



Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Comparative Economics 31 (2003) 620–652

Journal of
COMPARATIVE
ECONOMICS

www.elsevier.com/locate/jce

Why not a political Coase theorem? Social conflict, commitment, and politics

Daron Acemoglu

Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142, USA

Received 15 March 2003; revised 4 September 2003

Acemoglu, Daron—Why not a political Coase theorem? Social conflict, commitment, and politics

Do societies choose inefficient policies and institutions? An extension of the Coase theorem to politics would suggest the answer is no. This paper discusses 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. We conclude that these inefficient institutions and policies are chosen because they serve the interests of politicians or social groups that hold political power at the expense of the rest. The theoretical case depends on 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. *Journal of Comparative Economics* 31 (4) (2003) 620–652. Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142, USA.

© 2003 Association for Comparative Economic Studies. Published by Elsevier Inc. All rights reserved.

JEL classification: H2; N10; N40; O1

Keywords: Political economy; Institutions; Commitment; Social conflict; Belief differences; Appropriate institutions; Economic development; Colonialism

1. Introduction

Many economists and social scientists believe that government policies, economic, political and legal institutions, and more broadly, the organization of society affect economic outcomes. Nevertheless, despite important theoretical advances, we still lack an

E-mail address: daron@mit.edu.

0147-5967/\$ – see front matter © 2003 Association for Comparative Economic Studies. Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.jce.2003.09.003

organizational framework to analyze the determinants of policies and institutions.¹ Why do certain societies choose different policies, different institutions, and radically different ways of organizing their lives?² This paper provides a simple taxonomy of the existing approaches and argues for the relevance of one particular approach.

The first category in this taxonomy is designated as “the political Coase theorem” (PCT). The Coase theorem maintains that, if property rights are well-defined and there are no transaction costs, economic agents will contract to achieve an efficient outcome, irrespective of who holds the property rights on particular assets (Coase, 1960; Stigler, 1966). An extension of this reasoning to the political sphere suggests 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, because societies choose, at least approximately, the appropriate policies and institutions for their conditions.

The second category is designated “theories of belief differences” (or a modified political Coase theorem). According to this view, societies may choose different policies, with very different implications, because they or their leaders disagree about what would be good for the society. Prevailing uncertainty about the correct policies and institutions leads well-meaning political actors to differ about what is good for their own people. Societies in which the leaders or the electorate turn out to be right *ex post* are those that prosper. As with the PCT, strong forces prevent the implementation of policies that are known to be bad for the society at large; hence, we label this approach as modified PCT.

The third category is “theories of social conflict.” According to this perspective, societies choose different policies, some of which are 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 for which internal conflict within the society leads to inefficient choices and those for which inefficient institutions and policies are imposed on societies from the outside, e.g., by colonial powers.³

¹ Austen-Smith and Banks (1999) and Persson and Tabellini (2000) are excellent introductions to recent advances in political economy.

² There is an important distinction between policies and institutions. Policies can be considered as choices made within a given political and social structure, e.g., the tax rate and fiscal policies. In contrast, institutions can be thought of as determinants of the political and social structure that are more durable and, as such, constrain future choices and policies. These include whether the society is democratic or not, the nature of the legal constraints on the government, and 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. Our focus is on why both inefficient policies and inefficient institutions are chosen and persist. Hence, for most of this paper we will not distinguish between policies and institutions.

³ There are many interesting theories that combine features from the three groups. As long as they assign a prominent role to social conflict, this taxonomy classifies them in the third category. Several interactions are worth mentioning briefly. In a number of theories featuring social conflict, societies make different choices because of differences in their economic conditions, but generally strong forces pushing for efficient outcomes are not present

The major difference in the taxonomy is between the theories of social conflict, which emphasize the prevalence of systematically inefficient government policies and institutional arrangements and the first two approaches, which stress the presence of social forces that preclude these types of inefficiencies. We argue that the PCT, in its simple form or in its modified version depending on belief differences, is not an appropriate framework for analyzing policy and institutional differences across countries. 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. Hence, the theories of social conflict provide the correct perspective for analyzing policy and institutional differences.

A major challenge for the theories of social conflict is to identify the specific transaction costs that prevent the application of the PCT. 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?

This paper argues that even though the PCT may be a useful benchmark, its applicability is limited because of inherent commitment problems associated with political power. It develops a simple model where commitment problems prevent efficient predation by powerful groups and lead to inefficient policies and institutions.

Underlying the Coase theorem is the ability to write enforceable contracts. Therefore, any enforcement problem potentially limits the applicability of the Coase theorem.⁴ In the context of the PCT, widespread enforcement problems arise because most contracts are enforced by the state. Contracts that the state, or social groups controlling the state, would like to write with others, e.g., the citizens, will be non-enforceable by definition because groups controlling the state cannot commit to not using their power to renege on their promises or to not changing the terms of the contract. Hence, the allocation of political power creates an inherent commitment problem that undermines the potential to reach efficient outcomes.

(e.g., Engermann and Sokoloff, 1997). In other models, certain groups may attempt, or manage, to convince others that their most-preferred policies also benefit the society at large, e.g., Coate and Morris (1995). In yet others, some societies choose different policies initially because of differences in beliefs, but then these policies create or strengthen their own political constituencies who support the continued implementation of these inefficient policies, e.g., Acemoglu et al. (2002).

Another set of approaches, especially popular among sociologists and political scientists, is also worth mentioning. Although maintaining that many policies and institutions are inefficient, the literature does 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. Tilly (1990) and Herbst (2000), who stress the importance of the emergence of the nation-state for economic development, are characteristic of this approach. From this perspective, whether the nation-state emerges or not is a consequence of other unrelated factors, for example population density or the frequency of wars.

⁴ Enforcement problems may arise from incomplete information, contracting costs or bounded rationality, e.g., Anderlini and Felli (1998), Dixit and Olson (2000), Farrell (1987), and McKelvey and Page (1999). We focus on the commitment problems, because of their importance to the PCT. The literature on transaction costs and the organization of the firm, e.g., Williamson (1985), Grossman and Hart (1986), and Hart and Moore (1990), focuses on why the distribution of property rights may matter for incentives if contracts are incomplete. Informal attempts to extend the reasoning of transaction costs to politics include North (1990) and Spiller and Tommasi (2002).

The commitment problem associated with the PCT has two components. First, those in power cannot commit to not using their 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 former rulers no longer possess the political power to enforce such promises. This double commitment problem restricts the potential remedies available to combat inefficiencies. Nevertheless, because the relationship between the state and the citizens is repeated, some commitment based on reputation and supported by the threat of future punishments may be possible. As a result, the extent to which the PCT will be applicable depends on the possibility of commitment through constitutions or other institutions and on how good a substitute this type of reputation-based commitment is for enforceable contracts. The extent of distributional conflict between various social groups will also affect the relevance of the PCT.

To focus on commitment problems inherent in political situations, the theoretical analysis in this paper allows unrestricted transfers and taxes, including non-distortionary lump-sum taxes. The inefficiencies arise not because of any restrictions on the technology of taxation but because of the political-economic interactions between different groups and agents. Despite the availability of non-distortionary taxes, the model suggests that, in this type of repeated game environment, the equilibrium may involve distortionary taxes. The allocation must 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; hence, they will have a tendency to overinvest, which may violate the incentive compatibility constraint of the ruler. Consequently, distortionary taxation may be necessary to guarantee levels of investment consistent with the incentive compatibility constraint of the ruler.

Robinson (1998) surveys the literature on distortionary policies of governments and makes a similar distinction is made between bad policies due to belief differences and those originating from social conflict. The most influential models of distortionary policies are the voting models in which the median or the decisive voter may choose policies that redistribute resources from society as a whole to himself or to his group (Romer, 1975; Meltzer and Richards, 1981; Persson and Tabellini, 2000). The political science literature examines how voting behavior and the organization of parties interact to produce equilibrium policies, e.g., Aldrich (1983), Baron and Ferejohn (1989), Dixit and Londregan (1995), Myerson (1995), and Snyder (1990). Other papers emphasize the conflict between bureaucrats or politicians and the society, for example, Buchanan and Tullock (1962), Ferejohn (1986), Persson et al. (1997), and Shleifer and Vishny (1998). However, these papers do not consider why politically powerful groups cannot extract resources from the rest of the society in an efficient manner. In fact, much of this literature rules out efficient methods of redistribution and takes for granted that rent-maximizing behavior by rulers or the government will result in inefficiencies. Our focus is on why efficient policies fail to appear. In this respect, our paper is related more closely to North (1981), Libecap (1989), and Olson (2000), who discuss how inefficient policies result from distributional conflicts. For example, North (1981) suggests that rulers will choose the

system of property rights in order to maximize their return, leading to inefficiencies, but he also places considerable emphasis on differences in beliefs. However, neither North nor Libecap nor Olson considers explicitly why a version of the PCT would not be applicable.

By providing a rationale for inefficient methods of taxation, this paper also relates to a few studies that investigate the reasons why societies choose inefficient redistributive policies rather than lump-sum taxation and transfers. Rodrik (1986), Wilson (1990) and Becker and Mulligan (1998) argue that, if the amount of redistribution is endogenous, politicians may 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 a two-period political economy model, Besley and Coate (1998) show how certain types of inefficiencies may arise because efficient policies would affect the identity of who is in power and emphasize the importance of commitment problems.⁵ Acemoglu and Robinson (2000b, 2002) develop a theory in which elites may want to block the introduction of new and efficient technologies because this change may reduce their future political power. Rajan and Zingales (2000) show how organizations make inefficient choices because each group, or agent, is worried that others in the organization will get richer and demand more concessions in the future. Finally, Acemoglu and Robinson (2001) show that inefficient methods of redistribution, rather than more efficient alternatives, may arise as a strategy for maintaining future political power. In a related vein, Persson and Svensson (1989) and Aghion and Bolton (1990) consider the use of fiscal policy to affect future elections. None of these papers address the general issue of how commitment problems undermine the PCT, nor do they analyze repeated games in which punishment strategies may substitute for a lack of formal commitment.⁶

The rest of the paper is organized as follows. The next section revisits and formalizes the taxonomy of various approaches to the determination of policies and institutions. 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 PCT will not apply. This analysis also develops some simple comparative statics, and indicates why distortionary taxes may be necessary to reduce overinvestment by citizens. Section 6 concludes.

⁵ Other 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 Acemoglu and Robinson (2000a), who argue that the introduction of democracy in 19th century Europe was a commitment by the rich elite to redistribution in the future.

⁶ In focusing on infinite-horizon models with self-enforcing arrangements, this paper is also related to Dixit et al. (2000), who analyze self-enforcing political deals between groups with different interests.

2. A simple taxonomy

To emphasize the differences between various approaches and to build a simple taxonomy, consider the following model in which Y denotes aggregate output or consumption, which we take to represent social welfare.⁷ Moreover, suppose that

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

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

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

The PCT maintains that there are strong forces leading societies towards some $P^*(X)$ in $\mathbb{P}(\cdot | X)$. The basic idea is that, if a society is pursuing a policy $P(X) \notin \mathbb{P}(\cdot | X)$, a switch to $P^*(X) \in \mathbb{P}(\cdot | X)$ will create aggregate gains. If these gains correspond to a Pareto improvement, all political systems will implement this change. If the change creates only a potential Pareto improvement, part of the gains can be redistributed to those that lose using various mechanisms or, at the very least, the winners can lobby or vote for the beneficial change. Several authors espouse limited forms of this PCT. Becker (1983, 1985), demonstrates how competition between pressure groups could create a force toward efficient policies. Wittman (1995) develops this argument further and formulates an informal PCT for democratic societies. Wittman argues that democratic societies with rational voters generally produce Pareto efficient, even wealth-maximizing, outcomes. In fact, Wittman's argument does not rely much on democratic institutions so that his reasoning could apply to nondemocratic societies.

The PCT does not imply policy convergence, however. To the extent that $\mathbb{P}(\cdot | X)$ is not a singleton, there can be considerable policy differences between two identical societies. But the performance of these two societies should not be appreciably different. As an example, consider the differences in policies regarding the role of the government between the Anglo-Saxon economies, in particular, the US and the UK, and the Continental European countries. These do not seem to lead to major differences in the economic performance between these two sets of countries.⁸

However, across a broader cross-section of societies, we see major differences in policies and institutions, e.g., free-market policies in Hong Kong and heavy government involvement and widespread corruption in Indonesia. 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 its own situation, but they achieve different outcomes because their situations, i.e., their

⁷ This assumption avoids the complications raised by Pareto comparisons and helps us focus on the main point.

⁸ See, for example, the discussion in Hall and Soskice (2001) on the costs and benefits of different types of capitalism.

X 's, are different. More specifically, for two societies with characteristics X and $X' \neq X$, we have $F(X, P^*(X)) \neq F(X', P^*(X'))$; moreover, $F(X, P^*(X)) > F(X, P^*(X'))$ and $F(X', P^*(X')) > F(X', P^*(X))$. Thus, the PCT suggests that Indonesian institutions are chosen appropriately for its circumstances and not inefficiently.⁹ Hence, to refute the applicability of the PCT, we must find systematic evidence that there are societies choosing P while $F(X, P) < F(X, P')$ for some feasible alternative P' , or simply that $P \notin \mathbb{P}(\cdot | X)$. In other words, we must find societies that persistently pursue wrong policies with significant output and welfare consequences.¹⁰

The theories of belief differences (the modified PCT), on the other hand, emphasizes that some subset of X, X_u , is uncertain. To simplify the notation, 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 $X_u \neq X'_u$. In other words, these uncertain characteristics affect which policies are beneficial for the society. Suppose that politicians, or the societies at large, have beliefs, denoted by $G(X_u)$, over the actual distribution of X_u . Assume that social welfare maximization corresponds to the maximization of expected aggregate output. Define:

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

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, i.e., 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., $F(X_c, X_u, P^*(X_c, G)) \neq F(X_c, X_u, P^*(X_c, G'))$ for $G \neq G'$.

As an example, suppose that the North Koreans chose socialist policies and government ownership because they believed these policies to be welfare-enhancing, while South Korea, which presumably had the same characteristics, X_c and X_u , chose a capitalist development path. Ex post, the South Koreans are correct; hence, they actually adopted the appropriate policies, and prosper, while the North Koreans suffer poverty and famine.¹¹ To refute this class of models, we must establish the existence of societies pursuing policies that could not be beneficial for the society under any plausible scenario. Denoting the set of admissible beliefs by \mathbb{G} , if, for two feasible policies, denoted P and P' , $\int F(X_c, X_u, P') dG \geq \int F(X_c, X_u, P) dG$ for all $G \in \mathbb{G}$, P should never be chosen.

Finally, according to the theories of social conflict, societies often choose some policy vector $P(X) \notin \mathbb{P}(\cdot | X)$ knowingly to maximize the payoffs of those who hold political

⁹ Glaesar and Shleifer (2002) explain why Britain and France chose very different legal codes and systems that were appropriate to the underlying circumstances of each country.

¹⁰ By refuting the PCT, we mean that we can find significant and quantitatively important inefficiencies in the institutions and policies of some societies. Of course, this statement begs the question of what is significant and quantitatively important. A refutation of the PCT does not imply that there are no forces pushing toward more efficient arrangements.

¹¹ Piketty (1995) develops an interesting theory of policy differences arising from belief differences. Individuals vote over the degree of redistribution in the economy based on their beliefs of the importance of individual effort in economic success. In turn, these beliefs evolve as a result of various economic interactions and tax policies. Romer (1997) and Mukand and Rodrik (2002) also stress the importance of these issues.

power and not to maximize social welfare or aggregate income. To emphasize the difference between this approach and the PCT, consider another vector of variables, Z , which do not directly affect Y , so that $P^*(X)$ is independent of Z . Nonetheless, these variables may influence the equilibrium policy; hence we write $P(X, Z)$. Changes in Z will have no direct effect on output but they may have an indirect impact on it 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 because it influences X , but has no direct effect on F . At this level of generality, the theories of social conflict is a residual group; if we can find certain societies that pursue inefficient policies systematically and knowingly, they fall into this category. However, the usefulness of these theories depends on whether they can identify a mechanism to explain why political and economic bargains are not struck to achieve better policies and institutions. In other words, these theories need to identify the salient transaction costs that limit the scope of the PCT and also find institutional or other social variables, i.e., the Z 's, that affect the degree of inefficiency of equilibrium policies.

3. The historical data: an interpretation

This section briefly develops the arguments that cross-country differences in policies and institutions are important determinants of economic performance and that the origins of these differences do not lie in different perceptions of the peoples and the leaders, but rather in the social conflicts that exist between these leaders, or the social groups that these leaders represent, and the rest of the society. We provide examples of inefficient institutions and policies that account for quantitatively large variations in economic performance. Societies may choose inefficient policies and institutions because of internal conflict or because these choices are imposed on them externally. Although inefficiencies arising from internal conflict are important, this section presents two examples of inefficient institutions imposed by outside forces in the latter part of this section. These episodes establish clearly that institutional choices are not always made in response to different economic circumstances. In other words, these episodes identify sources of exogenous variation in policies and institutions.

3.1. Differences in institutions and economic outcomes

Significant cross-country differences exist in the organization of economic and political activity. To be concrete, let us focus on economic institutions, e.g., the degree of enforcement of stable property rights, the extent of equal opportunity, and the feasibility of imposing entry barriers. A voluminous literature documents large cross-country differences in economic institutions and documents a strong correlation between these institutions and economic performance. Knack and Keefer (1995) consider measures of property rights enforcement compiled by international business organizations, Mauro (1995) looks at measures of corruption, and Djankov et al. (2002) compile measures of entry barriers across countries. Many studies examine the variation in educational institutions and the corresponding differences in human capital, e.g., Ringer (1979), Krueger and Lindahl (2001), and Hanushek and Kimko (2000). These authors

find substantial differences in the measures of economic institutions and a significant correlation between these measures and various indicators of economic performance. For example, Djankov and his co-authors document that while the total cost of opening a medium-size business in the United States is less than 2 percent of GDP per capita in 1999, the same cost is 270 percent of GDP per capita in Nigeria, 116 percent in Kenya, 91 percent in Ecuador, and 495 percent in the Dominican Republic. Such entry barriers are highly correlated with various economic outcomes, including the rate of economic growth and the level of development.

According to the PCT, this type of correlation might not establish that countries are choosing the wrong institutions because the United States differs from Nigeria, Kenya and the Dominican Republic in its economic characteristics, i.e., its X 's. Different X 's require different optimal policies and institutions so that $X \neq X'$ implies $P^*(X) \neq P^*(X')$ and $F(X, P^*(X)) \neq F(X', P^*(X'))$. Consequently, these observations may be optimal responses of different societies to their own varying conditions. Given the circumstances in the Dominican Republic, it is not worth investing in the arrangements to reduce the costs of opening and doing business. Alternatively, in theories of belief differences (the modified PCT), the people of the Dominican Republic may believe that high entry barriers are good for the society.¹²

To refute the applicability of the PCT, we must show that otherwise identical, or at least similar, societies choose different institutions and policies for reasons that do not affect economic outcomes directly and that they experience different economic performances as a result of these choices, which is essentially an instrumental variables approach. In other words, we seek social experiments in which, for political or historical reasons, societies have significantly different institutions. In addition, to refute the modified PCT, we need to show that the variation captured by the Z 's is not due solely to differences in beliefs. These considerations make us focus on external sources of variation in policies and institutions to develop the empirical case. In practice, internal dynamics are likely to be at least as important as external factors, but also make it harder to develop the case that differences in policies and institutions do not reflect differences in X 's.

3.2. Colonialism and institutional development

European colonization provides almost a natural experiment to investigate these issues. Beginning in the late 15th century, Europeans dominated and colonized much of the rest of the world. Together with European dominance came the imposition of various types of institutions in the colonies. Most interesting for our purposes, Europeans imposed significantly different institutions and social power structures in different parts of the world. Acemoglu et al. (2001) document that, in a large number of colonies, especially those in

¹² An example of optimal non-enforcement of private property rights may be the case of the 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 because, without the fur market, only weak incentives for overhunting were present. This changed after the Indians started trading fur with the white Americans; at this point, the incentives for overhunting and the costs of no property rights increased, so that private property rights in land were introduced.

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 goal of the Europeans, in one form or another, was the extraction of resources from these colonies. In the Caribbean, the extractive institutions took the form of slave plantations; in parts of Central and Meso America, they constituted mining based on forced labor. In Africa, Europeans were first interested in the acquisition of slaves to employ on the plantations in the Americas; later, they developed other methods of extracting resources, including high taxes and the extraction of natural resources.¹³ Other economic institutions set up by Europeans in these colonies were also detrimental to economic advancement; little investment was made 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 contrasts with the institutions that the Europeans set up in colonies in which they settled in large numbers, e.g., the United States, Canada, Australia and New Zealand. In these colonies, life was modeled after that in the home country; the emphasis was on the enforcement of property rights for a broad cross-section of society, especially smallholders, merchants and entrepreneurs. Gann and Duignan (1962), Robinson and Gallagher (1961), Denoon (1983), and Cain and Hopkins (1993) discuss these situations. Acemoglu et al. (2001) conclude that the crucial determinant of whether or not Europeans chose extractive institutions was whether or not they settled in large numbers in the colony. If Europeans settled in a colony, institutions were developed for their own future benefits. If Europeans did not settle in a colony, they set up a highly centralized state apparatus and other similar institutions to oppress the native population and to facilitate the extraction of resources in the short run. Hence, Acemoglu et al. (2001) suggest that, in places in which the disease environments made it easy for Europeans to settle, the path of institutional development should have been different from areas in which Europeans faced high mortality rates.

During the colonization period, Europeans faced widely different mortality rates in colonies because of differences in the prevalence of malaria and yellow fever.¹⁴ Differences in the mortality rates of potential settlers, driven mostly by malaria and yellow fever, are a plausible instrumental variable; while these mortality rates should not influence current output directly, they likely had a significant effect on institutional development by shaping the settlement patterns of Europeans. Mortality rates should not have a direct effect because Malaria and yellow fever were fatal to Europeans who had no immunity,

¹³ Davis and Huttenback (1986) calculate that, before 1885, investment in the British empire had a rate of return that was 25 percent higher than that on domestic investment. Roberts (1976) 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) notes that taxation rates in Tunisia were four times higher than those in metropolitan France. 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 1920s and 1930s.

¹⁴ In Acemoglu et al. (2001, Table A.2) presents 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, which was lower than in Europe at that time, to 49 in India, 130 in Jamaica, and around 500 in West Africa.

and thus had a major effect on settlement patterns, but they had much more limited effects on natives who had developed various immunities over the centuries.¹⁵ The data indicate major differences in the institutional development of the high-mortality and low-mortality colonies. Acemoglu et al. (2001) show that expropriation risk is much greater in the colonies where Europeans faced higher death rates and did not settle.

Acemoglu et al. (2001) also show that these institutional differences induced by mortality rates and European settlement patterns have a major impact on income per capita.¹⁶ Their estimates imply that improving Nigeria's institutions to the level of those in Chile could lead to as much as a 7-fold increase in Nigeria's income in the long run. This evidence gives a clear example of how societies do choose very different institutions, and shows that these choices are not due to differences in output-relevant variables, the *X*'s, but rather to other political or historical circumstances, the *Z*'s (in this case the mortality rates faced by early European settlers). These results suggest that the PCT, which emphasizes the forces that push societies towards choosing the appropriate institutions and policies, does not provide a sufficiently useful framework for analyzing the major institutional and policy differences across countries.

3.3. *North versus South Korea*

The contrast between North and South Korea is another example of 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. 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 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, so 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 was the takeover of all of Korea, either by the Soviet Union or by communist forces under the control of the former guerrilla fighter, Kim Il Sung. Therefore, US authorities supported the influential nationalist leader Syngman Rhee, who was in favor of separation rather than a united communist Korea. Elections 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 con-

¹⁵ This 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 for British troops in Bengal and Madras were between 70 and 170 in 1000.

¹⁶ That paper 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 latitudinal distance from the equator, continent dummies, temperature, humidity, whether countries are land-locked, soil quality, and natural resource abundance. They also obtain similar results using only yellow fever prevalence, which is an attractive source of variation because yellow fever is mostly eradicated.

stitution and established the Republic of Korea to the south of the 38th parallel. The North became the Democratic People's Republic of Korea, under the control of Kim Il Sung.¹⁷

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; initially, only minor economic differences were present between the two areas. 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 (Ha-Cheong, 1988). After separation, policies and institutions diverged substantially in the two countries. Under the leadership of Kim Il Sung, the North adopted a centralized command economy with little role for individual enterprise. Kim Il Sung advocated and imposed a philosophy he named *Juche*, which played an important role both in the political and economic life in North Korea. This philosophy emphasizes 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 made the majority of the key economic decisions directly. 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 many small household industries and producers; these were forced to join the cooperatives of the Consumer Union so that they could be supervised closely and instructed by the state. For all practical purposes, there were no private property rights for individuals (Eberstadt, 1999; Hunter, 1999).

In contrast, South Korea 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, they refrained from most predatory policies. In fact, General Park was supportive of economic development; his regime is often credited with facilitating, or even encouraging, investment and rapid growth in Korea (Evans, 1995; Wade, 1990). Even though many South Korean economic policies, such as protected domestic markets, entry barriers and subsidized loans, favored the 'chaebol' directly, there were no major violations of the property rights of 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 pure model of free-market capitalism would suggest.

Under these two highly contrasting regimes, the economies of North and South Korea diverged. According to Maddison (2001), both North and South Korea had approximately the same income level in 1950, \$770 at 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 was around \$2841, less than one-third of the income per capita in South Korea, which stood at \$8704. 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 that year, income in North Korea was

¹⁷ Cumings (1997) and Buzo (2002) provide recent histories of Korea.

inflated by Soviet aid. Since then, the North Korean economy has been shrinking further, while South Korea continues to grow. According to Maddison (2001), the difference is now over 10-fold with income per capita at \$12,152 in South Korea versus \$1183 in North Korea. 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 a homogeneous country into two parts with very different policies and institutions provides another clear example of how political leaders often choose very different policies, leading to very different outcomes, despite the very similar economic conditions.

3.4. *Conscious choices or belief differences?*

Can the differences in institutional development across the European colonies or the divergence in the institutions and policies between the North and South Korea be interpreted as resulting from differences in beliefs? It is entirely possible that Rhee, Park, and other South Korean leaders believed in the superiority of capitalist institutions and private property, while Kim Il Sung and Communist Party members in the North believed that communist policies would be better for the country. However, even if differences in beliefs could explain the divergence in institutions in the immediate aftermath of separation, by the 1980s, clearly the communist economic policies in the North were not working. The continued efforts of the leadership to cling to these policies so as to retain power can be explained only by those leaders seeking their own interests at the expense of the interests of the population. 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 alternatives, while realizing fully the costs that the system imposes on the North Korean people, including the famine that much of the population has been suffering for the past several years.

Differences in colonial policies argue even more strongly that social and distributional conflict is the most important cause of inefficient policies and persistently inefficient institutions. The same British colonists 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 and supported by highly oppressive institutions. In contrast, the institutions they developed in areas in which they settled and which had no large population of Indians or slaves to be oppressed, e.g., the Northeastern United States, Canada, Australia and New Zealand, were democratic, encouraged participation, imposed checks and balances on politicians and political elites, and enforced the property rights of a broad cross-section of society.¹⁸ Moreover, the incentives of the colonists are easy

¹⁸ The experience of the Puritans in the New World is an interesting example of how the same groups adopted very different colonization strategies and organizations of society in response to different incentives. The colony of Massachusetts Bay, formed in 1630 by the English Puritans, is often hailed as an example of good institutions introduced in the colonies by a group of British colonists 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 adopted immediately in this colony and the most profitable endeavor for those settling on the island was attacking and pirating Spanish ships in the area (Kupperman, 2000).

to understand; if they did not settle, they were choosing institutions simply to extract resources from the native population. If they settled in large numbers, institutions and policies were set in place to protect the settlers in the future and to encourage both investment and prosperity.

4. Commitment problems and the political Coase theorem

4.1. The description of the model

The previous section developed the argument that the PCT, in its simple or modified form, does not provide a good framework for analyzing cross-country differences in institutions and policies, based on empirical evidence. This section develops the economic rationale for socially and politically powerful groups to extract resources from the rest of the population in inefficient ways and to set up bad institutions. The basic idea is that the Coase theorem requires transfers from one party to another and these transfers cannot be made at the same time. Therefore, enforceable contracts specifying future transfers are necessary. However, if such transfers are between the citizens and the state, or groups controlling the state, a major commitment problem arises because no outside party is available to enforce such contracts and those controlling the state can renege on their promises. Only incentive compatible or self-enforcing arrangements are feasible so that inherent commitment problems make the PCT inapplicable. To simplify the exposition, the model focuses on the case with no belief differences, although the results apply to the modified PCT as well.

Consider an infinite horizon economy in which time is discrete and indexed by t . There are two groups of agents, a ruler, and identical citizens whose mass is equal to one. All agents discount the future by the discount factor β , so that the appropriate utility function is

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

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

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

where y_t^i is market output and R is income from natural resources.

Another inferior production technology, which has the advantage of being non-taxable, is available. In particular, this alternative produces non-market income

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

where $b < 1$. We designate the decision to produce for the market by $m_t \in \{0, 1\}$; if $m_t = 0$, the individual uses the non-market technology and his market income is $y_t^i = R$ so that income from natural resources remains taxable.¹⁹

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

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

The ruler has the power to tax the citizens, but he does not contribute to the production process. However, because he has full means of coercion, he can take as much of the output in the market sector as he wishes. For simplicity, this characterization ignores useful roles of rulers, such as law enforcement, public good provision, regulation, and defense, since these roles are not essential for the theory.²⁰

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

$$T_t(Y) \leq Y \equiv \int y_t^i di, \quad (1)$$

where Y denotes aggregate output. In the most general case, the ruler specifies person-specific taxes, T_t^j , for each individual j . Rulers are assumed to have the same discount factor, β , as the citizens. However, because of the possibility of an internal power struggle, they can lose power to another ruler with exogenous probability q .

The current ruler can also decide to relinquish his power, which is denoted by $r = 1$; in that case, no more rulers take over in the future. We refer to this case as democratization, even though it lacks many of the interesting features of transitions to democracy.²¹ In democracy, citizens can also impose taxes, denoted by S_t , on themselves in order to make transfers to previous rulers.

We make two alternative assumptions regarding feasible contracts. First, as a useful benchmark, contracts between citizens and current or previous ruler can be, at least partly, enforced. Second, more plausibly, contracts between citizens and rulers are not enforceable.

The timing of events within each period is as follows. First, if contracts are available, parties sign them. If there has been no democratization in the past, i.e., $r = 0$ in all past periods, the ruler decides whether or not to relinquish his power, $r = 1$ or $r = 0$. Next, individuals choose how much to invest, e , and whether or not to produce in the market sector, $m = 1$ or 0 . 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 democratization has occurred, a citizen decides how much tax to impose on each individual in order to transfer to the previous ruler, denoted

¹⁹ The presence of market income, even if individuals withdraw from market production, ensures that rulers continue to get a positive return.

²⁰ Acemoglu and Verdier (1998) present a model in which the government plays a useful role, but government officials are also corrupt and their actions distort private incentives.

²¹ Acemoglu and Robinson (2000a) present a model of democratization.

$S(Y)$. Then, consumption takes place. Finally, 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 then rulers set taxes. Hence, some type of contracts, either implicit or explicit, are necessary.²² The history of play in this repeated game, h^t , includes all the actions up to time t . 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 because, 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 . The analysis focuses on symmetric equilibria in which all citizens use the same strategy, and only one strategy mapping for the citizens needs to be specified. The investment and 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, we focus on stationary equilibria, in which the same strategies are played at all dates.²³

4.2. The no-cooperation benchmark

To provide a benchmark, let us begin with the no-cooperation case in which no democratization occurs, i.e., $r = 0$, and no contracts are written between rulers and citizens.

Proposition 1. *If $r = 0$, there exists an equilibrium in which the ruler expropriated all income, i.e., he sets $T(Y) = Y$, and the citizens use the informal sector technology only, i.e., $m = 0$ and $e = b$.*

This allocation is an equilibrium because 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 . If an individual deviates from his strategy and produces in the market system, his consumption will not increase because the ruler is grabbing everything. Hence, the citizens choose $m = 0$, i.e., production with the non-market technology only, and they invest the optimal amount for this technology, $e = b$.

²² This 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, the goal of the analysis is not to clarify the distinction between institutions and policies, but rather to identify the forces that prevent the efficient choice of policies and institutions.

²³ The restriction to stationary equilibria is an important one. Acemoglu (2003) shows that non-stationary equilibria outperform stationary equilibria in a similar environment, and Dixit et al. (2000) present a thorough analysis of non-stationary equilibria in a related political economy environment. The restriction to stationary equilibria is useful, however, since these are easier to characterize and yield useful comparative static results.

For future reference, denote the values received by the citizens and the ruler in this equilibrium as \widehat{W} and \widehat{V} , which are given by

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

and

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

This equilibrium is highly inefficient. For example, a Pareto-improving contract can be designed in which the ruler relinquishes power and the citizens promise him a side payment of $R + \varepsilon$ for every period thereafter. Then, the citizens would choose market production and investment so that $e = 1$, which achieves the first-best equilibrium. The remainder of the analysis focuses on whether such Pareto-improving contracts can be written.

4.3. The political Coase theorem with commitment

Suppose that enforceable contracts between rulers and citizens can be written. Even in the absence of full property rights for citizens, the PCT applies if the economy generates the efficient allocation and when the distribution of political power between the citizens and the ruler is irrelevant for the allocation.

Three different situations are possible: full commitment by the ruler, full commitment by the citizens, and limited commitment. This section discusses the first two cases, and leaves the third, which is the central focus of the paper, to the next section.

First, suppose that the ruler can commit to impose a tax level T so that $T(Y) = \min(T; Y)$, and assume that $q = 0$, so that the ruler is never replaced. After paying a tax level of T , each citizen keeps any amount of production above T . The following program gives the equilibrium allocation that satisfies the PCT and yields the largest surplus to the ruler (this equilibrium also features $r = 0$, i.e., the ruler does not transfer power to the citizens):

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

subject to the feasibility constraint (1) and to the following 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 if they invest e and are taxed T , while in the right-hand side, \widehat{W} is the value that citizens can obtain by opting out of the formal sector and is given by (2). The solution to this problem is straightforward; $T = \alpha(1 - b) + R$ for the ruler and $e = 1$ and $m = 1$ for all the citizens. Importantly, the efficient allocation is achieved despite the fact that political power is in the hands of the ruler, i.e., $r = 0$. By committing to the tax schedule $T(Y) = \min(\alpha(1 - b); Y) + R$, the ruler makes the citizens residual claimants after a threshold level of investment, encouraging them to undertake the first-best level of investment.

The above program is special in that it gives all the bargaining power to the ruler. Alternatively, some of the rents from achieving the PCT may be shared between the ruler and the citizens. Suppose that rents between the citizens and the ruler are shared according to the generalized Nash bargaining procedure. Following the above reasoning, citizens choose the efficient level of investment, $e = 1$. Imposing this level of investment, the Nash solution is found from the following program:

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

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, production for the non-market sector, which has a net present value of $\alpha b/(1-\beta)$. 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 is $T = (1-\theta)\alpha(1-b) + R$ and $r = 0$ for the ruler, while $e = 1$ and $m = 1$ are still optimal for all citizens. The surplus accruing from citizens producing in the market sector and undertaking the first-best level of investment is $\alpha(1-b)$. This surplus is shared between the citizens and the ruler. Income from natural resources, i.e., R , goes entirely to the ruler because the ruler can obtain this even if citizens do not cooperate. As the bargaining power of citizens, θ , goes to zero, T limits to $\alpha(1-b) + R$, as above. Again, the efficient allocation is achieved because of the commitment power of the ruler. Moreover, this model 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 .

The optimal solution is more involved if the ruler can be replaced by a new ruler, i.e., $q > 0$. In this case, the ruler has a preference for front-loaded payments, because he may not be around in the future. Hence, the ruler discounts 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 a current ruler because if this ruler is replaced, these payments are lost and additional payments have to be made to a new ruler. These two effects cancel each other so that the problem is still stationary. In particular, the allocation will be a solution to the maximization problem

$$\max_T \left[\frac{\alpha}{1-\beta} + \frac{R}{1-\beta} - \frac{T}{1-\beta} - \frac{\alpha b}{1-\beta} \right]^\theta \left[\frac{T}{1-\beta(1-q)} - \frac{R}{1-\beta(1-q)} \right]^{1-\theta}, \tag{6}$$

subject to (4). The only difference between (5) and (6) is that the discount factor of the ruler is different due to the possibility that he may be replaced 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, if the ruler relinquishes power, citizens can commit to a future path of transfers, $\{S_t\}$. Now the PCT can be achieved through democratization, i.e., $r = 1$; the ruler transfers power to the citizens in return for their commitment to a future path of transfers. Therefore, the equilibrium allocation with commitment on the side of citizens is 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, either by the ruler or by the citizens, the PCT applies and the distribution of rents between various parties 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. *If either the ruler or the citizens can commit to future transfers, the optimal solution always has $m = 1$ and $e = 1$, so that the PCT applies.*

4.4. Equilibria with limited commitment

Contracts between the ruler and the citizens involving commitment are useful as a benchmarks; however, they have little practical relevance because they are not enforceable in the real world. Contract enforcement requires a third party, typically the state, that possesses the monopoly of legitimate coercion in the society. This monopoly of coercion gives the state the power to force contractors to abide by the terms of the contract, even if making the specified payments or the necessary delivery of goods is not in their interests *ex post*. If the state is one of these contractors, this type of outside enforcement is not possible. Hence, it is very difficult for any party with real power to commit to a path of future transfers, taxes or actions. Therefore, we cannot rely on outside enforcement because abiding by the conditions of the contract must be incentive compatible for the state as well as for the citizens.

To develop this point, we analyze the above game without such contracts starting with the Markov perfect equilibria (MPE), which do not allow repeated-game punishment strategies. Using the above notation 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 also 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 depend only on payoff-relevant state variables which are not present in this simple game. Therefore, 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 because the future play of the game and the continuation payoffs do not depend on history, and consequently, grabbing everything has no future repercussions. Hence, $T(Y) = Y$. Given this tax strategy, citizens prefer $m = 0$ so that there is no market production, as in Proposition 1. This is a highly inefficient outcome that both the 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 $T \leq \alpha(1 - b) + R$, that 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 trigger punishment strategies. Therefore, the PCT 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 MPE has $m = 0$ and $T(Y) = Y$.*

5. Incentive-compatible promises

5.1. Incentive-compatible commitments by the ruler

Suppose that commitment can be supported because of 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 in which the ruler promises not to grab everything because of the future rents available from continued market production by the citizens. Such promises must be self-enforcing or incentive compatible for the ruler. These issues can be modeled as the non-Markovian equilibria of the game in which citizens play trigger strategies to induce the ruler not to grab all the output. Let us begin with the case in which there is no replacement of rulers, i.e., $q = 0$. Moreover, assume that the citizens can coordinate their actions and that they all choose the level of investment e that maximizes their utility as a group so that the game can be characterized as one between two players. Section 5.3 considers free riding, in which each individual may prefer to choose a different level of investment than the one that maximizes the utility of the citizens as a group.

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 the 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. The resulting allocation will yield a tax revenue of T in each period and provide the ruler with utility equal to

$$V = \frac{T}{1 - \beta}. \quad (7)$$

Since the ruler cannot commit to future taxes of the form $T(Y) = \min\langle T; Y \rangle$, we must ensure that playing this strategy is optimal for the ruler. The obvious deviation is to grab everything in the current period so that we must check to insure that not grabbing everything today is incentive compatible. If the ruler follows the repeated game equilibrium, he obtains V as given by (7). Alternatively, if he deviates to grab everything today and switches to the non-cooperative equilibrium in Proposition 1, he obtains all the output today, $e^{1-\alpha} + R$; however, from today onwards, he obtains the payoff given by \widehat{V} in (3). Thus the ruler's return to deviating from the specified strategy is $e^{1-\alpha} + R + \beta\widehat{V}$.

Incentive compatibility for the ruler requires:

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

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

$$T \geq \Upsilon(e) \equiv (1 - \beta)e^{1-\alpha} + R, \quad (8)$$

where the function $\Upsilon(e)$ represents the flow value of grabbing all current output for the ruler if current investment is e . Condition (8) states that the tax in each period must be large enough so that the ruler is not tempted to grab everything.

The incentive compatibility for citizens must also be satisfied. If they stay in the market sector, they obtain

$$W(e) = \frac{1}{1 - \beta} [e^{1-\alpha} - (1 - \alpha)e + R - T], \quad (9)$$

which must be greater than \widehat{W} given by (2) for an equilibrium investment of e . In other words, the tax must satisfy

$$T \leq T^{\max}(e) \equiv R + e^{1-\alpha} - (1-\alpha)e - \alpha b, \quad (10)$$

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

Given these two incentive compatibility constraints, we check to see whether the first-best solution can be supported, i.e., whether the allocation with $m = 1$ and $e = 1$ can be achieved so that the PCT generalizes 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 the citizens' incentive compatibility constraint is given by

$$T^{\max}(e = 1) = R + \alpha(1 - b). \quad (11)$$

Whether the first-best allocation can be supported depends on whether the ruler's incentive compatibility constraint, (8), is satisfied for this tax level. From (8) and (11) we have:

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

which is equivalent to the condition that

$$1 - \beta \leq \alpha(1 - b). \quad (12)$$

If condition (12) is satisfied, the PCT applies. 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 monopoly of political power in the hands of the ruler.

Condition (12) is more likely to be satisfied if agents are patient, i.e., β is high, and the outside options of the citizens are not too attractive so that the ruler can raise enough taxes in every period to avoid being tempted to grab more than the prescribed amount. Suppose that (12) is not satisfied so that the first-best investment level, $e = 1$, cannot be maintained. Market participation by the citizens, $m = 1$ and some positive investment in the market sector, $e > 0$, can be nonetheless supported as an equilibrium if the maximum tax citizens are willing to pay is greater than the flow return to the ruler from grabbing everything, that is:

$$T^{\max}(e) = R + e^{1-\alpha} - (1-\alpha)e - \alpha b \geq \Upsilon(e) = (1-\beta)e^{1-\alpha} + R. \quad (13)$$

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 1 represents the left- and right-hand sides of this inequality (13) in the space of $e^{1-\alpha}$ and T . For low values of e , $T^{\max}(e)$ increases faster than $\Upsilon(e)$; hence, 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 (in the figure, the coordinate of this point is $\beta^{(1-\alpha)/\alpha}$, since the horizontal axis is for $e^{1-\alpha}$). After this point, $\Upsilon(e)$ grows

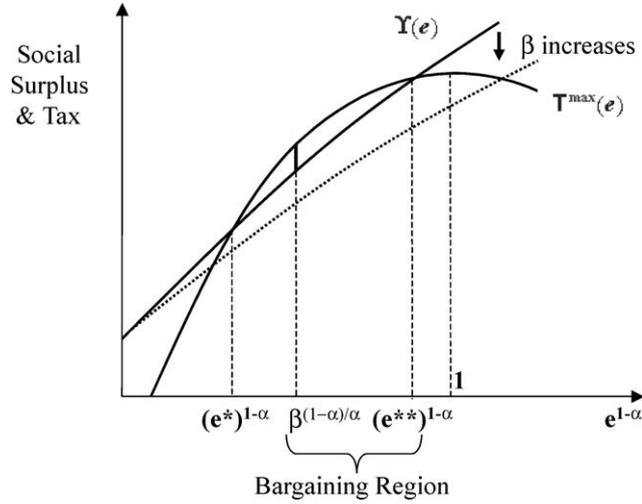


Fig. 1. Incentive compatible equilibria.

faster than $T^{\max}(e)$ because 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. Therefore, $T^{\max}(e = \beta^{1/\alpha}) - \Upsilon(e = \beta^{1/\alpha})$ is always greater than $T^{\max}(e = 1) - \Upsilon(e = 1)$ and it is easier to satisfy both incentive compatibility constraints at $e = \beta^{1/\alpha}$ than at the first-best level of investment. Therefore, at maximum effort, strong incentives exist for the ruler to grab everything today.

Figure 1 presents the case in which $T^{\max}(e = \beta^{1/\alpha}) - \Upsilon(e = \beta^{1/\alpha}) > 0 > T^{\max}(e = 1) - \Upsilon(e = 1)$. Although the first-best solution cannot be attained, 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 it will be explained 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. Let us therefore focus on the set $e \in [\beta^{1/\alpha}, e^{**}]$ as the set of potential equilibria. Clearly, this set changes with the underlying parameters. For example, if β increases, the set becomes larger; in particular, the highest investment that can be supported, e^{**} , increases. In fact, Fig. 1 presents the case in which the first-best level of investment can be supported with the increase in β .

From this analysis, a simple condition for market production to be supported arises. If both incentive compatibility constraints are not satisfied at $e = \beta^{1/\alpha}$, the set $[e^*, e^{**}]$ is empty. Therefore, the condition for $m = 1$ to be supported is:

$$\alpha\beta^{1/\alpha} \geq b. \tag{14}$$

If condition (14) is satisfied, multiple equilibria with different investment levels in the set $[e^*, e^{**}]$ are feasible. We have no selection criteria, and any of these investments, as well as others that are Pareto inferior to those on the frontier, may emerge in equilibrium. However, it is straightforward to identify which of these equilibria is the most preferred by the citizens and by the ruler, and also other points on the Pareto frontier of this society.

First, consider maximizing the ruler's utility, (7) subject to (8) and (10) by choosing e and T . The optimal solution is $e = e^{**}$ and $T^{\max}(e^{**})$. In other words, the ruler wants to maximize investment and choose the highest possible tax level given that investment. Note that e^{**} is also the level of investment that a social planner who wishes to maximize output would choose. In contrast, the citizens want to maximize (9) again subject to (8) and (10). As long as $e \in [e^*, e^{**}]$, the citizens' incentive compatibility constraint, (10), will be satisfied and the citizens will never give the ruler more than the minimum amount necessary to satisfy his incentive compatibility. Therefore, (8), must hold with equality. Then, substituting for T from (8) into (9), the citizens' maximization problem can be written as:

$$\max_e T^{\max}(e) - \mathcal{T}(e) = \max_e \beta e^{1-\alpha} - (1-\alpha)e.$$

In other words, the citizens want to maximize the difference between the left-hand side and the right-hand side of (13), which gives the solution that $e = \beta^{1/\alpha}$. Intuitively, increasing e further is costly to the citizens because they pay the additional investment costs, while the ruler obtains all the benefits. Since they do not internalize the ruler's gains, the citizens prefer $e = \beta^{1/\alpha}$ to the maximum supportable investment.

With direct analogy to the analysis above, we can also characterize the Pareto frontier by considering the generalized Nash bargaining solution between the ruler and the citizens subject to the incentive compatibility constraints. This is:

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

subject to the two incentive compatibility constraints, (8) and (10), and θ is again the bargaining power of the citizens. According to the PCT, the allocation of political power between the two groups should not affect the outcome. However, as long as (14) is satisfied and the set $[e^*, e^{**}]$ is non-empty, this will not be the case. We already saw that if $\theta = 0$, i.e., if the ruler has all the bargaining power, the outcome is $e = e^{**}$, and if $\theta = 1$ so that the citizens have all the power, the outcome is $e = \beta^{1/\alpha}$. The general solution to this maximization problem, $e(\theta)$, is decreasing in θ with $\hat{e}(\theta = 0) = e^{**}$ and $\hat{e}(\theta = 1) = \beta^{1/\alpha}$.²⁴ That greater bargaining power for citizens reduces investment and efficiency is a somewhat surprising result. Since the problem is the inability of the ruler to commit to not taxing the returns from citizens' investments, one might think that greater bargaining power for the citizens would reduce inefficiencies. However, bargaining power does not affect the incentive compatibility constraint of ruler; rather, it determines which point is chosen from among the possible set of equilibria. Since citizens bear the cost of investment and receive less than the full return, their preferred investment is always less than that of the ruler. With greater bargaining power, the citizens select an equilibrium closer to their desired point, which features lower investment but greater net returns for them.

Now 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 continued

²⁴ To demonstrate this result, note that the incentive compatibility constraint of the ruler, (8), must hold or both parties could 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: $\theta = 1 + (\beta \hat{e}^{1-\alpha} - \hat{e}) / ((\hat{e} - b)\alpha)$, which gives $d\hat{e}/d\theta < 0$ in the range of $\hat{e} \in [\hat{e}(\theta = 1) = \beta^{1/\alpha}, \hat{e}(\theta = 0) = e^{**}]$.

cooperation for the ruler. Taking this into account, the relevant comparison for the ruler is between grabbing everything today, 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. Hence, the ruler's incentive compatibility constraint becomes

$$T \geq \Upsilon(e) \equiv (1 - \beta(1 - q))e^{1-\alpha} + R, \quad (15)$$

while the citizens' incentive compatibility constraint remains unchanged. Condition (15) implies that the first-best outcome can now be supported if

$$1 - \beta(1 - q) \leq \alpha(1 - b), \quad (16)$$

which is more restrictive than (12) for all $q > 0$. Intuitively, the possibility of replacement reduces the value of future cooperation for the ruler and makes the first-best outcome more difficult to achieve. Hence, the PCT is less likely to apply.

The general solution also changes in the same direction. In terms of Fig. 1, the curve specifying $\Upsilon(e)$ shifts up and the range of investment levels that can be supported declines. The bargaining solution now corresponds to maximizing:

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

subject to (10) and (15). Obviously, the larger is q , i.e., the higher is the probability of replacement, the lower is the equilibrium level of investment.²⁵ The effect of the replacement probability in this situation is in contrast to the case of enforceable contracts, in which q does not matter. Finally, the corresponding condition for an equilibrium with $m = 1$ to be supported changes to

$$(\beta(1 - q))^{1/\alpha} > b. \quad (17)$$

These results are summarized in the following proposition.

Proposition 4. *If the ruler and the citizens cannot commit to future transfers, the PCT and the efficient allocation can be supported by trigger punishment strategies provided that (16) is satisfied. Otherwise, the level of investment is less than first best, $e^{fb} \equiv 1$. As long as condition (17) is satisfied, an equilibrium with market production, i.e., $m = 1$, but less than efficient investment, namely $e < 1$, 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 ruler, q .*

This analysis therefore establishes theoretical limits on the applicability of the PCT because of the inherent commitment problems in politics. Since there is no outside party with

²⁵ The mathematical argument is similar to that in the previous footnote. The relevant expression is $\theta = 1 + (\beta(1 - q)\hat{e}^{1-\alpha} - \hat{e})/((\hat{e} - b)\alpha)$, which gives $d\hat{e}/dq < 0$ in the range of $\hat{e} \in [\hat{e}(\theta = 1) = \beta^{1/\alpha}, \hat{e}(\theta = 0) = e^{**}]$.

the coercion capability to enforce contracts between the ruler and the citizens, promises of the ruler must be self-enforcing or incentive-compatible. This restriction puts limits on the society's capability to achieve efficient allocations and on the applicability of the PCT. The result holds, even though the model does not rule out lump-sum transfers between citizens and rulers.

5.2. *Determinants of policies and institutions*

The above analysis and comparative statics provide a simple interpretation of the potential determinants of equilibrium institutions and policies. First, if the PCT does not apply, the distribution of political power between the ruler and the citizens matters for the equilibrium outcome. More interesting, the decision-making horizon of the ruler matters. If the ruler is impatient, for example because he fears replacement by other competing rulers, self-enforcing agreements are harder to maintain because the incentives from the threat of future punishments are weaker. Therefore, better equilibrium policies will arise if the ruler has a longer horizon.²⁶ Finally, better outside options for citizens which leave only a small surplus to be shared between the ruler and the citizens make cooperation more difficult.

The identity of the agent who designs the game or writes the constitution may also be important, especially if institutions are imposed on a society by external groups, such as colonial powers. If the political system is constituted by the citizens, they will choose "democratic" institutions, in which the ruler does not have the power to tax them. In contrast, if some political elite or a ruler, who do not internalize the interests of the citizens, design the system, they will choose to keep all the power. Even though this observation may seem trivial, it is important for understanding why the European colonists introduced relatively democratic institutions with checks and balances on the state and the politicians' power in colonies where they settled in large numbers, i.e., where they became the citizens. In contrast, the same colonial powers established or maintained oppressive and extractive institutions in colonies where they did not settle but rather wished to transfer resources from the native population to themselves. Although our objective is not to construct a model that can be used to interpret a wide range of social situations, it is useful to incorporate checks and balances into the theoretical framework here by introducing a measure of institutional controls on politicians, namely costly replacement of rulers. This analysis provides comparative statics regarding checks and balances and establishes the interaction between these types of institutional constraints on rulers and the implicit constraints that the rulers place on themselves via self-enforcing agreements.

Let us now assume that the citizens can attempt to replace the politician, although such replacement is costly. Regarding the timing of events the replacement decision occurs after the ruler sets taxes. Let us also assume that the cost of replacement, denoted c , is incurred by all citizens irrespective of whether they support the replacement of the ruler, i.e., there

²⁶ Acemoglu and Robinson (2000b, 2002) make a different argument for why rulers who fear replacement may pursue the wrong policies for the society. In that paper, rulers who fear replacement are more likely to resist the introduction of superior technologies or institutions when these changes may erode their incumbency advantage and their potential future political power.

is no free-rider problem.²⁷ If citizens attempt to replace the ruler, we assume that they succeed with probability p . The parameter p can be interpreted as a measure of the quality of the checks and balances on politicians; if p is high, the citizens can control the ruler better. If the current ruler is ousted from power, a new ruler is put in place the following period. If the ruler is ousted, he does not receive the tax revenue from the current period and, to simplify the analysis, let us 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.

Suppose that the ruler has set the tax T and is expected to set the same tax in the future. Ignore the incentive compatibility constraint of the ruler initially, and suppose that the same equilibrium will be played over time irrespective of whether citizens have attempted to replace, or have replaced, the ruler. To simplify the analysis further, set $q = 0$. The citizens have a choice of whether to replace the ruler, at cost c , or continue with the implicit agreement. The value function of citizens is:

$$W(e, T) = \max \left\{ \begin{array}{l} e^{1-\alpha} - (1-\alpha)e + R - c - T + \beta W(e, T) \\ e^{1-\alpha} - (1-\alpha)e + R - T + \beta W(e, T) \end{array} \right\}. \quad (18)$$

The upper branch corresponds to the choice to replace and the lower branch applies when citizens do not attempt to replace the ruler. Notice that the continuation value with or without replacement is the same, $\beta W(e, T)$, because some ruler will follow the optimal policy after this point. The only difference between the two branches is the cost of replacement. This immediately implies that the citizens will never exercise their option to replace the ruler. Such replacement is costly, and, along the equilibrium path, it generates no benefits to the citizens.²⁸

Nevertheless, the ability of the citizens to replace the ruler may have an effect on the equilibrium because they can replace a ruler who deviates from the implicit agreement. In particular, consider a ruler who grabs all the output. Following this, the citizens and the ruler will play the no-cooperation game. If the citizens do not attempt to replace the ruler, their continuation value is

$$\tilde{W}(e) = -(1-\alpha)e + \beta \hat{W}, \quad (19)$$

where \hat{W} is the value of the citizens in the no-cooperation continuation game given by (2) and $-(1-\alpha)e$ is the flow return in the current period, because they have invested e and the entire output has been grabbed by the ruler. This expression also incorporates the fact that, if citizens do not replace the ruler now, they will not replace him at any subsequent date.

Now, let us determine the value to the citizens after they attempt to replace the ruler. Since it is sufficient to look at a one-step deviation, suppose that if the citizens' attempt to replace the ruler fails, they will never attempt to replace him again, and also suppose that in the continuation game, they will cooperate with a new ruler. Then the value of attempting

²⁷ This comment refers to free-riding on the decision of whether or not to oust the current ruler; it is different from free-riding on the investment decision, which is discussed in the next subsection.

²⁸ This conclusion holds a fortiori, if the ruler and the citizens revert to no-cooperation following an unsuccessful replacement attempt.

to replace the ruler is

$$\tilde{W}(e) = -(1 - \alpha)e - c + \beta(1 - p)\hat{W} + \beta pW(e, T), \quad (20)$$

where $W(e, T)$ is the equilibrium value. The comparison of (19) and (20) shows that as long as

$$c \leq \beta(\hat{W} - W(e, T)), \quad (21)$$

citizens will attempt to replace the ruler. By construction, $\hat{W} - W(e, T) > 0$, therefore condition (21) implies that, for a sufficiently low cost of replacement, i.e., for $c \rightarrow 0$, the citizens will attempt to replace rulers who deviate and grab all the output. This replacement option will clearly affect the incentive compatibility constraint of the ruler. In particular, if (21) holds, a ruler who deviates and grabs everything knows that he will be replaced with probability p . Since $\hat{W} - W(e, T) > 0$ in all future dates, the citizens will attempt to replace the ruler in every future period. Taking this into account, the ruler's incentive compatibility constraint changes from (8) 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

$$(1 - p)(1 - \beta) - \alpha(1 - b) \leq \frac{pR}{1 - (1 - p)\beta}, \quad (22)$$

which is equivalent to (12) if $p = 0$. Condition (22) is more likely to hold when p is high. Therefore, better checks and balances on rulers, measured by a higher value of p , make it more likely that the PCT will apply.

If (22) does not hold, the allocation most preferred by the ruler will be the one that maximizes the level of investment subject to the citizens' and the ruler's incentive compatibility constraints. In other words, the allocation most preferred by the ruler will pick the maximum e, \tilde{e} , that satisfies

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

or equivalently:

$$\begin{aligned} \alpha b - \frac{pR}{1 - (1 - p)\beta} &\equiv [1 - (1 - p)(1 - \beta)]\tilde{e}^{1-\alpha} - (1 - \alpha)\tilde{e} \\ &= \alpha b - \frac{pR}{1 - (1 - p)\beta}. \end{aligned} \quad (23)$$

Clearly, \tilde{e} is increasing in p . Therefore, the more effective are the citizens in their ability to replace the ruler, the greater is equilibrium investment. A similar argument to the one before establishes that market production, $m = 1$, can be supported in this case as long as

$$\left[1 - (1 - p)(1 - \beta) \right]^{1/\alpha} \geq b - \frac{pR}{\alpha[1 - (1 - p)\beta]}. \quad (24)$$

These results are summarized in the following proposition.

Proposition 5. *Suppose that in the game with replacement, the cost of replacement, c , is small, i.e., $c \rightarrow 0$. If (22) holds, the PCT applies and the efficient level of investment can be supported. Better checks and balances, measured by greater p , make it more likely that (22) holds. When (22) does not hold, the efficient level of investment cannot be achieved. However, as long as (24) holds, market production can be supported, and in this case, equilibrium investment is given by \tilde{e} that satisfies (23). Better checks and balances, i.e., a greater p , increase equilibrium investment also in this case.*

5.3. Free riding, overinvestment, and the form of taxation

The model also raises interesting issues related to free riding among the citizens and the form of taxation. Once the assumption that citizens coordinate their actions is relaxed, a motive for distortionary taxation arises.

Suppose that citizens do not coordinate their actions, and that the ruler continues to observe the aggregate income level, Y , and sets a lump-sum tax T that applies to each individual. Let us focus on the case in which (16) does not hold and (14) holds so that the first-best solution, i.e., $e = 1$, is not possible, but equilibria with market production exist, with $e < 1$. Each individual faces a lump-sum tax T and since individuals are atomistic, they do not take their own effect on Y into account. Hence, the equilibrium with $e < 1$ is no longer possible. Each individual would like to invest up to $e = 1$ because each is both infinitesimal and the residual claimant of the returns from the additional investment at the margin. Aggregating the behavior of all individuals yields output equal to $Y = 1 + R$, which violates the incentive compatibility constraint of the ruler and precludes the equilibrium with market production.

Can the self-enforcing equilibrium of the previous section be supported by some tax scheme other than lump-sum taxation? If tax schedules can be conditioned on individual income, i.e., y_j , the equilibrium tax schedule can be made sufficiently distortionary to induce exactly the right amount of investment. For example, suppose that the ruler imposes 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 desired equilibrium has $\hat{e} < 1$ and a tax level equal to \hat{T} . To support this equilibrium, the ruler must impose 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 must discourage investment enough so that individuals do not overinvest and violate the incentive compatibility constraint of the ruler. These points are summarized in the following proposition.

Proposition 6. *If the 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 $e = 1$.*

This model therefore not only helps in analyzing commitment problems and the limitations of the applicability of the PCT, it also suggests a rationale for observing apparently inefficient methods of taxation, even when non-distortionary lump-sum taxes are available. At face value, these tax schedules appear to distort incentives. Nevertheless,

in the realm of self-enforcing agreements between rulers and citizens, an important goal of the tax system may be to prevent citizens from overinvesting. Hence, citizens should not be full residual claimants to the returns from their investments. The model provides a preliminary explanation for why distortionary taxes may be preferred to non-distortionary alternatives. A more detailed investigation of the importance of this rationale in practice is left to future work.

6. Concluding remarks

The determinants of policies and the institutional choices that societies make have attracted much interest recently. Why do some societies choose high taxes, while others opt for lower taxation? Why are bureaucracies more corrupt in some countries than in others? Why are some societies democratic, some parliamentary, and some majoritarian? The first step in analyzing these issues is to decide who makes the policy and institutional choices and for whose interests. For example, do collective choices maximize the welfare of society as a whole or do they select policies and institutions that benefit certain politically powerful groups at the expense of other segments of society?

This paper provides a simple taxonomy to be used in searching for answers to this question. The first category is called the political Coase theorem, and denoted PCT; it argues that societies make efficient choices and distribute the gains from these choices between various groups and individuals. According to this approach, if societies choose inefficient policies, strong political and social forces will push them back towards efficient policies. Alternatively, societies may choose inefficient policies, not due to failures in the political process, but because the beliefs of politicians and citizens are mistaken. Finally, the third category involves 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 society as a whole, but for the benefit of those who control political power.

Much of the paper is devoted to arguing that theories of social conflict provide the appropriate empirical and theoretical framework for analysis. However, the reason why politically powerful groups choose policies that reduce aggregate output rather than choosing efficient policies and redistributing the gains to themselves requires explanation. The analysis highlights the commitment problems in politics as the explanation and therefore as the limit on the application of the PCT. In other words, efficiency considerations cannot be separated from distributional conflicts. The PCT presumes that political and economic trades can be made between various individuals and groups. However, these trades are intertemporal and rely on contracts and promises. Typically, contracts and explicit promises are enforced by the state. Hence, contracts that the state, or social groups controlling the state, wish to write with the rest of the society are non-enforceable. This implies that the allocation of political power creates an inherent commitment problem that undermines the potential to achieve efficient outcomes. This is not to deny that political and economic forces may sometimes push towards more efficient social arrangements. In fact, the theoretical model shows how incentive-compatible promises can make up for this lack of enforceable contracts. The main point is that the forces pushing towards efficiency in general and incentive-compatible promises in particular typically fall short of achieving effi-

cient outcomes (or validating the application of the PCT). We should therefore study the conflict of interest between different social groups to understand policy and institutional differences.

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. The evidence suggests that theories of social conflict provide the appropriate framework for further analysis. However, factors other than commitment problems may be important in limiting the applicability of the PCT. Furthermore, even if commitment problems are of primary importance, our modeling strategy may not be the most fruitful approach. Hence, this paper is an invitation for future research on the causes of inefficient policies and the factors preventing the application of the PCT.

Acknowledgments

I thank Raphael Auer Miriam Bruhn for outstanding research assistance and the seminar participants at the World Bank Appropriate Institutions conference and at the International Society for New Institutional Economics conference. I also thank Simeon Djankov, Simon Johnson, James Robinson and two anonymous referees for comments.

References

- Acemoglu, Daron, 2003. State power and economic development. In preparation.
- Acemoglu, Daron, Aghion, Philippe, Zilibott, Fabrizio, 2002. Distance to frontier, selection and economic growth. NBER working paper No. 9066. Cambridge, MA.
- Acemoglu, Daron, Johnson, Simon, Robinson, James A., 2001. The colonial origins of comparative development: an empirical investigation. *American Economic Review* 91 (5), 1369–1401.
- Acemoglu, Daron, Robinson, James A., 2000a. Political losers as a barrier to economic development. *American Economic Review* 90 (2), 126–130.
- Acemoglu, Daron, Robinson, James A., 2000b. Why did the West extend the franchise? Democracy, inequality and growth in historical perspective. *Quarterly Journal of Economics* 115 (4), 1167–1199.
- Acemoglu, Daron, Robinson, James A., 2001. Inefficient redistribution. *American Political Science Review* 95 (3), 649–661.
- Acemoglu, Daron, Robinson, James A., 2002. Economic backwardness in political perspective. NBER working paper No. 8831. Cambridge, MA.
- Acemoglu, Daron, Verdier, Thierry, 1998. Property rights, corruption the allocation of talent: a general equilibrium approach. *Economic Journal* 108 (450), 1381–1403.
- Aghion, Philippe, Bolton, Patrick, 1990. Government and the risk of default: a political-economic model of strategic role of debt. In: Dornbusch, Rudi, Draghi, Mario (Eds.), *Public Debt Management: Theory and History*. Cambridge Univ. Press, Cambridge, pp. 121–145.
- Aldrich, John H., 1983. A Downsian spatial model with party activism. *American Political Science Review* 77 (4), 974–990.
- Anderlini, Luca, Felli, Leonardo, 1998. Costly Coasean contracts. Theoretical economics discussion paper TE/98/362. LSE, London.
- Austen-Smith, David, Banks, Jeffrey S., 1999. *Positive Political Theory I: Collective Preference*. Univ. of Michigan Press, Ann Arbor, MI.
- Baron, David P., Ferejohn, John, 1989. Bargaining in legislatures. *American Political Science Review* 83 (4), 1181–1206.

- Becker, Gary S., 1983. A theory of competition among pressure groups for political influence. *Quarterly Journal of Economics* 98 (3), 371–400.
- Becker, Gary S., 1985. Public policies, pressure groups and dead weight costs. *Journal of Public Economics* 28 (3), 329–347.
- Becker, Gary S., Mulligan, Casey, 1998. Deadweight costs and the size of government. NBER working paper No. 6789. Cambridge, MA.
- Besley, Timothy, Coate, Stephen, 1998. Sources of inefficiency in a representative democracy: a dynamic analysis. *American Economic Review* 88 (1), 139–156.
- Buchanan, James M., Tullock, Gordon, 1962. *The Calculus of Consent*. Univ. of Michigan Press, Ann Arbor, MI.
- Buzo, Adrian, 2002. *The Making of Modern Korea*. Routledge, London.
- Cain, Philip J., Hopkins, Anthony G., 1993. *British Imperialism: Innovation and Expansion 1688–1914*. Longman, New York.
- Coase, Ronald, 1960. The problem of social cost. *Journal of Law and Economics* 3 (1), 1–44.
- Coate, Stephen, Morris, Stephen, 1995. On the form of transfers to special interests. *Journal of Political Economy* 103 (6), 1210–1235.
- Cumings, Bruce, 1997. *Korea's Place in the Sun: A Modern History*. Norton, New York. Footnote.
- Curtin, Philip D., 1964. *The Image of Africa*. Univ. of Wisconsin Press, Madison.
- Davis, Lance E., Huttenback, Robert A., 1986. *Mammon and the Pursuit of Empire: The Political Economy of British Imperialism, 1860–1912*. Cambridge Univ. Press, Cambridge.
- Demsetz, Harold, 1967. Toward a theory of property rights. *American Economic Review* 57 (2), 347–359.
- Denoon, Donald, 1983. *Settler Capitalism: The Dynamics of Dependent Development in the Southern Hemisphere*. Clarendon Press, Oxford.
- Dixit, Avinash, Grossman, Gene M., Gul, Faruk, 2000. The dynamics of political compromise. *Journal of Political Economy* 108 (3), 531–568.
- Dixit, Avinash, Londregan, John, 1995. Redistributive politics and economic efficiency. *American Political Science Review* 89 (4), 856–866.
- Dixit, Avinash, Olson, Mancur, 2000. Does voluntary participation undermine the Coase theorem? *Journal of Public Economics* 76 (3), 309–335.
- Djankov, Simeon, La Porta, Rafael, Lopez-de-Silanes, Florencio, Shleifer, Andrei, 2002. The regulation of entry. *Quarterly Journal of Economics* 117 (1), 1–37.
- Eberstadt, Nicholas, 1999. *The End of North Korea*. The AEI Press, Washington, DC.
- Engerman, Stanley L., Sokoloff, Kenneth L., 1997. Factor endowments, institutions, and differential paths of growth among new world economies. In: Haber, S.H. (Ed.), *How Latin America Fell Behind*. Stanford Univ. Press, Stanford, CA.
- Evans, Peter, 1995. *Embedded Autonomy: States and Industrial Transformation*. Princeton Univ. Press.
- Farrell, Joseph, 1987. Information and the Coase theorem. *Journal of Economic Perspectives* 1 (2), 113–129.
- Ferejohn, John, 1986. Incumbent performance and electoral control. *Public Choice* 50 (1–3), 5–25.
- Gann, Lewis H., Duignan, Peter, 1962. *White Settlers in Tropical Africa*. Penguin, Baltimore, MD.
- Glaesar, Edward, Shleifer, Andrei, 2002. Legal origins. *Quarterly Journal of Economics* 117 (4), 1193–1229.
- Grossman, Sanford, Hart, Oliver, 1986. The costs and benefits of ownership: a theory of vertical and lateral integration. *Journal of Political Economy* 94 (4), 691–719.
- Ha-Cheong, Yeon, 1988. *Economy*. In: *A Comparative Study of South North Korea*. National Unification Board, Seoul.
- Hall, Peter A., Soskice, David, 2001. *Varieties of capitalism: the institutional foundations of comparative advantage*. Oxford Univ. Press, New York.
- Hanushek, Eric A., Kimko, Dennis D., 2000. Schooling, labor-force quality and the growth of nations. *American Economic Review* 90 (5), 1184–1208.
- Hart, Oliver, Moore, John, 1990. Property rights and the nature of the firm. *Journal of Political Economy* 98 (6), 1119–1158.
- Herbst, Jeffery I., 2000. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton Univ. Press, Princeton, NJ.
- Hunter, Helen-Louise, 1999. *Kim Il-Song's North Korea*. Preager, Westport, CT.
- Knack, Steven, Keefer, Philip, 1995. Institutions and economic performance: cross-country tests using alternative measures. *Economics and Politics* 7 (3), 207–227.

- Krueger, Alan, Lindahl, Mikael, 2001. Education for growth: why and for whom? *Journal of Economic Literature* 39 (4), 1101–1136.
- Kupperman, Karen Ordahl, 2000. England's dream of Caribbean empire. In: Kupperman, Karen Ordahl (Ed.), *Major Problems in American Colonial History: Documents and Essays*. Houghton Mifflin, Boston, MA.
- Libecap, Gary D., 1989. *Contracting for property rights*. Cambridge Univ. Press, Cambridge.
- Maddison, Angus, 2001. *The World Economy: A Millennial Perspective*. Development Centre of the OECD, Paris.
- Manning, Patrick, 1982. *Slavery, Colonialism, and Economic Growth in Dahomey 1640–1980*. Cambridge Univ. Press, New York, NY.
- Mauro, Paulo, 1995. Corruption and growth. *Quarterly Journal of Economics* 110 (3), 681–712.
- McKelvey, Richard D., Page, Talbot, 1999. Taking the Coase theorem seriously. *Economics and Philosophy* 15 (2), 235–247.
- Meltzer, Allan H., Richards, Scott, 1981. A rational theory of the size of government. *Journal of Political Economy* 89 (5), 914–927.
- Mukand, Sharun, Rodrik, Dani, 2002. In search of holy grail: Policy convergence, experimentation and economic growth. Mimeo. Harvard Kennedy School, Cambridge, MA.
- Myerson, Roger B., 1995. Analysis of democratic institutions: structure, conduct and performance. *Journal of Economic Perspectives* 9 (1), 77–89.
- North, Douglass C., 1981. *Structure and Change in Economic History*. Norton, New York, NY.
- North, Douglass C., 1990. A transaction cost theory of politics. *Journal of Theoretical Politics* 2 (4), 355–367.
- North, Douglass C., Weingast, Barry R., 1989. Constitutions and commitment: The evolution of institutions governing public choice in seventeenth-century England. *Journal of Economic History* 49 (4), 803–832.
- Olson, Mancur, 2000. *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships*. Basic Books, New York, NY.
- Peemans, Jean-Philippe, 1975. Capital accumulation in the Congo under colonialism: the role of the state. In: Gann, Lewis H., Duignan, Peter (Eds.), *Colonialism in Africa 1870–1960*, Vol. 4. Hoover Institution Press, Stanford, CA.
- Persson, Torsten, Svensson, Lars, 1989. Why a stubborn conservative would run a deficit: policy with time-inconsistent preferences. *Quarterly Journal of Economics* 104 (2), 325–346.
- Persson, Torsten, Roland, Gerard, Tabellini, Guido, 1997. Separation of powers and political accountability. *Quarterly Journal of Economics* 112 (4), 1163–1202.
- Persson, Torsten, Tabellini, Guido, 2000. *Political Economics: Explaining Economic Policy*. MIT Press, Cambridge, MA.
- Piketty, Thomas, 1995. Social mobility and redistributive politics. *Quarterly Journal of Economics* 110 (3), 551–584.
- Rajan, Raghuram G., Zingales, Luigi, 2000. The tyranny of inequality. *Journal of Public Economics* 76 (3), 521–558.
- Ringer, Fritz K., 1979. *Education and Society in Modern Europe*. Indiana Univ. Press, Bloomington, IN.
- Roberts, Andrew, 1976. *A History of Zambia*. Heinemann, London.
- Robinson, Ronald E., Gallagher, John, 1961. *Africa and the Victorians: The Official Mind of Imperialism*. MacMillan, London.
- Robinson, James A., 1998. Theories of 'bad' policy. *Journal of Policy Reform* 2 (1), 1–46.
- Rodrik, Dani, 1986. Tariffs, subsidies, and welfare with endogenous policy. *Journal of International Economics* 21 (34), 285–299.
- Romer, Thomas, 1975. Individual welfare majority voting, and the properties of a linear income tax. *Journal of Public Economics* 4 (2), 163–185.
- Romer, David, 1997. Misconceptions and political outcomes. NBER working paper No. 6117. Cambridge, MA.
- Shleifer, Andrei, Vishny, Robert W., 1998. *The Grabbing Hand: Government Pathologies and their Cures*. Harvard Univ. Press, Cambridge, MA.
- Snyder, James M., 1990. Resource allocation in multiparty elections. *American Journal of Political Science* 34 (1), 59–73.
- Spiller, Pablo, Tommasi, Mariano, 2002. The institutional foundations of public policy: A transactions approach with application to Argentina. Mimeo. Univ. of California, Berkeley, CA.
- Stigler, George, 1966. *The Theory of Price*. Macmillan, New York, NY.

- Tilly, Charles, 1990. *Coercion Capital and European States, AD 990–1990*. Blackwell, Cambridge, MA.
- Wade, Robert, 1990. *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization*. Princeton Univ. Press, Princeton, NJ.
- Weingast, Barry R., 1998. Political stability and civil war: institutions, commitment and American democracy. In: Bates, Robert H., Greif, Avner, Levi, Margaret, Rosenthal, Jean-Laurent, Weingast, Barry R. (Eds.), *Analytic Narratives*. Princeton Univ. Press, Princeton, NJ.
- Williamson, Oliver, 1985. *The Economic Institutions of Capitalism*. Free Press, New York, NY.
- Wilson, James D., 1990. Are efficiency improvements in government transfer policies self-defeating in political equilibrium? *Economics and Politics* 2 (3), 241–258.
- Wittman, Donald, 1995. *The Myth of Democratic Failure: Why Political Institutions are Efficient*. Univ. of Chicago Press, Chicago, IL.
- Young, Crawford, 1994. *The African Colonial State in Comparative Perspective*. Yale Univ. Press, New Haven, CT.