful groups not “predate” efficiently? The theoretical analysis is intended to highlight some of the issues that arise in thinking about these questions.

The argument I develop in Sections 4 and 5 of the paper is that, although the PCT may be a useful benchmark, its applicability is limited because of the inherent *commitment problems* associated with political power. Underlying the Coase Theorem is the ability to write enforceable contracts. Any enforcement problem will therefore potentially limit the applicability of the Coase Theorem.³ In the context of the PCT, there is a natural reason for widespread enforcement problems. Most contracts are enforced by “the state”. When it comes to contracts that the state or social groups controlling the state would like to write with others (e.g., the citizens), they will, by definition, be non-enforceable because groups controlling the state cannot *commit* not to using their power to renege on their promises or not to changing the terms of the contract. This implies that the allocation of political power creates an inherent commitment problem, undermining the potential to reach efficient outcomes.

The commitment problem associated with the PCT is twofold: first, those in power, e.g., the rulers, cannot commit to not using this power—as long as they do not relinquish it—in ways that benefit them in the future. Second, if the rulers relinquish their power, the citizens cannot commit to making side payments to them in the future, because the

³Enforcement problems may arise from incomplete information, contracting costs or bounded rationality (e.g., Anderlini and Felli, 1998, Dixit and Olson, 2000, Farrell, 1987, McKelvey and Page, 1999). I focus on the commitment problems here, because I believe these are much more central when it comes to the PCT. The recent literature on transaction costs and the organization of the firm (e.g., Williamson, 1981, Grossman and Hart, 1986, Hart and Moore, 1990) similarly focuses on why the distribution of property rights may matter for incentives when contracts are incomplete. There are a number of informal attempts to extend the reasoning of transaction costs to politics, see, for example, North (1990) or Spiller and Tommasi (2002).

former rulers no longer possess the political power to enforce such promises. This double commitment problem restricts the potential remedies to inefficiencies. Nevertheless, because the relationship between the state and the citizens is repeated, there may be some amount of commitment based on reputation—supported by the threat of future punishments. As a result, the extent to which the PCT will provide a good approximation to reality depends on the possibility of commitment via constitutions or other institutions and on how good a substitute this type of reputation-based commitment is for enforceable contracts, as well as on the extent of distributional conflict between various social groups in society.

To focus on the inherent commitment problems present in political situations, the simple model analyzed here allows unrestricted transfers and taxes, including non-distortionary lump-sum taxes. The inefficiencies arise *not* because of some restrictions on the technology of taxation, but because of the political-economic interactions between different groups and agents. Interestingly, despite the availability of non-distortionary taxes, the model also suggests that in this type of repeated game environment, the equilibrium may involve distortionary taxes. The reason is that the allocation has to satisfy the incentive compatibility constraint of the ruler, which requires the current output not to be too large; otherwise, the ruler would prefer to grab all the output rather than stick to the agreement. With lump-sum taxes, individuals are *the residual claimants* of the returns they generate from their investments, and will have a tendency to “overinvest”, violating the incentive compatibility constraint of the ruler. Distortionary taxation may then be necessary to guarantee levels of investment consistent with the incentive compatibility constraint of the ruler.

There is a large literature on distortionary policies of governments, which is nicely surveyed in Robinson (1998), with a similar distinction to the one here between bad policies that arise due to belief differences and those originating from social conflict. The most celebrated models of distortionary policies are the voting models, where the median or the decisive voter may choose policies that redistribute resources from the society as a whole to himself or to his group (e.g., Romer, 1975, Meltzer and Richards, 1981, or see Persson and Tabellini, 2000). There is also a large literature in political science on how voting behavior and the organization of parties interact to produce equilibrium policies (see, for example, Aldrich, 1983, Baron and Ferejohn, 1989, Dixit and Londregan, 1995, Myerson, 1995, Snyder, 1990). A range of other papers emphasize conflict between

bureaucrats or politicians and the society, for example, Buchanan and Tullock (1962), Ferejohn, (1986), Persson, Roland and Tabellini (1997), and Shleifer and Vishny (1998). These papers do not consider why politically powerful groups cannot extract resources from the rest of the society in an efficient manner, however. In fact, much of this literature rules out efficient methods of redistribution and takes it for granted that rent-maximizing behavior by rulers or the government will result in inefficiencies. The focus here is instead on why efficient policies fail to arise. In this respect, this paper is related to North (1981), Libecap (1989) and Olson (2000), who emphasize inefficient policies resulting from distributional conflicts. North, for example, suggests that rulers will choose the system of property rights in order to maximize their return, and this will lead to inefficiencies, but he also places considerable emphasis on differences in beliefs. Neither North nor Libecap nor Olson is explicit, however, on why a version of the Political Coase Theorem would not apply.

By providing a rationale for inefficient methods of taxation, this paper also relates to a few existing studies investigating the reasons why, in many instances, societies use inefficient redistributive policies rather than lump-sum taxation and transfers. Rodrik (1986), Wilson (1991) and Becker and Mulligan (1998) argue that if the amount of redistribution is endogenous, then politicians might want to commit to using inefficient methods in order to reduce total redistribution. Coate and Morris (1995) argue that inefficient redistribution arises when politicians exploit the uncertainty of voters regarding which policies are efficient. In this context, most closely related are Besley and Coate (1998), Acemoglu and Robinson (2000b and 2002), Rajan and Zingales (2000), and Acemoglu and Robinson (2001). Besley and Coate analyze a two-period political economy model, and show how certain types of “inefficiencies” may arise because efficient policies would affect the identity of who is in power (though they limit the politicians to linear taxes), and similarly emphasize the importance of commitment problems.⁴ Acemoglu and Robinson (2000b and 2002) develop a theory where elites may want to block the introduction of new and efficient technologies because this will reduce their future political power, and similarly, Rajan and Zingales (2000) show how organizations make inefficient choices because each

⁴Other related papers emphasizing the importance of commitment issues in politics include: North and Weingast (1989), who argue that the introduction of the English Parliament in the seventeenth century was a commitment to low taxes in the future; Weingast (1998) who interprets the Missouri compromise as a commitment by Northerners not to attempt to abolish slavery in the South; and finally Acemoglu and Robinson (2000a), who argue that the introduction of democracy was a commitment by the rich elite to future redistribution.

group (or agent) is worried that others in the organization will get richer, and demand more concessions in the future. Finally, Acemoglu and Robinson show that inefficient methods of redistribution, rather than more efficient alternatives, may arise as a method of maintaining future political power (see also Persson and Svensson, 1989, and Aghion and Bolton, 1990, on the use of fiscal policy to affect future elections). None of these papers address the more general issue of how commitment problems undermine the “Political Coase Theorem”, which is the focus of the current paper, nor do they analyze repeated games where punishment strategies may substitute for lack of formal commitment.⁵

The rest of the paper is organized as follows. The next section revisits the above taxonomy of various approaches to the determination of policies and institutions introduced above. Section 3 argues that in practice, neither the PCT nor the Modified PCT provide a satisfactory framework for studying cross-country differences in institutions and policies. Sections 4 and 5 analyze a simple model of conflict between the ruler and the citizens, highlight the commitment problems inherent in political transactions, and show why the reasoning of the PCT will not apply in general. This analysis also develops some simple comparative statics from the model, and shows why distortionary taxes may be necessary to reduce overinvestment by citizens, which would otherwise violate the incentive compatibility constraint of rulers.

2 A Simple Taxonomy

To emphasize the differences between various approaches and build a simple taxonomy, consider the following setup, with Y denoting aggregate output or consumption, which I take to represent social welfare (thus avoiding some of the complications that come from Pareto comparisons, and focusing on the main point here). Moreover, suppose that we can write

$$Y = F(X, P),$$

where X is a vector of economic, geographic, social or other characteristics that are taken as given and directly influence economic outcomes, and P is a vector of policies and institutions that can potentially affect the outcomes of interest. I define $\mathbb{P}(\cdot | X)$ as the

⁵In focusing on infinite-horizon models with self-enforcing arrangements, this model is also similar to Dixit, Grossman, and Gul (2000) who analyze self-enforcing political deals between groups with different interests.

set of policies that maximize output, given a vector of characteristics X , i.e.,

$$P^*(X) \in \mathbb{P}(\cdot | X) \iff P^*(X) \in \arg \max_P F(X, P).$$

The Political Coase Theorem maintains that there are strong forces leading societies towards some $P^*(X)$ in $\mathbb{P}(\cdot | X)$. The underlying idea is that if a society is pursuing a policy $P(X) \notin \mathbb{P}(\cdot | X)$, then a switch to $P^*(X) \in \mathbb{P}(\cdot | X)$ will create aggregate gains. If these gains correspond to a Pareto improvement, then all political systems will implement this change. If the change creates only a potential Pareto improvement, then part of the gains can be redistributed to those that are losing out via various mechanisms, or at the very least, the winners can lobby or vote for the beneficial change. A number of social scientists have proposed limited forms of this Political Coase Theorem. Becker (1983, 1985), for example, pointed out how competition between pressure groups could create a force towards efficient policies. Wittman (1995) pushed this argument further and formulated an informal Political Coase Theorem for democratic societies. Wittman argued that with rational voters, democratic societies generally produce Pareto efficient, even wealth-maximizing, outcomes. In fact, Wittman's argument relies little on democratic institutions, and his reasoning could even apply to nondemocratic societies.

To the extent that $\mathbb{P}(\cdot | X)$ is not a singleton, we can observe considerable policy differences across two identical societies, but the performance of these two societies should not be appreciably different. An example could be differences in policies regarding the role of the government in the economy between the Anglo-Saxon economies, in particular, the U.S. and the UK, and Continental European countries, which do not seem to lead to major differences in the economic performance between these two sets of countries.⁶

When we look at a cross-section of societies in the data, however, we also see more major differences in policies and institutions, for example, free-market policies in some societies like Hong Kong, and heavy government involvement and widespread corruption in some others like Indonesia. But according to the PCT, various government interventions and corruption in Indonesia are not the reason why this country is poorer than Hong Kong. Each is choosing the policies and institutions that are *appropriate* for their own situations, but they achieve different outcomes because their situations, their X 's, are different. More specifically, for two societies with characteristics X and $X' \neq X$, we typically have $F(X, P^*(X)) \neq F(X', P^*(X'))$, and moreover, $F(X, P^*(X)) > F(X, P^*(X'))$ and

⁶See Hall and Soskice (2001) for a discussion of the costs and benefits of various different types of capitalism.

$F(X', P^*(X')) > F(X', P^*(X))$. Thus, the PCT suggests that Indonesian institutions are not chosen inefficiently, but appropriately for their circumstances.⁷

This discussion implies that to refute the applicability of the Political Coase Theorem, we need to find systematic evidence that there are societies choosing P while $F(X, P) < F(X, P')$ for some feasible alternative P' , or simply that $P \notin \mathbb{P}(\cdot | X)$. That is, we need to show that there are societies that persistently pursue wrong policies, with significant output and welfare consequences.⁸

Theories of Belief Differences (Modified PCT), on the other hand, emphasize that some subset of X , X_u , is uncertain. To simplify the notation while elaborating on this, suppose that $\mathbb{P}(\cdot | X)$ is a singleton, in particular $\mathbb{P}(\cdot | X) = P^*(X)$. Moreover, imagine that $X = (X_c, X_u)$, and suppose that $P^*(X_c, X_u) \neq P^*(X_c, X'_u)$ whenever $X_u \neq X'_u$, that is, these uncertain characteristics affect which policies are right for the society. Suppose that politicians (or the society at large) have beliefs, denoted by $G(X_u)$, over the actual distribution of X_u . Also suppose that social welfare maximization corresponds to the maximization of expected aggregate output. Then define

$$P^*(X_c, G) \in \arg \max_P \int F(X_c, X_u, P) dG.$$

Now two societies with the same X_c , and the same ex post realization of X_u , may choose different policies because their ex ante beliefs over the payoff-relevant characteristics, the X_u 's, are different. Given a particular realization of X_u , some societies among those with the same X_c and X_u will be richer than others, i.e., typically $F(X_c, X_u, P^*(X_c, G)) \neq F(X_c, X_u, P^*(X_c, G'))$ for $G \neq G'$.

For example, the North Koreans may be choosing socialist policies and government ownership because they believe those are the policies that will increase welfare, while South Korea, which presumably had the same characteristics, X_c and X_u , chose a capitalist development path. Ex post, the South Koreans turned out to be right, hence they were the ones who adopted the right policies, and the ones who prospered, while North Koreans today suffer poverty and famine.⁹

⁷See Glaesar and Shleifer (2002) for an explanation for why Britain and France chose very different legal codes and systems that were appropriate to their underlying circumstances.

⁸Throughout by “refuting” the PCT, I mean showing that there are significant and quantitatively important inefficiencies in the institutions and policies of some societies—of course, this definition poses the question of what is “significant and quantitatively important”. A refutation of the PCT does not imply that there are no forces towards more efficient arrangements.

⁹An interesting theory of policy differences arising from belief differences is developed by Piketty

To refute the class of models in this group, we need to show that there are societies that pursue policies that could not be the right policies under any plausible scenario. In other words, denoting the set of admissible beliefs by \mathbb{G} , if, for two feasible policies, P and P' , $\int F(X_c, X_u, P') dG \geq \int F(X_c, X_u, P) dG$ for all $G \in \mathbb{G}$, then we should never observe P .

Finally, according to the *Theories of Social Conflict*, societies often, *knowingly*, choose some policy vector $P(X) \notin \mathbb{P}(\cdot | X)$, because policies and institutions are chosen to maximize the payoffs of those who hold political power, *not* to maximize social welfare or aggregate income. To emphasize the difference between this approach and the Political Coase Theorem, imagine another vector of variables Z , which *do not* directly affect Y , thus $P^*(X)$ is independent of Z . These variables may nonetheless influence the “equilibrium” policy, so we can have $P(X, Z)$. Changes in Z will have no direct effect on output, but may have a powerful indirect impact by influencing the gap between $P(X, Z)$ and $P^*(X)$. In other words, we need to find a variable, Z , that is like an instrument in econometrics: it influences X , but has no direct effect on F .

At this level of generality, Theories of Social Conflict are more like a residual group; if we can show that certain societies systematically, and knowingly, pursue inefficient policies, we are in the realm of Theories of Social Conflict. But the usefulness of these theories depends, in turn, on whether they can pinpoint an interesting mechanism for why political and economic bargains are not struck to achieve better policies and institutions (i.e., what are the salient “transaction costs” preventing the PCT from applying?), and whether we can identify a range of institutional or other social variables, the Z 's, that affect the degree of inefficiency of policies.

3 What the Data Say

In this section, I briefly develop the argument that cross-country differences in policies and institutions are important determinants of economic performance, and the origins of these differences do not lie in different perceptions of the peoples and the leaders, but in the social conflicts that exist between these leaders, or the social groups that these leaders represent, and the rest of the society. As noted above, the purpose of this exercise

(1995) where individuals vote over the degree of redistribution in the economy as a function of their beliefs on the importance of individual effort in economic success. These beliefs, in turn, evolve as a result of various economic interactions and tax policies. See Romer (1997) and Mukand and Rodrik (2002) for recent studies arguing for the importance of these issues.

is not to argue that there are no economic and political forces towards more efficient social arrangements, but to show that there are salient examples of inefficient institutions and policies, accounting for quantitatively large variations in economic performance.

Recall also that societies may choose inefficient policies and institutions both because of internal conflict and because these choices are imposed on them externally. Although inefficiencies arising from internal conflict are at least as important, in the latter part of this section I focus on two examples of inefficient institutions imposed by outside forces, because these episodes make it clear that these institutional choices were not in response to different economic circumstances (i.e., they exploit sources of exogenous variation from history).

3.1 Differences in Institutions and Economic Outcomes

There are tremendous cross-country differences in the way that economic and political life is organized. Let us focus here on a range of characteristics which we can think of as “economic institutions,” for example, the degree of stable property rights enforcement, the extent of equal opportunity and of entry barriers. A voluminous literature documents large cross-country differences in economic institutions, and a strong correlation between these institutions and economic performance.

To pick a few examples, Knack and Keefer (1995) look at measures of property rights enforcement compiled by international business organizations (in particular Political and Risk Services), Mauro (1995) looks at measures of corruption, and Djankov, La Porta, Lopez-De-Silanes and Shleifer (2002) compile measures of entry barriers across countries, while many studies look at variation in educational institutions and the corresponding differences in human capital (e.g., Ringer, 1979, Krueger and Lindahl, 2001, Hanushek and Kimko, 2000). All of these authors find substantial differences in these measures of economic institutions, and significant correlation between these measures and various indicators of economic performance. For example, Djankov et al. document that, while the total cost of opening a medium-size business in the United States is less than 0.02 percent of GDP per capita in 1999, the same cost is 2.7 percent of GDP per capita in Nigeria, 1.16 percent in Kenya 0.91 percent in Ecuador and 4.95 percent in the Dominican Republic. These entry barriers are highly correlated with various economic outcomes, including the rate of economic growth and the level of development.

A defender of the PCT could counter these empirical patterns with the following

argument: this type of correlation does not establish that countries are choosing the wrong institutions. After all, the United States differs from Nigeria, Kenya and the Dominican Republic in its economic characteristics, i.e., its X 's. And different X 's require different optimal (appropriate) policies and institutions. In terms of the notation introduced above, it may be the case that $X \neq X'$, and consequently, $P^*(X) \neq P^*(X')$, and $F(X, P^*(X)) \neq F(X', P^*(X'))$, so we might simply be observing the optimal response of different societies to their own varying conditions. Perhaps, given the circumstances in the Dominican Republic, it is not worth investing in the arrangements to reduce the costs of opening and doing business (or in the modified form of the PCT, perhaps the people of the Dominican Republic believe that high entry barriers are good for the society).¹⁰

To refute the general applicability of the PCT for analyzing differences in institutions and policies across countries, and their impact on economic outcomes, we need to show that otherwise identical, or at least similar, societies choose different institutions and policies because of reasons that do not directly affect economic outcomes, and experience differential economic performances as a result of these choices. This is essentially the reasoning of *instrumental variables*. Therefore, to refute the PCT, we have to find a source of variation, the Z 's, that do not directly influence economic outcomes, but affect the choice of policies and institutions, and then show that these differences matter for economic outcomes. In other words, we have to find some type of natural (or “unnatural”) social experiment where for political or historical reasons some societies end up with very different institutions than others (in addition, if we also want to refute the modified PCT, we would have to show that the variation captured by the Z 's is not working solely through belief differences). Naturally, the focus on historical and sources of variation in institutions and policies does not mean that these provide the most major reason for cross-country differences. Internal dynamics leading to different policies institutions are likely to be at least as important, but for the purposes of this exercise, external sources of variation make identification easier.

¹⁰ An example of optimal non-enforcement of private property rights may be the case of North American Indians before the eighteenth century. Demsetz (1967) argues that despite the potential for overhunting of game, the costs of enforcing property rights in land were higher than the benefits, since without the fur market, there were only weak incentives for overhunting. This changed after the Indians started trading fur with the white Americans, at which point the incentives for overhunting and the costs of lack of property rights increased, and private property rights in land were duly introduced.

3.2 Colonialism and Institutional Development

European colonization of the rest of the world provides the best laboratory (almost a natural experiment) to investigate these issues. From the late 15th century, Europeans dominated and colonized much of the rest of the Globe. Together with European dominance came the imposition of various types of institutions in the colonies. Most interesting for our purposes, Europeans imposed very different institutions and social power structures in different parts of the world.

Acemoglu, Johnson and Robinson (2001) document that in a large number of colonies, especially those in Africa, Central America, the Caribbean and South Asia, European powers set up “extractive states”. These institutions did not introduce much protection for private property, nor did they provide checks and balances against government expropriation. The explicit aim of the Europeans in these colonies was the extraction of resources, in one form or another. In the Caribbean, this took the form of slave plantations, in parts of Central and Meso America, mining based on forced labor. In Africa, Europeans were first interested in the extraction of slaves to employ on the plantations in the Americas, and later developed other methods of extracting resources, including high taxes and extraction of natural resources.¹¹ Other economic institutions that Europeans set up in these colonies were similarly detrimental to economic advancement; there was little investment in the human capital of the majority of the population, and access to key resources was concentrated in the hands of a few.

This colonization strategy and the associated institutions contrast with the institutions Europeans set up in colonies where they settled in large numbers, for example, the United States, Canada, Australia and New Zealand. In these colonies, life was modeled after the home country, and the emphasis was on the enforcement of property rights for a broad cross-section of the society, especially smallholders, merchants and entrepreneurs. See for example Gann and Duignan (1962), Robinson and Gallagher (1961), Denoon (1983), Cain and Hopkins (1993).

¹¹For example, Davis and Huttenback (1986, p. 307) calculate that before 1885, investment in the British empire had a return 25 percent higher than that on domestic investment. Roberts (1976, p. 193) calculates a large transfer of resources from Northern Rhodesia to Britain, in return for minimal investment. Manning (1982) estimates that between 1905 and 1914, 50 percent of GDP in Dahomey was extracted by the French, Young (1994, p.125) notes that taxation rates in Tunisia were four times higher than those in metropolitan France, and Peemans (1975) documents the amount of resources extracted from the Belgian Congo and calculates that tax rates on Africans approached 60 percent of their income during the 1920’s and 1930’s.

Acemoglu, Johnson and Robinson (2001) document that the crucial determinant of whether Europeans chose the path of extractive institutions was whether or not they settled in large numbers. In colonies where Europeans settled, institutions were developed for their own future benefits. In colonies where Europeans did not settle, they often set up a highly centralized state apparatus, and other similar institutions, to oppress the native population and facilitate the extraction of resources in the short run. Based on this idea, Acemoglu, Johnson and Robinson (2001) suggest that in places where the disease environments made it easy for Europeans to settle, the path of institutional development should have been different from areas where Europeans faced high mortality rates.

In practice, during the time of colonization, Europeans faced widely different mortality rates in colonies because of differences in the prevalence of malaria and yellow fever.¹² The argument in Acemoglu, Johnson and Robinson (2001) is that differences in mortality rates of potential settlers, driven mostly by malaria and yellow fever, provide a possible candidate for the variable Z above: these mortality rates should not influence output today directly; but by affecting the settlement patterns of Europeans, they may have had a first-order effect on institutional development. The idea that these mortality rates should not have a direct effect is plausible. Malaria and yellow fever were fatal to Europeans who had no immunity, thus having a major effect on settlement patterns, but they had much more limited effects on natives who, over centuries, had developed various types of immunities.¹³

The data support the notion that there were major differences in the institutional development of the high-mortality and low-mortality colonies. Figure 1 shows a measure of property rights enforcement—the protection against expropriation risk—against the logarithm of potential European settler mortality in 1000 mean strength soldiers (see Acemoglu, Johnson and Robinson, 2001 for details). Expropriation risk is much greater in places where Europeans faced higher death rates and did not settle.

¹²See Appendix Table A2 in Acemoglu, Johnson, and Robinson (2001) for the variation in the mortality rates of European military and clergy in the various colonies. Before 1850, the annual mortality rates for a settlement size maintained at 1000 (via replacement) ranged from 8.55 in New Zealand (lower than in Europe at that time) to 49 in India, 130 in Jamaica, and around 500 in West Africa.

¹³This “exclusion restriction” is supported by the death rates of natives in these areas. For example, Curtin (1964) reports that the annual death rates of native troops serving in Bengal and Madras were respectively 11 and 13 in 1000. These numbers are similar to the annual death rates of British troops serving in Britain, which were approximately 15 in 1000. In contrast, the death rates of British troops serving in these colonies were much higher because of their lack of immunity. For example, death rates in Bengal and Madras for British troops were between 70 and 170 in 1000.

Acemoglu, Johnson and Robinson (2001) also show that these institutional differences induced by mortality rates and European settlement patterns have a major effect on income per capita.¹⁴ For example, the estimates imply that improving Nigeria's institutions to the level of those in Chile could, in the long run, lead to as much as a 7-fold increase in Nigeria's income. These results have a clear interpretation: in practice, societies do choose very different institutions, and not because of differences in output-relevant variables, the X 's, but because of other political/historical circumstances, the Z 's, in this case the mortality rates faced by early European settlers. Moreover, these institutional choices have a major effect on economic performance. These results suggest that the PCT, which emphasizes the forces that push societies towards the correct institutions and policies, does not provide a useful framework for analyzing the major institutional and policy differences across countries.

3.3 Another Experiment: North Versus South Korea

Another example that illustrates how societies with very similar conditions, but different histories or political equilibria, may end up with very different economic and political institutions, and consequently with divergent economic performances, is the contrast between North and South Korea.

Until the end of World War II, Korea was under Japanese occupation. Korean independence came shortly after the Japanese Emperor Hirohito announced the Japanese surrender on August 15, 1945. After this date, Soviet forces entered Manchuria and North Korea and took over the control of these provinces from the Japanese. The United States did not want to leave the control of the Korean peninsula to the Soviet Union, and with "General Order No. 1," President Truman proposed a joint occupation of Korea, with the division between the north and south at the 38th parallel. The major fear of the United States during this time period was the takeover of the entire Korea either by the Soviet Union or by communist forces under the control of the former guerrilla fighter, Kim Il Sung. U.S. authorities therefore supported the influential nationalist leader Syngman Rhee, who was in favor of separation rather than a united communist Korea. Elections

¹⁴That paper also documents that this effect of institutions on economic performance is robust to excluding Australia, New Zealand, Canada, and the United States, or Africa, to controlling for various geography variables such as distance from the equator (latitude), continent dummies, temperature, humidity, whether countries are land-locked, soil quality, natural resource abundance etc. They also obtain similar results using only yellow fever prevalence, which is an attractive source of variation, since yellow fever is now mostly eradicated.

in the South were held in May 1948, amidst widespread boycott by Koreans opposed to separation. The newly elected representatives proceeded to draft a new constitution and established the Republic of Korea to the south of the 38th parallel. The North became the Democratic People's Republic of Korea, under to control of Kim Il Sung.¹⁵

A distinguishing feature of Korea before separation was its ethnic, linguistic and economic homogeneity. The north and south are inhabited by essentially the same people, with the same culture, and there were only minor economic differences between the two areas. If anything, at the time of separation, the North was more industrialized; for example, production levels of heavy industrial output were almost four times as high in the North as in the South, despite the larger size and population of the South (production of light industry in turn was greater in the South, see Ha-cheong, 1988).

After separation, policies and institutions diverged substantially in the two countries. The North, under the leadership of Kim Il Sung, adopted a centralized command economy with little role for individual enterprise. Kim Il Sung advocated and largely imposed a philosophy he named "Juche", which played an important role both in the political and economic life in North Korea. This philosophy emphasized self-reliance and the control of resources by the communist party and the state which, in turn, were supposed to represent the people. All non-labor factors of production were under the control of the state, which directly made the majority of the key economic decisions. Before the separation, industries in North Korea were concentrated in mining, electricity, steel, chemicals, transportation, communication and cement. Most of these were quickly nationalized. There were also many small household industries and producers, and these were forced to join the cooperatives of the Consumer Union, where they were closely supervised and instructed by the state. For all practical purposes, there were no private property rights for individuals (see, for example, Koo, 1992, or Eberstadt, 1999).

In contrast, South Korea, though far from a free-market economy, relied on a capitalist organization of the economy, with private ownership of the means of production, and legal protection for a range of producers, especially those under the umbrella of the chaebols, the large family conglomerates that dominated the South Korean economy. Although Syngman Rhee, and subsequently General Park, were highly dictatorial, for a variety of reasons beyond the scope of this paper, they refrained from the most predatory policies. In fact, General Park was generally supportive of economic development, and his regime is

¹⁵See Cumings (1997) and Buzo (2002) for recent histories of Korea.

often credited with facilitating, or even encouraging, investment and rapid growth in Korea (e.g., Evans, 1998, Wade, 1990). Even though many South Korean economic policies, such as protected domestic markets, entry barriers and subsidized loans, directly favored the chaebol, there were no major violations of property rights for the rest of the society, and the state actively subsidized and encouraged education. Overall, South Korean economic institutions were highly capitalistic, even though the government intervened more than the simplest textbook model of free-market capitalism would suggest.

Under these two highly contrasting regimes, the economies of North and South Korea diverged. In 1950, according to Maddison (2001), both North and South Korea had approximately the same income level, \$770 (in 1990 International Geary-Khamis dollars). In the 1990s, before the collapse of the Soviet system and the cessation of Soviet aid, Maddison (2001) estimates per capita income in North Korea to be around \$2,841, less than one-third of the income per capita in South Korea, which stood at \$8,704. The South Korean government estimates the North Korean GDP per capita to be less than 1/6th of the South Korean per capita in 1990 (see www.bok.or.kr). In any case, there is no doubt that income in North Korea in 1990 was inflated by Soviet aid, and since then, the North Korean economy has been shrinking, while South Korea continued to grow rapidly. According to Maddison (2001), there is now an over 10-fold difference: income per capita is \$12,152 in South Korea vs. \$1,183 in North Korea. Figure 2 shows the divergence in income per capita between South and North Korea using Maddison's (2001) numbers.

Overall, since 1950, South Korea grew rapidly under capitalist institutions and policies, while North Korea experienced minimal growth, under communist institutions and policies. This “experiment” of dividing this homogeneous country into two parts with very different policies and institutions gives another clear example of how, despite the very similar economic conditions, political leaders often choose very different policies, with very different outcomes.

3.4 Conscious Choices or Belief Differences?

Can we interpret the differences in institutional development across the European colonies or the divergence in the institutions and policies between the North and South of Korea as resulting from differences in beliefs? For example, could it be the case that while Rhee, Park, and other South Korean leaders, believed in the superiority of capitalist institutions and private property rights enforcement, Kim Il Sung and Communist Party members in

the North believed that communist policies would be better for the country?

In the case of South versus North Korea, this is certainly a possibility. However, even if differences in beliefs could explain the divergence in institutions in the immediate aftermath of separation, by the 1990s or even by the 1980s, it was clear that the communist economic policies in the North were not working. The continued efforts of the leadership to cling to these policies and to power can only be explained as those leaders wishing to look after their own interests at the expense of the interests of the population at large. Currently, North Korean leaders, the Communist Party and the bureaucratic elites are prolonging the current system, which gives them greater economic and political returns than the alternative, while fully realizing the costs that the system imposes on the North Korean people, including the famine that much of the population has been suffering over the past several years.

Differences in colonial policies make an even clearer case for the importance of social and distributional conflict leading to inefficient policies and persistently inefficient institutions. It was the same British colonists who established different institutions in very different parts of the world: in the Caribbean and Southern United States, they set up plantation societies based on slavery, supported by highly oppressive institutions. In contrast, the institutions they developed in areas where they settled, and where there was no large population of Indians or slaves to be oppressed, such as Northeastern United States, Canada, Australia and New Zealand, were democratic, encouraged participation, imposed checks and balances on politicians and political elites, and generally enforced the property rights of a broad cross-section of society.¹⁶ Moreover, differences in the incentives of the colonists in various colonies are easy to understand: when they did not settle, they were choosing the institutions simply to extract resources from the native population. When they settled in large numbers, institutions and policies were set in place in order to protect them in the future and encourage investment and prosperity.

¹⁶An interesting example of how the same groups adopted very different colonization strategies and organizations of society in response to different incentives is the experience of the Puritans in the New World. While the colony of Massachusetts Bay, formed in 1630 by the English Puritans is hailed as an example of good institutions introduced in the colonies by a group seeking economic and religious freedom, at the same time a group of Puritans sponsored by the powerful Puritan interests in England formed a colony in Providence Island in the Western Caribbean. Slavery was immediately adopted in this colony, and the most profitable endeavor for those settling on this island was attacking and pirating Spanish ships in the area (see Kupperman, 1988).

4 Commitment Problems and The Political Coase Theorem

In the previous section, I developed the argument that the PCT, in its simple or modified form, does not provide a good framework for thinking about the major cross-country differences in institutions and policies. The argument was essentially empirical. In this section, I turn to a discussion of why we should expect socially and politically powerful groups to often extract resources from the rest of the population in inefficient ways and set up bad institutions. In other words, why should we expect the PCT not to hold?

The basic idea is that the reasoning of the Coase Theorem requires transfers from one party to another, and all such transfers cannot be made at the same time. Therefore, some type of “enforceable contracts” specifying future transfers are necessary. And yet, when such transfers are between the citizens and the state (or groups controlling the state), there is a major commitment problem: there is no outside party to enforce such contracts, and those controlling the state can renege on their promises. Only “incentive compatible” or “self-enforcing” arrangements can be made. These inherent commitment problems often rule out the reasoning of the PCT.

To simplify the discussion, I focus on the case with no belief differences, i.e., the PCT, though obviously the discussion here applies to the Modified PCT as well.

4.1 Description

Consider the following infinite horizon economy. Time is discrete and indexed by t . There are two groups of agents, a ruler, and a mass 1 of identical citizens. All agents discount the future with the discount factor β , so have a utility function

$$u_t = \sum_{j=t}^{\infty} \beta^j [c_{t+j} - (1 - \alpha) e_{t+j}],$$

where c_{t+j} is consumption and e_{t+j} is effort (investment), and the term $(1 - \alpha)$ is introduced to simplify some of the algebra below.

$$y_t^i = f(e_t^i) = (e_t^i)^{1-\alpha} + R$$

There is also another inferior production technology, which has the advantage of being non-taxable. In particular, this alternative produces nonmarket income

$$n_t^i = b^\alpha (e_t^i)^{1-\alpha}$$

where $b < 1$. I denote the decision to produce for the market by $m_t \in \{0, 1\}$, and when $m_t = 0$ and the individual uses the non-market technology, his market income is $y_t^i = R$, so income from natural resources, R , remains taxable.¹⁷

In the first-best—efficient—allocation, only the superior market technology would be used, and the level of investment would satisfy:

$$\frac{\partial f}{\partial e_t} = 1 \iff e_t = e^{fb} \equiv 1$$

There is a ruler with the power to tax the citizens. The ruler does not contribute to the production process, but because he has all the means of coercion, he can take as much of the output in the market sector as he wishes. This clearly ignores useful roles of rulers and of the state, such as law enforcement, public good provision, regulation and defense. This is only for simplicity, since these roles are not essential for the theory here.¹⁸

The feasibility constraint that determines the maximum tax per person that the ruler can impose is

$$T_t(Y) \leq Y \equiv \int y_t^i di, \tag{1}$$

where Y denotes aggregate output. In addition, the taxation technology needs to be specified. The most general case would be when the ruler specifies the person-specific taxes, T_t^j , for each individual j . I return to a discussion of this issue below.

Rulers have the same discount factor, β , as the citizens, but because of internal power struggle, they can also lose power to another ruler with exogenous probability q (see Acemoglu, 2002, for a similar model where this probability is endogenized). The current ruler can also decide to relinquish his power, in which case there are no more rulers in the future, and nobody can tax the citizens. I refer to this case loosely as “democratization,” though it lacks many of the interesting features of real-life experiences of transitions to democracy (see Acemoglu and Robinson, 2000a, for a model of democratization). I will make two alternative assumptions regarding what type of contracts can be enforced: (i) contracts that citizens write with a current or previous ruler can be, at least partly,

¹⁷The presence of market income, even when individuals withdraw from market production will ensure that rulers continue to get positive return.

¹⁸See Acemoglu and Verdier (1998) for a model where government plays a useful role, but also government officials are corrupt and their actions distort private incentives.

“enforced”; (ii) no enforcement of such contracts. The plausible case is clearly (ii), but (i) is useful to analyze as a benchmark.

The timing of events within each period is as follows:

1. If contracts are available, parties sign contracts. If there has been no democratization in the past, i.e., $r = 0$ in all past periods, then the ruler decides whether to relinquish his power, $r = 0$ or $r = 1$.
2. Individuals choose how much to invest, e , and whether or not to produce in the market sector, $m = 0$ or 1 .
3. If $r = 0$ in all previous periods, the ruler decides how much aggregate tax $T(Y)$ to impose on the citizens, as a function of aggregate income Y . If $r = 1$ in some previous period so that there is “democracy”, then a citizen decides how much tax $S(Y)$ to impose on each individual to be given to a previous ruler.
4. Consumption takes place.
5. If there has been no democratization, it is revealed whether or not the ruler will be in power in the next period (he is replaced with probability q).

This timing of events introduces the assumption that not all transactions can be made at the same time; citizens invest first, and rulers set taxes after. So some type of “contracts”, implicit or explicit, are necessary.¹⁹ The history of play in this repeated game, h^t , includes all the actions up to that point. The strategy of a ruler consists of a mapping $\sigma(\cdot | h^t)$, which determines $(r, T(Y))$ in every period for a given history h^t . The level of taxes T is in turn conditioned on the level of output, since, according to the timing of events, taxes are determined after citizens make their investment and sector choices, and also because taxes can never exceed the level of output. The strategy of citizens consists of a mapping $\rho(\cdot | h^t)$, which determines $(m, e, S(Y))$ for a given history of the game h^t (I am focusing on symmetric equilibria where all citizens use the same strategy, hence I specify only one strategy mapping for the citizens, $\rho(\cdot | h^t)$). The investment and

¹⁹This game also introduces a possible distinction between institutions and policies: institutions may correspond to whether the society is democratic, i.e., who has the right set taxes, while policies correspond to the choice of actual taxes. Nevertheless, my focus here is not to clarify the distinction between institutions and policies, but highlight the forces that prevent the efficient choice of policies (and institutions).

sector choices of citizens are conditioned on the actions of the ruler in the same period that are observed before the citizens' actions.

A subgame perfect equilibrium is defined as a strategy $\sigma(\cdot | h^t)$ for the ruler and a strategy $\rho(\cdot | h^t)$ for all citizens that are best responses to each other in all subgames, i.e., for all h^t . To simplify the discussion throughout, I focus on stationary equilibria, where the same strategies are played at all dates.

4.2 The No-Cooperation Benchmark

Let us start with “no-cooperation benchmark,” which features $r = 0$, i.e., no democratization, and no contracts between rulers and citizens.

Proposition 1 Suppose $r = 0$. Then there exists an equilibrium in which all citizens expect the ruler to grab everything, so they use the informal sector technology, $m = 0$ and $e = b$. The ruler sets $T(Y) = Y$.

It is straightforward to see that this allocation is an equilibrium. It is a weakly dominant strategy for the ruler to grab everything, which along the equilibrium path will simply be the income from natural resources, R . And if an individual deviates and produces more, this will not increase his consumption, since the ruler is grabbing everything. So the citizens choose $m = 0$, i.e., production with the nonmarket technology, and they invest the optimal amount for this technology, $e = b$.

For future reference, denote the values of the citizens and the ruler in this equilibrium as \widehat{W} and \widehat{V} :

$$\widehat{W} = \frac{\alpha b}{1 - \beta}, \quad (2)$$

and

$$\widehat{V} = \frac{R}{1 - \beta(1 - q)}. \quad (3)$$

This equilibrium is highly inefficient. For example, a contract along the following lines would constitute a Pareto improvement: the ruler relinquishes power, and the citizens promise him a side payment of $R + \varepsilon$ every period thereafter. Then, they would all choose market production, $e = 1$, achieving the first-best equilibrium. The focus below will be on whether this type of arrangement can be made to improve the allocation of resources away from that specified in Proposition 1.

4.3 The Political Coase Theorem With Commitment

Next I discuss the equilibrium allocation when enforceable contracts between rulers and citizens are possible. In this discussion, I will say that a Political Coase Theorem (PCT) applies when, even in the absence of full property rights for citizens, the economy generates the same allocation (and when the distribution of political power between the citizens and the ruler is irrelevant for the allocation).

There are three different ways in which this result may emerge: with full commitment on the side of the ruler, with full commitment on the side of citizens, and with limited commitment. In this subsection, I discuss the first two cases, leaving the third, which is the central focus of this analysis, to the next section.

First suppose that the ruler can commit to impose a tax level T during this period (more specifically, $T(Y) = \min \langle T; Y \rangle$), and assume that $q = 0$, so that there is no ruler replacement. This means that each citizen will pay a tax level T . Then, each citizen knows that he will be able to keep any level of production above T . The following program gives the equilibrium allocation that satisfies the PCT, yielding the largest surplus to the ruler, with $r = 0$ (i.e., no “democratization”):

$$\max_{T,e} \frac{T}{1-\beta},$$

subject to the feasibility constraint (1), and an incentive compatibility constraint for citizens:

$$W(e) = \frac{1}{1-\beta} [e^{1-\alpha} - (1-\alpha)e + R - T] \geq \widehat{W}, \quad (4)$$

where the left-hand side of (4) is the return to citizens when they invest e , and are taxed T , and the right-hand side, \widehat{W} , is the value that citizens can obtain opting out of the formal sector, given by (2). The solution to this problem is straightforward: $T = \alpha(1-b) + R$ (or more formally, $T(Y) = \min \langle \alpha(1-b) + R; Y \rangle$) and $r = 0$ for the ruler, and $e = 1$ and $m = 1$ for all the citizens. The important point is that the efficient allocation is achieved despite the fact that political power is in the hands of the ruler. The reason for this is the commitment power of the ruler: by committing to the tax schedule $T(Y) = \min \langle \alpha(1-b); Y \rangle$, the ruler is making the citizens the residual claimant after a certain level of investment, and this encourages them to undertake the first-best level of investment.

The above program was special in that it gave all the “bargaining power” to the ruler. Alternatively, we can have some of the rents from achieving the PCT shared between the

ruler and the citizens. The simplest way to illustrate this is to assume that rents between citizens and rulers are shared by Generalized Nash Bargaining. By the same reasoning as above, citizens will choose the efficient level of investment, $e = 1$, so imposing this level of investment, the Generalized Nash Bargaining solution can be found from the following program:

$$\max_T \left[\frac{\alpha}{1-\beta} + \frac{R}{1-\beta} - \frac{T}{1-\beta} - \frac{\alpha b}{1-\beta} \right]^\theta \left[\frac{T}{1-\beta} - \frac{R}{1-\beta} \right]^{1-\theta}, \quad (5)$$

subject to (4), where θ is the bargaining power of the citizens. The first bracket is the return to citizens net of their outside option, which is production for the non-market sector with net present value, $\alpha b / (1 - \beta)$. The second bracket is the net return to the ruler above his outside option of taxing, only the income from natural resources. The solution to this problem has $T = (1 - \theta) \alpha (1 - b) + R$ (or more formally, $T(Y) = \min \langle (1 - \theta) \alpha (1 - b) + R; Y \rangle$) and $r = 0$ for the ruler, and $e = 1$ and $m = 1$ (production in the market sector) for all citizens. Note that the surplus from citizens producing in the market sector and undertaking the first-best level of investment (over the alternative of non-market production) is $\alpha (1 - b)$. This surplus is being shared between the citizens and the ruler. Income from “natural resources,” R , goes entirely to the ruler, since the ruler can obtain this even when citizens do not “cooperate”. As the bargaining power of citizens, θ , goes to zero, we obtain the tax level above, $T = \alpha (1 - b) + R$. Once again, the important point is that the efficient allocation is achieved, thanks to the commitment power of the ruler. Moreover, the model also illustrates that the distribution of political power between the ruler and the citizens, θ , does not affect the efficiency of the allocation; $m = 1$ and $e = 1$ irrespective of θ .

This solution is slightly more involved when the ruler can be replaced by a new ruler, i.e., $q > 0$. In this case, rulers have a preference for front-loaded payments, since they may not be around in the future, i.e., they discount the future at the rate $\beta(1 - q)$, which is less than the discount factor of citizens, β . However, citizens dislike making front-loaded payments to current rulers, since in case these rulers are replaced, these payments are lost, and additional payments have to be made to new rulers. These two effects cancel each other, and the problem is still stationary. In particular, now the allocation will be a solution to maximization problem:

$$\max_T \left[\frac{\alpha}{1-\beta} + \frac{R}{1-\beta} - \frac{T}{1-\beta} - \frac{\alpha b}{1-\beta} \right]^\theta \left[\frac{T}{1-\beta(1-q)} - \frac{R}{1-\beta(1-q)} \right]^{1-\theta},$$

subject to (4). This only differs from (5) because the discount factor of the ruler is different due to potential replacement at the end of the period. The solution is straightforward to characterize, and is identical to above. With complete contracts, the discount factor of the ruler does not matter for the equilibrium allocation.

Next suppose that the ruler cannot commit to a tax level T , but citizens can commit to a future path of transfers, $\{S_t\}$ if the ruler relinquishes his power. Now the PCT can be achieved via democratization, i.e., $r = 1$; the ruler transfers power to the citizens in return for their commitment to a future path of transfers. The equilibrium allocation with commitment on the side of citizens is therefore similarly a solution to the maximization problem:

$$\max_S \left[\frac{\alpha}{1-\beta} + \frac{R}{1-\beta} - \frac{S}{1-\beta} - \frac{\alpha b}{1-\beta} \right]^\theta \left[\frac{S}{1-\beta(1-q)} - \frac{R}{1-\beta(1-q)} \right]^{1-\theta},$$

subject to (4). The solution is: $S(Y) = \min \langle (1-\theta)\alpha(1-b) + R; Y \rangle$ and $r = 1$ for the ruler, and $e = 1$ and $m = 1$ for all citizens. Therefore, with commitment to future taxes and transfers on the side either the ruler or the citizens, the basic logic of the PCT is powerful, and the distribution of rents between various parties (here the ruler and the citizens) can be separated from efficiency considerations. The first-best investment level is achieved, and the distribution of power, here captured by θ , has no effect on the allocation.

Proposition 2 When either rulers or citizens can commit to future transfers, we always have $m = 1$ and $e = 1$, i.e., the PCT applies.

4.4 Limited Commitment

The above analysis with “contracts” between rulers and citizens is useful as a benchmark, but has little practical relevance. These types of contracts are not enforceable in the real world. Contract enforcement requires a third party, typically the state, which possesses the monopoly of legitimate coercion in the society. Its monopoly of coercion gives the state the power to force contractors to abide by the terms of the contract, even if ex post making the specified payments or the necessary delivery of goods is not in their interests.

When the state is one of the “contractors”, this type of outside enforcement is not possible. For this reason, it is very difficult for any party with real power to commit to a

path of future transfers, taxes or actions. Therefore, we cannot hope for the outside enforcement when we are dealing with interactions with the state; abiding by the conditions of the “contract” has to be incentive compatible for the state (and for the citizens).

To highlight these issues, I now analyze the above game without such contracts. I start with the Markov Perfect Equilibria (MPE) of this game, which does not allow repeated-game punishment strategies. In the next section, I will analyze non-Markovian equilibria.

Using the above notation, in this simple game, an MPE is defined as a strategy combination $\sigma(\cdot | h^t)$ for the ruler and $\rho(\cdot | h^t)$ for the citizens that are best responses to each other, and history-independent, i.e., $\sigma(\cdot | h^t) = \sigma(\cdot | h^{t'})$ and $\rho(\cdot | h^t) = \rho(\cdot | h^{t'})$ for any h^t and $h^{t'}$. Thus strategies in a MPE do not depend on the history of the game (more generally, they depend only on the payoff-relevant state variables, and here there are no such state variables).

This implies that within each period, we can solve the game by backward induction. In the last stage, the ruler in power sets the tax. The best action for the ruler is to grab everything, since the future play of the game, and therefore the continuation payoffs, does not depend on history, so whether or not he has grabbed everything will not have future repercussions. Hence, $T(Y) = Y$. Given this, citizens prefer $m = 0$, and there is no market production. We are back to Proposition 1.

This is a highly inefficient outcome, and one that both citizens and the ruler would like to prevent. For example, the ruler would like to promise to set a lower tax, e.g., $T(Y) = \min\langle T; Y \rangle$ for some $T \leq \alpha(1 - b) + R$, which would encourage citizens to stay in the market and invest up to the first-best level of investment. However, no such promises can be credible (without some type of trigger punishment strategies). The PCT, therefore, does not apply because of lack of commitment.

Proposition 3 Without the possibility of commitment by the ruler or the citizens to future actions, the unique Markov Perfect Equilibrium has $m = 0$, and $T(Y) = Y$.

5 Incentive-Compatible Promises

5.1 Incentive-Compatible Commitments by the Ruler

The discussion with MPEs ignored possible commitment that can be supported thanks to the repeated nature of the game. For example, if we allow strategies to depend on the history of the game, the citizens and the ruler may enter into an implicit agreement

where the ruler promises not to grab everything because of future rents from continued market production by the citizens. The important point is that such promises have to be “self-enforcing” or incentive compatible for the ruler. We can capture these issues by analyzing the non-Markovian equilibria of this game, where citizens play trigger strategies to induce the ruler not to grab all the output.

I will undertake this analysis in stages, starting with the case where there is no replacement of rulers, i.e., $q = 0$. Moreover, I first assume that the citizens can coordinate their actions and all choose the level of investment e , maximizing their utility as a group. Thus we can think of the game as one between two players. Later in subsection 5.3, I will come back to the issues of “free-riding,” where each individual may prefer to choose a different level of investment than the one maximizing the utility of the citizens as a group. These considerations will have interesting implications for the equilibrium form of taxation.

Consider the following strategy combination for the ruler and the citizens: the ruler sets the tax $T(Y) = \min \langle T; Y \rangle$ as long as citizens have played $e' = e$ in all past periods, and $T(Y) = Y$ otherwise; citizens play $m = 1$ and $e' = e$ as long as the ruler has set the tax $T(Y) = \min \langle T; Y \rangle$ in all past periods, and $m = 0$ otherwise, and we have $T \leq e^\alpha + R$. The resulting allocation will yield a tax revenue of T in each period, and give the ruler utility:

$$V = \frac{T}{1 - \beta}. \quad (6)$$

Since the ruler cannot commit to future taxes of the form $T(Y) = \min \langle T; Y \rangle$, we have to ensure that playing the above specified strategy is optimal for the ruler. The obvious, and the best, deviation for the ruler from this strategy profile is clearly to grab everything in the current period. So we have to check that not deviating from this strategy profile (i.e., not to grab everything today) is incentive compatible. If the ruler follows the repeated game equilibrium, he obtains V as given by (6). Alternatively, if he deviates to grab everything today and then switches to the non-cooperative equilibrium of Proposition 1, he obtains all the output today, $e^{1-\alpha} + R$, and from today onwards, he obtains his payoff in the non-cooperative equilibrium of Proposition 1, \widehat{V} . Thus the return to deviating is $e^{1-\alpha} + R + \beta\widehat{V}$.

Incentive compatibility for the ruler therefore requires:

$$e^{1-\alpha} + R + \beta\widehat{V} \leq V,$$

or written more compactly, the incentive compatibility constraint for rulers is

$$T \geq \Upsilon(e) \equiv (1 - \beta) e^{1-\alpha} + R, \quad (7)$$

where the function $\Upsilon(e)$ is defined for future reference, and represents the flow value of grabbing all current output for the ruler when current investment is e . Condition (7) states that the tax that the ruler receives in each period should be large enough so that he is not tempted to grab everything.

We also have to satisfy incentive compatibility for citizens, to convince them to stay in the market sector. When they do so, they obtain

$$W(e) = \frac{1}{1 - \beta} [e^{1-\alpha} - (1 - \alpha)e + R - T], \quad (8)$$

which needs to be greater than \widehat{W} given by (2) for the “equilibrium” e . In other words:

$$T \leq T^{\max}(e) \equiv R + e^{1-\alpha} - (1 - \alpha)e - \alpha b. \quad (9)$$

Here $T^{\max}(e)$ is the maximum tax that citizens are willing to pay before they switch to the non-market sector.

I now look for an equilibrium which satisfies these two incentive compatibility constraints. First, I will check whether the first best can be supported, i.e., whether the allocation with $m = 1$ and $e = 1$ can be achieved, so that the basic insights of the PCT generalize to this case without commitment.

To investigate the conditions under which the first-best allocation with $e = 1$ can be supported, observe that the maximum tax rate consistent with citizens incentive compatibility constraint is

$$T^{\max}(e = 1) = R + \alpha(1 - b).$$

Then the question of whether we can support the first-best allocation simply boils down to whether we can satisfy the ruler’s incentive compatibility constraint, (7) with this tax level. Combining this with (7), we have:

$$T^{\max}(e = 1) = R + \alpha(1 - b) \geq \Upsilon(e = 1) = 1 - \beta + R,$$

which is equivalent to the condition:

$$1 - \beta \leq \alpha(1 - b) \quad (10)$$

When (10) is satisfied, the basic message of the PCT goes through. Agents can enter into implicit agreements because the threat of punishment implied by the trigger strategies is sufficient to overcome the inherent commitment problems, and the first-best allocation can be achieved despite the fact that all political power is in the hands of the ruler. Condition (10) is more likely to be satisfied when agents are patient, and the outside options of citizens are not too attractive, so that the ruler can raise enough taxes in every period not to be tempted to tax more than the prescribed amount.

Next consider the case where (10) is not satisfied, so that the first-best investment level, $e = 1$, cannot be maintained. Can market participation by the citizens, $m = 1$ and some positive investment in the market sector, $e > 0$, be supported as an equilibrium? The answer is yes provided that we can satisfy the inequality that the maximum tax citizens are willing to pay is greater than the flow return to the ruler from grabbing everything:

$$T^{\max}(e) = R + e^{1-\alpha} - (1 - \alpha)e - \alpha b \geq \Upsilon(e) = (1 - \beta)e^{1-\alpha} + R, \quad (11)$$

The left-hand side, $T^{\max}(e)$, represents the incentive compatibility condition of the citizens, while the right-hand side, $\Upsilon(e)$, corresponds to the incentive compatibility condition of the ruler.

Figure 3 draws the left- and right-hand sides of this equation in the space of $e^{1-\alpha}$ and T . For low values of e , $T^{\max}(e)$ increases faster than $\Upsilon(e)$, so greater investment levels make it easier to satisfy both incentive compatibility conditions. However, the gap between $T^{\max}(e)$ and $\Upsilon(e)$ reaches its maximum at $e = \beta^{1/\alpha} < 1$, where the slopes of the two curves are equalized. After this point, $\Upsilon(e)$ grows faster than $T^{\max}(e)$. This reflects the fact that the incentive compatibility constraint of the ruler depends on output, whereas the incentive compatibility of the citizens depends on the difference between output and the cost of investment, which grows less than output.

As a result, we always have that $T^{\max}(e = \beta^{1/\alpha}) - \Upsilon(e = \beta^{1/\alpha}) > T^{\max}(e = 1) - \Upsilon(e = 1)$, and it is easier to satisfy both incentive compatibility constraints at $e = \beta^{1/\alpha}$ than at the first-best level of investment. This reflects the fact that at maximum effort, there are strong incentives for the ruler to grab everything today.

Figure 3 shows a case where $T^{\max}(e = \beta^{1/\alpha}) - \Upsilon(e = \beta^{1/\alpha}) > 0 > T^{\max}(e = 1) - \Upsilon(e = 1)$, so while the first-best cannot be supported, a range of investment levels, $e \in [e^*, e^{**}]$, can be supported as an equilibrium of the repeated game between the ruler and the citizens. As we will see in greater detail below, an equilibrium with $e = \beta^{1/\alpha}$ is preferred to any equilibrium with $e \in [e^*, \beta^{1/\alpha})$ by both the citizens and the ruler. So

we can focus on the set $e \in [\beta^{1/\alpha}, e^{**}]$ at the set of potential equilibria. It is clear that this set changes with the underlying parameters. For example, when β increases, the set becomes larger, and in particular, the highest investment that can be supported, e^{**} , increases (Figure 3, in fact, shows the case where, with the decline in β , the first-best level of investment can be supported).

This analysis also gives us a simple condition to check to see whether market production can be supported. If we cannot satisfy both incentive compatibility constraints when $e = \beta^{1/\alpha}$, then the set $[e^*, e^{**}]$ will be empty. Therefore, the condition for $m = 1$ to be supported is

$$\alpha\beta^{1/\alpha} \geq b \tag{12}$$

Now suppose that condition (12) is satisfied, so that the set $[e^*, e^{**}]$ is non-empty. Which of the many equilibria, or which of the investment levels in the set $[e^*, e^{**}]$, will be chosen? This paper has little to say on equilibrium selection, so any of these investments, as well as many others that are Pareto inferior to those on the frontier, may emerge in equilibrium. Nevertheless, it is useful to briefly discuss which of these equilibria will be the most preferred by the citizens and by the ruler, and how changes in the distribution of “power” between the two groups might affect equilibrium selection.

First, consider maximizing the ruler’s utility, (6) subject to (7) and (9) by choosing e and T . The solution to this problem is $e = e^{**}$ and $T^{\max}(e^{**})$: the ruler would like to maximize investment and choose the highest possible tax level given that investment. Incidentally, this is exactly what a social planner who wishes to maximize output would also choose.

In contrast, citizens would like to maximize (8) again subject to (7) and (9). We know that as long as $e \in [e^*, e^{**}]$, citizens’ incentive compatibility constraint, (9), will be satisfied, and citizens will never give the ruler more than the minimum amount to satisfy his incentive compatibility. Therefore, the ruler’s incentive productivity constraint, (7), has to hold as an equality. A simple way to characterize the solution is then simply to take this equation and substitute it back into (8). This gives the maximization problem:

$$\max T^{\max}(e) - \Upsilon(e) = \beta e^{1-\alpha} - (1 - \alpha)e,$$

i.e., citizens would like to maximize the difference between the left-hand and the right-hand sides of (11), which has the solution $e = \beta^{1/\alpha}$. Intuitively, increasing effort further is costly for the citizens, because they pay the additional investment costs, while the ruler

gets the benefits. Since they do not internalize the ruler's gains, they prefer $e = \beta^{1/\alpha}$ to the maximum investment that can be supported.

The Generalized Nash Bargaining solution between the ruler and the citizens is, in turn, given by:

$$\max_{e,T} \left[\frac{1}{1-\beta} e^{1-\alpha} - \frac{(1-\alpha)e}{1-\beta} + \frac{R}{1-\beta} - \frac{T}{1-\beta} - \frac{\alpha b}{1-\beta} \right]^\theta \left[\frac{T}{1-\beta} - \frac{R}{1-\beta} \right]^{1-\theta}$$

subject to the two incentive compatibility constraints, (7) and (9), where as before θ is the bargaining power of the citizens. According to the PCT, the allocation of political power between the two groups should not matter for the outcome. Here, it will clearly matter (as long as (12) is satisfied, and the set $[e^*, e^{**}]$ is non-empty). We already saw that when $\theta = 0$, i.e., when the ruler has all the bargaining power, we obtain $e = e^{**}$, whereas when $\theta = 1$, we obtain the citizens' most preferred solution, $e = \beta^{1/\alpha}$. Analysis of this maximization problem establishes that the general solution is $e(\theta)$, which is decreasing in θ with $\hat{e}(\theta = 0) = e^{**}$ and $\hat{e}(\theta = 1) = \beta^{1/\alpha}$.²⁰ That greater bargaining power for citizens reduces investment and efficiency is a somewhat surprising result, but is intuitive ex post. Recall that the problem is the inability of the ruler to commit to not taxing the returns from citizens' investments; so a naive intuition may have been that greater bargaining power for the citizens would reduce inefficiencies. However, the bargaining power θ does not affect the incentive compatibility constraint of rulers—instead, it determines which point we choose from the possible set of allocations. Since citizens bear the cost of investment and receive less than the full return, their preferred investment is less than that of the ruler. Greater bargaining power for the citizens selects an equilibrium closer to their desired point, with lower investment and greater net returns for them.

Next consider the case with ruler replacement, i.e., $q > 0$. Since replacement happens at the end of the period, the only difference from the above analysis is the value of

²⁰To see this, first note that the incentive compatibility constraint of the ruler, (7) has to hold (otherwise, both parties can be made better off). Using this condition and factoring out constants, the maximization problem can be rewritten as

$$\max_e [\beta e^{1-\alpha} - (1-\alpha)e - \alpha b]^\theta [e^{1-\alpha}]^{1-\theta}$$

Differentiating and simplifying, we obtain

$$\theta = 1 + \frac{\beta \hat{e}^{1-\alpha} - \hat{e}}{(\hat{e} - b)\alpha},$$

which gives $d\hat{e}/d\theta < 0$ in the range $\hat{e} \in [\hat{e}(\theta = 1) = \beta^{1/\alpha}, \hat{e}(\theta = 0) = e^{**}]$.

continued cooperation for the ruler. Taking this into account, the relevant comparison for the ruler is between grabbing everything, which has payoff $e^{1-\alpha} + R/(1 - \beta(1 - q))$, and taxing at the prescribed rate, which yields $T/(1 - \beta(1 - q))$. Both of these are different from the above expressions because the value of the future is less for the ruler due to the possibility of replacement. The ruler's incentive compatibility constraint then becomes:

$$T \geq \Upsilon(e) \equiv (1 - \beta(1 - q))e^{1-\alpha} + R, \quad (13)$$

and citizens' incentive compatibility constraint remains unchanged. This implies that the first-best can now be supported when

$$1 - \beta(1 - q) \leq \alpha(1 - b), \quad (14)$$

which is more restrictive than (10) for all $q > 0$. Intuitively, the possibility of replacement reduces the value of future cooperation for the ruler, and makes the first-best and the PCT more difficult to achieve.

The general solution also changes in the same direction. In terms of Figure 3, the curve for $\Upsilon(e)$ shifts up, and the range of investment levels that can be supported declines. The bargaining solution now corresponds to maximizing:

$$\max_{e,T} \left[\frac{1}{1-\beta}e^{1-\alpha} - \frac{(1-\alpha)e}{1-\beta} + \frac{R}{1-\beta} - \frac{T}{1-\beta} - \frac{\alpha b}{1-\beta} \right]^\theta \left[\frac{T}{1-\beta(1-q)} - \frac{R}{1-\beta(1-q)} \right]^{1-\theta},$$

subject to (9) and (13). It is straightforward to show that a greater q , i.e., a higher probability of replacement, reduces the equilibrium level of investment.²¹ This effect of the replacement probability, q , contrasts with the case with enforceable contracts, where q did not matter.

Finally, the corresponding condition for an equilibrium with $m = 1$ to be supported changes to

$$(\beta(1 - q))^{1/\alpha} > b. \quad (15)$$

Summarizing:

²¹The mathematical argument mirrors that of footnote 20, and now the relevant expression is

$$\theta = 1 + \frac{\beta(1 - q)\hat{e}^{1-\alpha} - \hat{e}}{(\hat{e} - b)\alpha},$$

which gives $d\hat{e}/dq < 0$ in the range $\hat{e} \in [\hat{e}(\theta = 1) = \beta^{1/\alpha}, \hat{e}(\theta = 0) = e^{**}]$.

Proposition 4 When rulers and citizens cannot commit to future transfers, the PCT and the efficient allocation can be supported by trigger punishment strategies provided that (14) is satisfied. Otherwise, the level of investment is less than the first best, $e = 1$. As long as condition (15) is satisfied, an equilibrium with market production, i.e., $m = 1$, but $e < 1$, can be supported. In this equilibrium, the level of investment is a decreasing function of the bargaining power of the citizens, θ , and of the replacement probability of the rulers, q .

This analysis therefore establishes limits on the reasoning of the PCT because of the inherent commitment problems in politics. Since there is no outside party with the coercion capability to enforce contracts between rulers and citizens, promises of rulers have to be self-enforcing or incentive-compatible. This puts limits on the society’s capacity to achieve efficient allocations, and on the applicability of the PCT. This is so, even though the model does not rule out any types of transfers between citizens and rulers on technological grounds.

5.2 Determinants of Policies and Institutions

The analysis and the comparative statics above provide us with a simple interpretation of the potential determinants of equilibrium institutions and policies. First, the distribution of political power between rulers and citizens, θ , matters for the equilibrium outcome when the PCT does not apply. More interesting, the horizon of the ruler matters. When rulers are impatient because they fear replacement by other competing rulers, self-enforcing agreements are harder to maintain because of standard repeated game reasoning. Therefore, better equilibrium policies will arise when rulers have longer horizons.²² Finally, better outside options for citizens (leaving only a small surplus to be shared between rulers and citizens in the market) make cooperation between citizens and rulers more difficult.

Who designs the game, or the “constitution”, may also be important, especially in thinking about institutions imposed on a society by external groups, such as colonial powers. If the political system is set up by the citizens, they will immediately choose $r = 1$, i.e., democratic institutions, where rulers do not have the power to tax them.

²²See Acemoglu and Robinson (2000b, 2002) for a different argument for why rulers who fear replacement may pursue the wrong policies for the society. There, rulers who fear replacement are more likely to resist the introduction of superior technologies or institutions when these changes can erode part of their incumbency advantage and potential future political power.

In contrast, if the original design is by the ruler, or by some political elite, who does not internalize the interests of the citizens, they will opt for $r = 0$. Though trivial, this observation may be important in thinking about why the European colonists introduced relatively democratic institutions with checks and balances on state and politicians' power in colonies where they settled in large numbers (i.e., where they became the "citizens" in terms of the model above), while establishing or maintaining oppressive and extractive institutions in colonies where they did not settle and wished to transfer resources from the native population to themselves.

Although the focus of this paper is not to construct a model that can be used to interpret a wide range of social situations, it is also instructive to attempt to incorporate "checks and balances" in this simple framework. To do this, I now extend the model in one dimension to introduce a measure of institutional controls on politicians: costly replacement of rulers. This analysis is useful both for providing comparative statics with respect to the extent of checks and balances, and to show the interaction between these types of institutional constraints on rulers and the implicit constraints that the rulers place on themselves via self-enforcing agreements.

The only difference from the baseline model is that I now assume that after the taxes are set, citizens can attempt to replace the politician, but such replacement costs c (in terms of the timing of events in subsection 4.1, the replacement decision is after step 3). I assume that this cost is incurred by all the citizens irrespective of whether they "support" the replacement of the ruler, so there is no "free-rider" problem.²³ If citizens attempt to replace the ruler, this is successful with probability p . The parameter p can be interpreted as a measure of the quality of checks and balances on politicians: when p is high, citizens can control the ruler better.

If the current ruler is successfully ousted from power, a new ruler is put in place the following period. I assume that if the ruler is ousted, he does not receive the tax revenue from the current period, and to simplify the analysis I also assume that this tax revenue is not received by the citizens either. Similar results are obtained with alternative assumptions, but the current set of assumptions simplifies the analysis.

Now suppose the ruler has set the tax T , and is expected to set the same tax in the future. Let us first ignore the incentive compatibility constraint of the ruler, and suppose that the same equilibrium will be played over time irrespective of whether citizens have

²³This refers to free-riding in the decision whether to oust the current ruler, different from "free-riding" in the investment decisions discussed in the next subsection.

attempted to replace, or have replaced, the ruler or not. Also to further simplify the discussion, set $q = 0$.

The citizens have a choice of whether to replace the ruler, at cost c , or go along with the implicit agreement. The value function of citizens at this point is:

$$W(e, T) = \max \left\{ \begin{array}{l} e^{1-\alpha} - (1-\alpha)e + R - c - T + \beta W(e, T); \\ e^{1-\alpha} - (1-\alpha)e + R - T + \beta W(e, T) \end{array} \right\}. \quad (16)$$

The upper branch corresponds to the choice to replace, and the lower branch applies when citizens do not attempt to replace. Notice that the continuation value with or without replacement is the same, $\beta W(e, T)$, since there will be a ruler following the optimal policy after this point, and we are assuming that this new ruler will play the same strategy. The only difference between the two branches is the cost of replacement, $-c$, immediately implying that citizens will never exercise their option to replace the ruler. It is costly, and along the equilibrium path, it creates no benefits.²⁴

Nevertheless, the option to replace may have an effect on the equilibrium because it may be beneficial for the citizens to replace a ruler *who has deviated*. In particular, consider a ruler who deviates and grabs all the output. Following this, the citizens and the ruler will play the no-cooperation game. So the citizens' continuation value, if they do not attempt to replace the ruler, is

$$\tilde{W}(e) = -(1-\alpha)e + \beta \widehat{W}, \quad (17)$$

where \widehat{W} is the value of the citizens and the no-cooperation continuation game given by (2), and $-(1-\alpha)e$ is the flow return this period, since they have invested e and all the output is grabbed by the ruler. This expression also incorporates the fact that if citizens do not replace the ruler now, they will not replace him in any of the following dates.

Next consider the value to the citizens after they attempt to replace the ruler (and reverting back to never replacing him after this period, if this attempt does not work), presuming that in the continuation game, citizens will cooperate with a new ruler:

$$\tilde{W}(e) = -(1-\alpha)e - c + \beta(1-p)\widehat{W} + \beta p W(e, T), \quad (18)$$

where $W(e, T)$ is the equilibrium value. Comparing (17) and (18), we see that as long as

$$c \leq p\beta \left(\widehat{W} - W(e, T) \right), \quad (19)$$

²⁴This conclusion holds a fortiori if, following an unsuccessful replacement attempt, the ruler and the citizens revert to no-cooperation.

citizens will attempt to replace the ruler. Since, by construction, $\widehat{W} - W(e, T) > 0$, condition, (19), implies that for sufficiently low costs of replacement, i.e., for $c \rightarrow 0$, citizens will attempt to replace rulers who deviate and grab all the output. Citizens' replacement option will then affect the incentive compatibility constraint of the ruler. In particular, when (19) holds, a ruler who deviates and grabs everything knows that he will be replaced with probability p . Since $\widehat{W} - W(e, T) > 0$ also in all future dates, the citizens will attempt to replace the ruler in every future period as well. Taking this into account, the ruler's incentive compatibility constraint changes from (7) to

$$T \geq (1 - \beta)(1 - p) \left[e^{1-\alpha} + \frac{R}{1 - \beta(1 - p)} \right],$$

Thus, the condition for the PCT to hold becomes

$$(1 - p)(1 - \beta) - \alpha(1 - b) \leq \frac{pR}{1 - (1 - p)\beta}, \quad (20)$$

which boils down to (10) when $p = 0$. This condition is more likely to hold when p is high. Therefore, better checks and balances on rulers, as captured by a higher value of p , make PCT more likely to hold.

When (20) does not hold, the allocation most preferred by the ruler will be given by the maximum investment that satisfies both the citizens's and the ruler's incentive compatibility constraints, thus by the maximum e that satisfies:

$$(1 - p)(1 - \beta) \left[e^{1-\alpha} + \frac{R}{1 - (1 - p)\beta} \right] \leq T^{\max}(e = 1) \equiv e^{1-\alpha} + R - (1 - \alpha)e - \alpha b,$$

or by \tilde{e} such that:

$$\alpha b - \frac{pR}{1 - (1 - p)\beta} \equiv [1 - (1 - p)(1 - \beta)] \tilde{e}^{1-\alpha} - (1 - \alpha) \tilde{e} = \alpha b - \frac{pR}{1 - (1 - p)\beta}. \quad (21)$$

It is straightforward to check that the solution to this equation, \tilde{e} , is increasing in p . Therefore, a better "technology" for citizens to replace the ruler will lead to greater equilibrium investments. A similar argument to the one before establishes that market production, $m = 1$, can be supported in this case as long as:

$$[1 - (1 - p)(1 - \beta)]^{1/\alpha} \geq b - \frac{pR}{\alpha[1 - (1 - p)\beta]}. \quad (22)$$

Summarizing this discussion:

Proposition 5 Consider the game with replacement in this subsection, and suppose that the cost of replacement, c , small, i.e., $c \rightarrow 0$. Then, if (20) holds, the PCT applies and the efficient level of investment can be supported. Better checks and balances, as captured by greater p , make (20) more likely to hold. When (20) does not hold, the efficient level of investment cannot be achieved. But as long as (22) holds, market production, $m = 1$, can be supported, and in this case, equilibrium investment is given by \tilde{e} that satisfies (21). Better checks and balances, i.e., greater p , increase equilibrium investment.

5.3 Free-Riding, Overinvestment and the Form of Taxation

The simple model discussed above also raises another set of interesting issues related to free-riding among the citizens, and the form of taxation. Once we relax this assumption, the model delivers a motivation for distortionary taxation.

Suppose that citizens do not coordinate their actions, and the ruler, like before, observes the aggregate income level, Y , and sets a lump-sum tax T that applies to each individual. Let us focus on the case where (14) does not hold but (12) holds. This implies that the first-best is not possible, but there exist equilibria with $m = 1$, market production, which feature $e < 1$. Now each individual is facing a lump-sum tax T , and since individuals are atomistic, they do not take their effect on Y into account. This immediately implies that the equilibrium with $e < 1$ is no longer possible. Each individual would like to invest up to $e = 1$, since each is infinitesimal and is, at the margin, the residual claimant of the returns from the additional investment. This behavior by all individuals will take aggregate output to $Y = 1 + R$, violating the incentive compatibility constraint of the ruler, and destroying the equilibrium with market production.

Is there a way to prevent this type of unraveling of the self-enforcing equilibrium? The answer is yes: move away from lump-sum taxation. As long as simple tax schedules conditional on individual income, y_j , are possible, the equilibrium tax schedule can be made *sufficiently distortionary* to induce exactly the right level of investment. For example, imagine now that the ruler sets the following linear tax schedule $T_j(y_j) = \tau_0 + \tau_1 y_j$. In response, investment in the market sector would be $e = (1 - \tau_1)^{1/\alpha}$. Suppose that the equilibrium that we would like to support has $\hat{e} < 1$ and tax level \hat{T} . Then in order to be able to support this equilibrium, and the ruler has to set the following tax schedule $\hat{e} = (1 - \tau_1)^{1/\alpha}$ and $\tau_0 = \hat{T} - \tau_1 (1 - \tau_1)^{(1-\alpha)/\alpha}$. In other words, the tax schedule has to discourage investment enough so that individuals do not overinvest and violate the

incentive compatibility constraint of the ruler. Summarizing:

Proposition 6 When citizens choose their investment levels individually, and the first-best level of investment cannot be supported, equilibrium taxes have to be distortionary to discourage citizens from investing up to $e = 1$.

Therefore, this simple model not only helps in analyzing the commitment problems, and the limitations of the PCT, but also suggests a reason for apparently inefficient methods of taxation (even when non-distortionary lump-sum taxes are available). These non-lump-sum taxes, at face value, appear to distort incentives. Nevertheless, once we are in the realm of self-enforcing agreements between rulers and citizens, an important consideration is to prevent citizens from overinvesting. Thus, the form of taxation has to be such that citizens do not have an incentive to invest too much, that is, citizens should not be the full residual claimant of the returns from their investments. At this stage, this explanation for why distortionary taxes may be preferred to non-distortionary alternatives is simply an implication of the model, and it is not currently clear how important this rationale is in practice. I leave a more detailed investigation of these issues to future work.

6 Concluding Remarks

There is growing interest and work on the determinants of policies and on the institutional choices that societies make. Why do some societies choose high taxes, while others opt for lower taxation? Why do some societies close their borders to trade, while others are more open? Why are bureaucracies more corrupt in some countries than in others? Why are some societies democratic, some parliamentary, some majoritarian, some relying on common law, etc.?

Much recent work is attempting to answer these questions. The first step towards an answer is to decide who makes the policy and institutional choices, and for whose interests. Or in other words, do collective choices maximize the output, surplus or welfare of the society as a whole, or do they select policies and institutions that benefit certain politically powerful groups, while being detrimental for aggregate output, surplus or welfare? Even though this latter question is, in some sense, antecedent to the ones posed above, there is not yet general agreement on the answers.

This paper provides a simple taxonomy for the possible answers to this question. Either we could subscribe to what I dubbed the “Political Coase Theorem,” which argues that

societies make efficient (output- or surplus-maximizing) choices, and distribute the gains from these choices between various groups and individuals. According to this approach, when societies choose inefficient policies, there will be strong political and social forces pushing them back towards efficient policies. Alternatively, societies may choose inefficient policies, not because of a failure in the political process, but because politicians' and citizens' beliefs were "mistaken". Finally, we can be in the realm of Theories of Social Conflict, which maintain that societies often choose the wrong policies and institutions, or even pursue disastrous courses of action, because these choices are not made for the benefit of the society as a whole, but for the benefit of those who control political power.

The bulk of the paper was devoted to arguing that Theories of Social Conflict provide the right empirical and theoretical framework for the analysis of the questions posed above. But why do politically powerful groups choose policies that reduce aggregate output rather than choosing the "right" policies, and redistributing the gains to themselves? The answer this paper developed is that there are serious commitment problems in politics placing severe limits on the reasoning of the Political Coase Theorem. In other words, *efficiency considerations cannot be separated from distributional conflicts*. The reasoning of the Political Coase Theorem presumes the possibility of political and economic trades between various individuals and groups. But these trades are intertemporal, and need to rely on contracts and promises. Typically, contracts and explicit promises are enforced by "the state". When it comes to contracts that the state, or social groups controlling the state, would like to write with the rest of the society, they will be non-enforceable. This implies that the allocation of political power creates an inherent commitment problem, undermining the potential to reach efficient outcomes (naturally, this does not deny the fact that political and economic forces will sometimes push towards more efficient social arrangements). I investigated how incentive-compatible promises can make up for lack of enforceable contracts, but generally fall short of achieving the efficient outcome (or of validating the Political Coase Theorem). I also developed some comparative statics from the simple model I used for this purpose.

This paper is only a preliminary attempt to highlight some of the important issues that are implicit in much of the recent political economy literature. I believe that the evidence is clear that Theories of Social Conflict provide the right framework for further analysis. But there are certainly factors other than commitment problems important in preventing the Political Coase Theorem from applying, and even if commitment problems

are what matter most, the way this paper modeled them may not be the most fruitful approach. So this paper may be thought of as a call for future research on investigating the major factors that cause inefficient policies and prevent the Political Coase Theorem from applying.

7 References

Acemoglu, Daron, Philippe Aghion and Fabrizio Zilibotti (2002) “Distance to Frontier, Selection and Economic Growth” NBER Working Paper, No. 9066.

Acemoglu, Daron, Simon Johnson and James A. Robinson (2001) “The Colonial Origins of Comparative Development: An Empirical Investigation,” *American Economic Review*, volume 91, No. 5, 1369-1401.

Acemoglu, Daron and James A. Robinson (2000a) “Why did the West Extend the Franchise? Democracy, Inequality and Growth in Historical Perspective,” *Quarterly Journal of Economics*, CXV, 1167-1199.

Acemoglu, Daron and James A. Robinson (2000b) “Political Losers as a Barrier to Economic Development,” *American Economic Review Papers and Proceedings*, 90, 126-130.

Acemoglu, Daron and James A. Robinson (2001) “Inefficient Redistribution,” *American Political Science Review*, v95, n3 (September 2001): 649-61.

Acemoglu, Daron and James Robinson (2002) “Economic Backwardness in Political Perspective” NBER Working Paper, No. 8831.

Acemoglu, Daron and Thierry Verdier (1998). “Property Rights, Corruption and the Allocation of Talent: A General Equilibrium Approach.” *Economic Journal*, 108, 1381-1403.

Aghion, Philippe and Patrick Bolton (1990) “Government and the Risk of Default: a Political-Economic Model of Strategic Role of Debt” In Rudi Dornbusch and Mario Draghi, editors, *Public Debt Management: Theory and History*, Cambridge, Cambridge University press, 1990, 121-45.

Anderlini, Luca and Leonardo Felli (1998) “Costly Coasian Contracts” London School of Economics working paper.

Aldrich, John H. (1983) “A Downsian Spatial Model with Party Activism.” *American Political Science Review* 77: 974-90.

Austen-Smith, David, and Jeffrey S. Banks (1999) *Positive Political Theory I: Collective Preference*. Ann Arbor MI: University of Michigan Press.

Baron, David P., and John Ferejohn (1989) “Bargaining in Legislatures.” *American Political Science Review* 83: 1181-206.

Becker, Gary S. (1983) “A Theory of Competition among Pressure Groups for Political Influence,” *Quarterly Journal of Economics*, 98, 371-400.

Becker, Gary S. (1985) “Public Policies, Pressure Groups, and Dead Weight Costs,” *Journal of Public Economics*, 28, 329-347.

Becker, Gary S. and Casey Mulligan (1998) “Deadweight Costs and the Size of Government” NBER Working Paper No.6789.

Besley, Timothy, and Stephen Coate (1998) “Sources of Inefficiency in a Representative Democracy: a Dynamic Analysis” *American Economic Review*, 88 139-156.

Buchanan, James M., and Gordon Tullock (1962) *The Calculus of Consent*. Ann Arbor MI: University of Michigan Press.

Coase, Ronald (1960) “The Problem of Social Cost” *Journal of Law and Economics*, vol. 3, pp. 1-31.

Coate, Stephen and Stephen Morris (1995) “On the Form of Transfers to Special Interests,” *Journal of Political Economy*, 103, 1210-1235.

Coate, Stephen and Stephen Morris (1998) “Policy Persistence,” *American Economic Review*.

Cain, Philip J. and Anthony G. Hopkins (1993) *British Imperialism: Innovation and Expansion 1688-1914*, Longman, New York.

Coatsworth, John H. (1999) “Economic and Institutional Trajectories in 19th Century Latin America,” in John H. Coatsworth and Alan M. Taylor eds. *Latin America and the World Economy since 1800*, Harvard University Press, Cambridge MA.

Crowder, Michael (1968) *West Africa under Colonial Rule*, Northwestern University Press, Evanston IL.

Curtin, Philip D. (1964) *The Image of Africa*, University of Wisconsin Press, Madison.

Davis, Lance E. and Robert A. Huttenback (1987) *Mammon and the Pursuit of Empire: The Political Economy of British Imperialism, 1860-1912*, Cambridge University Press, Cambridge UK.

Denoon, Donald (1983) *Settler Capitalism: The Dynamics of Dependent Development in the Southern Hemisphere*, Clarendon Press, Oxford UK.

Demsetz, Harold (1967) “Toward a Theory of Property Rights” *American Economic Review*, volume 57.

Dixit, Avinash, Gene M. Grossman, and Faruk Gul (2000) “The Dynamics of Political Compromise.” *Journal of Political Economy* 108: 531-568.

Dixit, Avinash and John Londregan (1995) “Redistributive Politics and Eco-

conomic Efficiency,” *American Political Science Review*, 89, 856-866.

Dixit, Avinash and Mancur Olson (2000) “Does Voluntary Participation Undermine the Coase Theorem?” *Journal of Public Economics* vol 76, n3, 309-35.

Djankov, Simeon, Rafael, La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer (2002), ”The Regulation of Entry,” *Quarterly Journal of Economics*, CXVII,
Evans, Peter, (1998) *Embedded Autonomy: States and Industrial Transformation*, Princeton University Press.

Farrell, Joseph (1987) “Information and the Coase Theorem” *Journal of Economic Perspectives* vol 1, 113-29.

Ferejohn, John (1986) “Incumbent Performance and Electoral Control,” *Public Choice* 50: 5-25.

Gann, Lewis. H. and Peter Duignan (1962) *White Settlers in Tropical Africa*, Penguin, Baltimore MD.

Gann, Lewis. H. and Peter Duignan (1979) *The Rulers of Belgian Africa*, Princeton University Press, Princeton NJ.

Glaesar, Edward and Andrei Shleifer (2002) “Legal Origins” forthcoming *Quarterly Journal Economics*.

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

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

Grossman, Herschel and M-S. Kim (1995) “Swords or Ploughshares? A Theory of the Security of Claims to Property,” *Journal of Political Economy*, 103, 1275-1288.

Grossman, Sanford and Oliver Hart (1986) “The Costs and Benefits of Ownership: a Theory of Vertical and Lateral Integration” *Journal of Political Economy*, vol. 94, pp. 691-719.

Ha-Cheong Yeon (1988) “Economy” in *A Comparative Study of South and North Korea*, edited by the national Unification Board, Seoul, Korea.

Hall, Peter A. and David Soskice (2001) *Varieties of capitalism : the institutional foundations of comparative advantage*, New York; Oxford University Press.

Hart, Oliver and John Moore (1990) “Property Rights and the Nature of the Firm” *Journal of Political Economy* Vol 98, pp 1119-1158.

Herbst, Jeffery I. (2000) *States and Power in Africa: Comparative Lessons in*

Authority and Control, Princeton University Press, Princeton NJ.

Knack, Steven and Philip Keefer (1995) “Institutions and Economic Performance: Cross-Country Tests Using Alternative Measures,” *Economics and Politics*, 7, 207-227.

Krueger, Alan, and Mikael Lindahl (2001), “Education for Growth in Sweden and the World,” *Journal of Economic Literature*.

Kupperman, Karen Ordahl (1988) “England’s Dream of Caribbean Empire” *William and Mary Quarterly*, vol 45, pp. 70-99, reprinted in Karen Ordahl Kupperman (editor) *Major problems in American colonial history : documents and essays*, Boston, MA : Houghton Mifflin, 2000.

Libecap, Gary D. (1989) *Contracting for property rights*, Cambridge, Cambridge University Press.

Maddison, Angus (2001) *The World Economy: A Millennial Perspective*, Development Centre of the Organization for Economic Cooperation and Development, OECD, Paris.

Manning, Patrick (1982) *Slavery, Colonialism, and Economic Growth in Dahomey, 1640-1980*, Cambridge University Press, New York NY.

Mauro, Paulo (1995) “Corruption and Growth,” *Quarterly Journal of Economics*, CX, 681-712.

McKelvey, Richard D. and Talbot Page (1999) “Taking the Coase Theorem Seriously” *Economics and Philosophy* vol 15, n2, 235-47.

Meltzer, Allan H. and Scott Richards (1981) “A Rational Theory of the Size of Government” *Journal of Political Economy* volume 89, #5, 914-927.

Mukand Sharun and Dani Rodrik (2002) “In Search of Holy Grail: Policy Convergence, Experimentation and Economic Growth” Harvard Kennedy School mimeo.

Myerson, Roger B (1995) “Analysis of Democratic Institutions: Structure, Conduct and Performance” *Journal of Economic Perspectives* 9: 77-89.

North, Douglass C. (1981) *Structure and Change in Economic History*, W.W. Norton & Co., New York.

North, Douglass C. (1990) “A Transaction Cost Theory of Politics” *Journal of Theoretical Politics* 2(4).

Olson, Mancur, (1965) *The Logic of Collective Action*, Harvard University Press, Cambridge MA.

Olson, Mancur, (2000) *Power and Prosperity : Outgrowing Communist and Capitalist Dictatorships*, New York, NY, Basic Books.

Peemans, Jean-Philippe (1975) “Capital Accumulation in the Congo under Colonialism: The Role of the State,” in Lewis. H. Gann and Peter Duignan eds. *Colonialism in Africa 1870-1960, Volume 4 The Economics of Colonialism*, Hoover Institution Press, Stanford CA.

Persson, Torsten and Guido Tabellini (2000) *Political Economics: Explaining Economic Policy*, The MIT Press, Cambridge MA

Persson, Torsten and Lars Svensson (1989) “Why a Stubborn Conservative Would Run a Deficit: Policy with Time-Inconsistent Preferences” *Quarterly Journal of Economics*, 104, 325-46.

Piketty, Thomas, (1995) “Social Mobility and Redistributive Politics, *Quarterly Journal of Economics*, 110.

Rajan, Raghuram G. and Zingales, Luigi (2000) “The Tyranny of Inequality” *Journal of Public Economics* vol 76, n3 521-58.

Reno, William (1995) *Corruption and State Politics in Sierra Leone*, Cambridge University Press, New York NY.

Ringer, F. (1979) *Education and Society in Modern Europe*, Indiana University Press, Bloomington.

Roberts, Andrew (1976) *A History of Zambia*, Heinemann, London.

Robinson, Ronald E. and John Gallagher (1961) *Africa and the Victorians: The Official Mind of Imperialism*, MacMillan, London.

Robinson James A. (1998) “Theories of ”Bad” Policy”, (1998) *Journal of Policy Reform*, 3, 1-46.

Rodrik, Dani (1986) “Tariffs, Subsidies and Welfare with Endogenous Policy,” *Journal of International Economics*,”, 21, 285-299.

Romer, David (1997) “Misconceptions and Political Outcomes” NBER Working Paper No.6117.

Romer, Thomas (1975) “Individual Welfare, Majority Voting and the Properties of a Linear Income Tax” *Journal of Public Economics*, 4, 163-85.

Shleifer, Andrei and Robert W. Vishny (1998) *The Grabbing Hand: Government Pathologies and their Cures*, Harvard University Press, Cambridge MA.

Snyder, James M., Jr. (1990) “Resource Allocation in Multiparty Elections.”

American Journal of Political Science 34: 59-73.

Spiller, Pablo and Mariano Tommasi (2002) “The Institutional Foundations of Public Policy: A Transactions Approach with Application to Argentina” Berkeley mimeo.

Staiger, Robert W. and Guido Tabellini (1987) “Discretionary Trade Policy and Excessive Protection,” *American Economic Review*, 77, 340-348.

Stigler, George (1966) *The Theory of Price*, Third Edition, Macmillan.

Tilly, Charles (1990) *Coercion, Capital and European States, AD 990-1990*, Blackwell, Cambridge MA.

Tullock, Gordon (1983) *The Economics of Income Redistribution*, Kluwer-Nijhoff, Boston MA.

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

Weingast, Barry R. (1998) “Political Stability and Civil War: Institutions, Commitment and American Democracy,” in R.H. Bates et al. eds. *Analytic Narratives*, Princeton University Press, Princeton NJ.

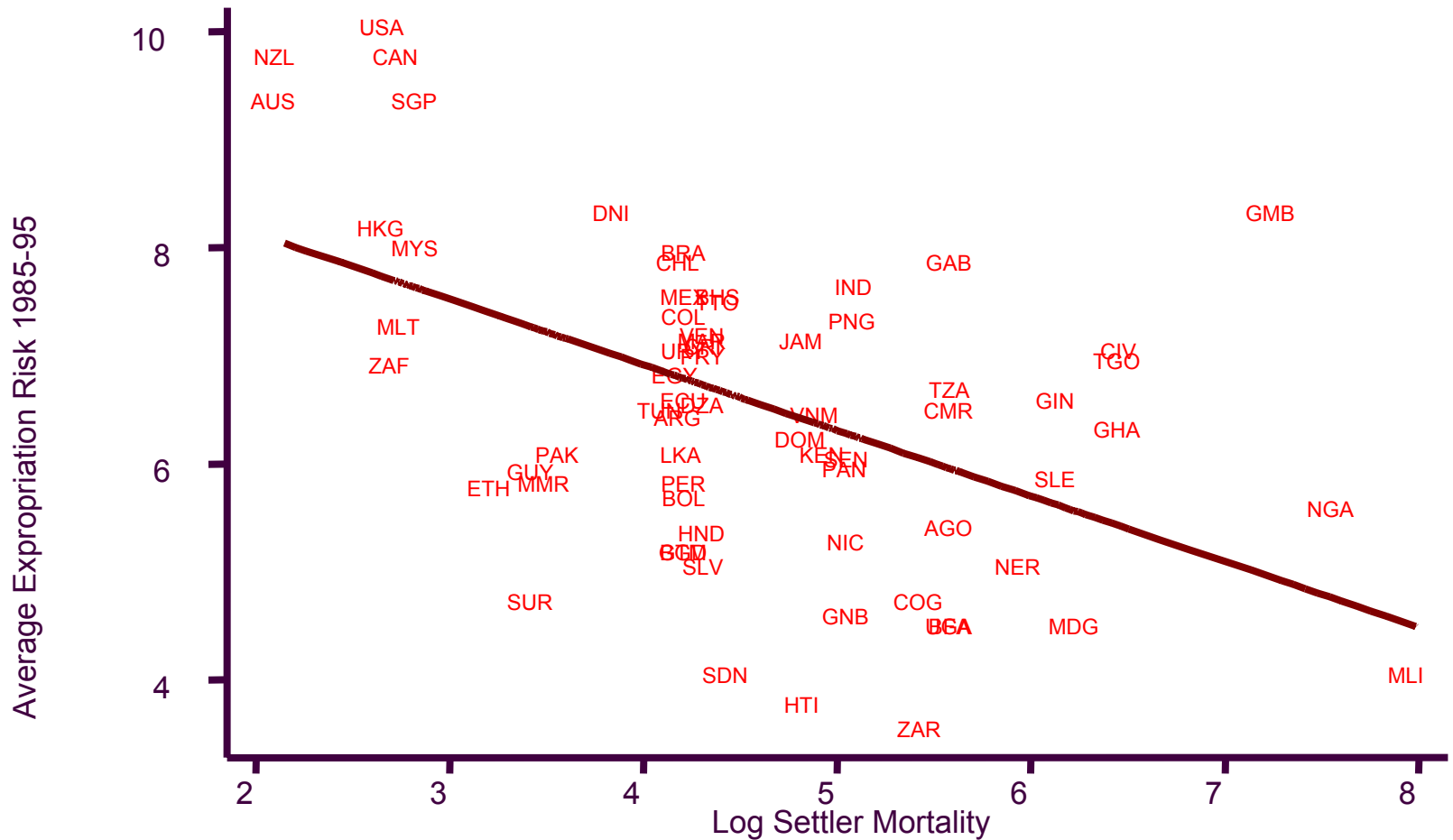
Wittman, Donald (1989) “Why Democracies Produce Efficient Results,” *Journal of Political Economy*, 97, 1395-1426.

Wittman, Donald (1995) *The Myth of Democratic Failure: Why Political Institutions are Efficient*, University of Chicago Press, Chicago IL.

Wilson, James D. (1990) “Are Efficiency Improvements in Government Transfer Policies Self-Defeating in Political Equilibrium,” *Economics and Politics*, 2, 241-258.

Young, Crawford (1994) *The African Colonial State in Comparative Perspective*, Yale University Press, New Haven CT.

Figure 1: Log European settler mortality and institutions among the former colonies



**Figure 2: GDP per capita
in South and North Korea
from Maddison (2001)**

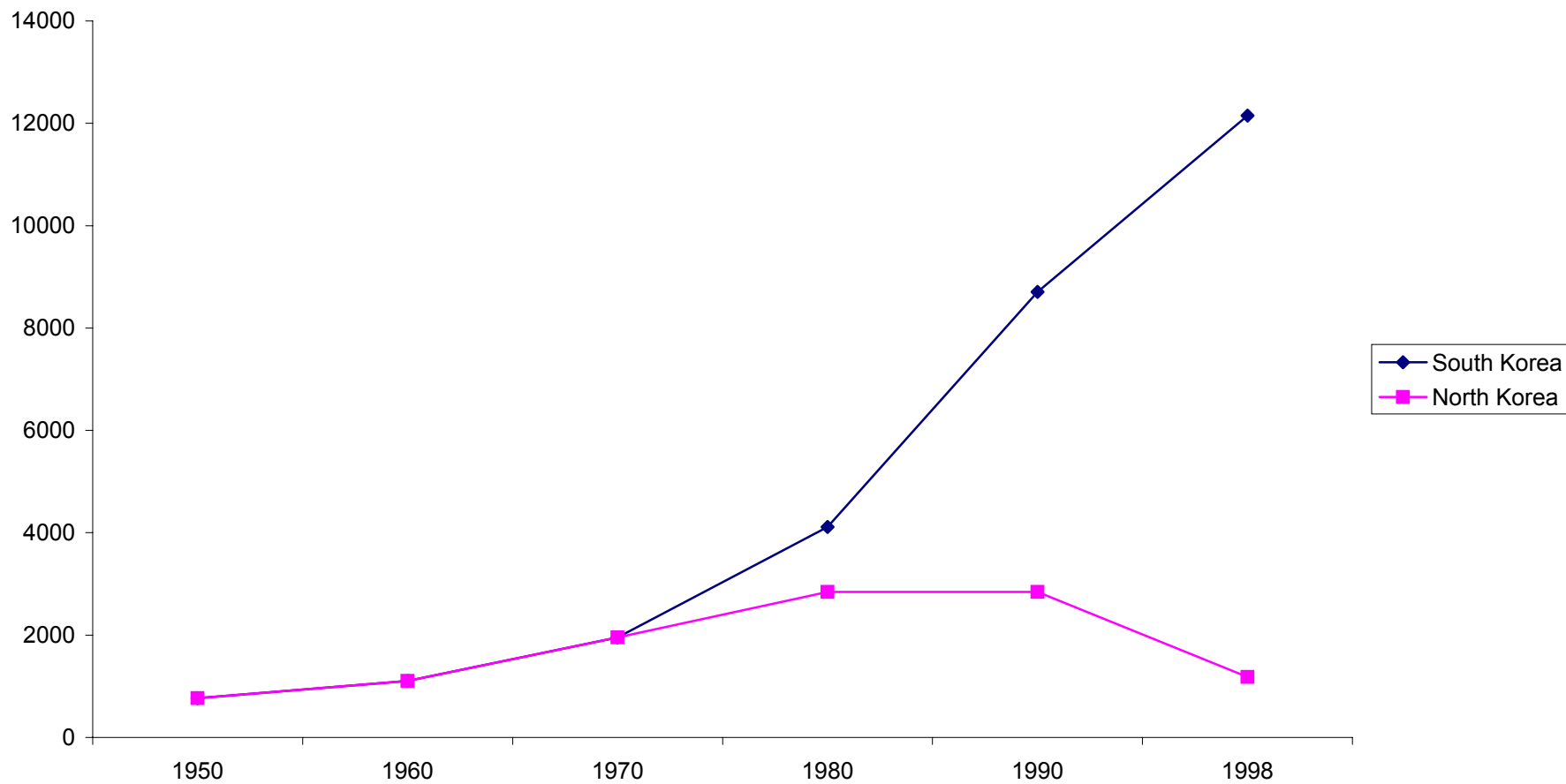


Figure 3: Equilibria

