Chapter 6

INSTITUTIONS AS A FUNDAMENTAL CAUSE OF LONG-RUN GROWTH

DARON ACEMOGLU

Department of Economics, MIT, 50 Memorial Drive E52-380b, Cambridge, MA 02142
e-mail: daron@mit.edu

SIMON JOHNSON

Sloan School of Management, MIT, 50 Memorial Drive, Cambridge, MA 02142
e-mail: sjohnson@mit.edu

JAMES A. ROBINSON

Department of Government, WCFIA, Harvard University, 1033 Massachusetts Avenue, Cambridge, MA 02138
e-mail: jrobinson@gov.harvard.edu

Contents

Handbook of Economic Growth, Volume 1A. Edited by Philippe Aghion and Steven N. Durlauf

DOI: 10.1016/S1574-0684(05)01006-3

Abstract

This paper develops the empirical and theoretical case that differences in economic institutions are the fundamental cause of differences in economic development. We first document the empirical importance of institutions by focusing on two "quasi-natural experiments" in history, the division of Korea into two parts with very different economic institutions and the colonization of much of the world by European powers starting in the fifteenth century. We then develop the basic outline of a framework for thinking about why economic institutions differ across countries. Economic institutions determine the incentives of and the constraints on economic actors, and shape economic outcomes. As such, they are social decisions, chosen for their consequences. Because different groups and individuals typically benefit from different economic institutions, there is generally a conflict over these social choices, ultimately resolved in favor of groups with greater political power. The distribution of political power in society is in turn determined by political institutions and the distribution of resources. Political institutions allocate *de jure* political power, while groups with greater economic might typically possess greater *de facto* political power. We therefore view the appropriate theoretical framework as a dynamic one with political institutions and the distribution of resources as the state variables. These variables themselves change over time because prevailing economic institutions affect the distribution of resources, and because groups with de facto political power today strive to change political institutions in order to increase their de jure political power in the future. Economic institutions encouraging economic growth emerge when political institutions allocate power to groups with interests in broad-based property rights enforcement, when they create effective constraints on power-holders, and when there are relatively few rents to be captured by power-holders. We illustrate the assumptions, the workings and the implications of this framework using a number of historical examples.

Keywords

institutions, growth, development, political power, rents, conflict, property rights, efficiency, distributions

JEL classification: D7, H1, O10, O40

1. Introduction

1.1. The question

The most trite yet crucial question in the field of economic growth and development is: Why are some countries much poorer than others? Traditional neoclassical growth models, following Solow (1956), Cass (1965) and Koopmans (1965), explain differences in income per capita in terms of different paths of factor accumulation. In these models, cross-country differences in factor accumulation are due either to differences in saving rates (Solow), preferences (Cass–Koopmans), or other exogenous parameters, such as total factor productivity growth. In these models there are institutions, for example agents have well defined property rights and exchange goods and services in markets, but differences in income and growth are not explained by variation in institutions.

The first wave of the more recent incarnations of growth theory, following Romer (1986) and Lucas (1988) differed in the sense that they emphasized that externalities from physical and human capital accumulation could induce sustained steady-state growth. However, they also stayed squarely within the neoclassical tradition of explaining differences in growth rates in terms of preferences and endowments. The second wave of models, particularly Romer (1990), Grossman and Helpman (1991) and Aghion and Howitt (1992), endogenized steady-state growth and technical progress, but their explanation for income differences is similar to that of the older theories. For instance, in the model of Romer (1990), a country may be more prosperous than another if it allocates more resources to innovation, but what determines this is essentially preferences and properties of the technology for creating 'ideas'.[1]

Though this theoretical tradition is still vibrant in economics and has provided many insights about the mechanics of economic growth, it has for a long time seemed unable to provide a *fundamental* explanation for economic growth. As North and Thomas (1973, p. 2) put it: "the factors we have listed (innovation, economies of scale, education, capital accumulation, etc.) are not causes of growth; they *are* growth" (italics in original). Factor accumulation and innovation are only *proximate* causes of growth. In North and Thomas's view, the fundamental explanation of comparative growth is differences in *institutions*.

What are institutions exactly? North (1990, p. 3) offers the following definition: "Institutions are the rules of the game in a society or, more formally, are the humanly devised constraints that shape human interaction". He goes on to emphasize the key implications of institutions since, "In consequence they structure incentives in human exchange, whether political, social, or economic".

[1] Although, as we discuss later, some recent contributions to growth theory emphasize the importance of economic policies, such as taxes, subsidies to research, barriers to technology adoption and human capital policy, they typically do not present an explanation for why there are differences in these policies across countries.

Of primary importance to economic outcomes are the *economic institutions* in society such as the structure of property rights and the presence and perfection of markets. Economic institutions are important because they influence the structure of economic incentives in society. Without property rights, individuals will not have the incentive to invest in physical or human capital or adopt more efficient technologies. Economic institutions are also important because they help to allocate resources to their most efficient uses, they determine who gets profits, revenues and residual rights of control. When markets are missing or ignored (as they were in the Soviet Union, for example), gains from trade go unexploited and resources are misallocated. Societies with economic institutions that facilitate and encourage factor accumulation, innovation and the efficient allocation of resources will prosper.

Central to this chapter and to much of political economy research on institutions is that economic institutions, and institutions more broadly, are *endogenous*; they are, at least in part, determined by society, or a segment of it. Consequently, the question of why some societies are much poorer than others is closely related to the question of why some societies have much "worse economic institutions" than others.

Even though many scholars including John Locke, Adam Smith, John Stuart Mill, Arthur Lewis, Douglass North and Robert Thomas, and recently many papers in the literature on economic growth and development, have emphasized the importance of economic institutions, we are far from a useful framework for thinking about how economic institutions are determined and why they vary across countries. In other words, while we have good reason to believe that economic institutions matter for economic growth, we lack the crucial *comparative static* results which will allow us to explain why equilibrium economic institutions differ (and perhaps this is part of the reason why much of the economics literature has focused on the proximate causes of economic growth, largely neglecting fundamental institutional causes).

This chapter has three aims. First, we selectively review the evidence that differences in economic institutions are a fundamental cause of cross-country differences in prosperity. Second, we outline a framework for thinking about why economic institutions vary across countries. We emphasize the potential comparative static results of this framework and also illustrate the key mechanisms through a series of historical examples and case studies. Finally, we highlight a large number of areas where we believe future theoretical and empirical work would be very fruitful.

1.2. The argument

The basic argument of this chapter can be summarized as follows:

1. Economic institutions matter for economic growth because they shape the incentives of key economic actors in society, in particular, they influence investments in physical and human capital and technology, and the organization of production. Although cultural and geographical factors may also matter for economic performance, differences in economic institutions are the major source of cross-country differences in economic growth and prosperity. Economic institutions not only determine the ag-

gregate economic growth potential of the economy, but also an array of economic outcomes, including the distribution of resources in the future (i.e., the distribution of wealth, of physical capital or human capital). In other words, they influence not only the size of the aggregate pie, but how this pie is divided among different groups and individuals in society. We summarize these ideas schematically as (where the subscript t refers to current period and $t+1$ to the future):

$$\text{economic institutions}_t \Longrightarrow \begin{cases} \text{economic performance}_t \\ \text{distribution of resources}_{t+1} \end{cases}.$$

2. Economic institutions are endogenous. They are determined as collective choices of the society, in large part for their economic consequences. However, there is no guarantee that all individuals and groups will prefer the same set of economic institutions because, as noted above, different economic institutions lead to different distributions of resources. Consequently, there will typically be a *conflict of interest* among various groups and individuals over the choice of economic institutions. So how are equilibrium economic institutions determined? If there are, for example, two groups with opposing preferences over the set of economic institutions, which group's preferences will prevail? The answer depends on the *political power* of the two groups. Although the efficiency of one set of economic institutions compared with another may play a role in this choice, political power will be the ultimate arbiter. Whichever group has more political power is likely to secure the set of economic institutions that it prefers. This leads to the second building block of our framework:

$$\text{political power}_t \Longrightarrow \text{economic institutions}_t.$$

3. Implicit in the notion that political power determines economic institutions is the idea that there are conflicting interests over the distribution of resources and therefore indirectly over the set of economic institutions. But why do the groups with conflicting interests not agree on the set of economic institutions that maximize aggregate growth (the size of the aggregate pie) and then use their political power simply to determine the distribution of the gains? Why does the exercise of political power lead to economic inefficiencies and even poverty? We will explain that this is because there are commitment problems inherent in the use of political power. Individuals who have political power cannot commit not to use it in their best interests, and this commitment problem creates an inseparability between efficiency and distribution because credible compensating transfers and side-payments cannot be made to offset the distributional consequences of any particular set of economic institutions.

4. The distribution of political power in society is also endogenous, however. In our framework, it is useful to distinguish between two components of political power, which we refer to as *de jure (institutional)* and *de facto political power*. Here de jure political power refers to power that originates from the *political institutions* in society. Political institutions, similarly to economic institutions, determine the constraints on and the incentives of the key actors, but this time in the political sphere. Examples of political institutions include the form of government, for example, democracy vs. dictatorship or

autocracy, and the extent of constraints on politicians and political elites. For example, in a monarchy, political institutions allocate all de jure political power to the monarch, and place few constraints on its exercise. A constitutional monarchy, in contrast, corresponds to a set of political institutions that reallocates some of the political power of the monarch to a parliament, thus effectively constraining the political power of the monarch. This discussion therefore implies that:

$$\text{political institutionst}_t \Longrightarrow \text{de jure political power}_t.$$

5. There is more to political power than political institutions, however. A group of individuals, even if they are not allocated power by political institutions, for example as specified in the constitution, may nonetheless possess political power. Namely, they can revolt, use arms, hire mercenaries, co-opt the military, or use economically costly but largely peaceful protests in order to impose their wishes on society. We refer to this type of political power as de facto political power, which itself has two sources. First, it depends on the ability of the group in question to solve its collective action problem, i.e., to ensure that people act together, even when any individual may have an incentive to free ride. For example, peasants in the Middle Ages, who were given no political power by the constitution, could sometimes solve the collective action problem and undertake a revolt against the authorities. Second, the de facto power of a group depends on its economic resources, which determine both their ability to use (or misuse) existing political institutions and also their option to hire and use force against different groups. Since we do not yet have a satisfactory theory of when groups are able to solve their collective action problems, our focus will be on the second source of de facto political power, hence:

$$\text{distribution of resources}_t \Longrightarrow \text{de facto political power}_t.$$

6. This brings us to the evolution of one of the two main *state variables* in our framework, political institutions (the other state variable is the distribution of resources, including distribution of physical and human capital stocks, etc.). Political institutions and the distribution of resources are the state variables in this dynamic system because they typically change relatively slowly, and more importantly, they determine economic institutions and economic performance both directly and indirectly. Their direct effect is straightforward to understand. If political institutions place all political power in the hands of a single individual or a small group, economic institutions that provide protection of property rights and equal opportunity for the rest of the population are difficult to sustain. The indirect effect works through the channels discussed above: political institutions determine the distribution of de jure political power, which in turn affects the choice of economic institutions. This framework therefore introduces a natural concept of *a hierarchy of institutions*, with political institutions influencing equilibrium economic institutions, which then determine economic outcomes.

Political institutions, though slow changing, are also endogenous. Societies transition from dictatorship to democracy, and change their constitutions to modify the constraints on power holders. Since, like economic institutions, political institutions are collective choices, the distribution of political power in society is the key determinant of their evolution. This creates a tendency for persistence: political institutions allocate de jure political power, and those who hold political power influence the evolution of political institutions, and they will generally opt to maintain the political institutions that give them political power. However, de facto political power occasionally creates changes in political institutions. While these changes are sometimes discontinuous, for example when an imbalance of power leads to a revolution or the threat of revolution leads to major reforms in political institutions, often they simply influence the way existing political institutions function, for example, whether the rules laid down in a particular constitution are respected as in most functioning democracies, or ignored as in current-day Zimbabwe. Summarizing this discussion, we have:

$$\text{political power}_t \Longrightarrow \text{political institutions}_{t+1}.$$

Putting all these pieces together, a schematic (and simplistic) representation of our framework is as follows:

$$\begin{array}{ccccccc}
\begin{array}{c}\text{political}\\ \text{institutions}_t\end{array} & \Longrightarrow & \left.\begin{array}{c}\text{de jure}\\ \text{political}\\ \text{power}_t\\ \&\\ \text{de facto}\\ \text{political}\\ \text{power}_t\end{array}\right\} & \begin{array}{c}\Longrightarrow\\ \\ \\ \\ \Longrightarrow\end{array} & \begin{array}{c}\text{economic}\\ \text{institutions}_t\\ \\ \\ \text{political}\\ \text{institutions}_{t+1}\end{array} & \Longrightarrow & \left\{\begin{array}{c}\text{economic}\\ \text{performance}_t\\ \&\\ \text{distribution}\\ \text{of resources}_{t+1}\end{array}\right. \\
\begin{array}{c}\text{distribution}\\ \text{of resources}_t\end{array} & \Longrightarrow & & & & &
\end{array}$$

The two state variables are political institutions and the distribution of resources, and the knowledge of these two variables at time t is sufficient to determine all the other variables in the system. While political institutions determine the distribution of de jure political power in society, the distribution of resources influences the distribution of de facto political power at time t. These two sources of political power, in turn, affect the choice of economic institutions and influence the future evolution of political institutions. Economic institutions determine economic outcomes, including the aggregate growth rate of the economy and the distribution of resources at time $t+1$. Although economic institutions are the essential factor shaping economic outcomes, they are themselves endogenous and determined by political institutions and distribution of resources in society.

There are two sources of persistence in the behavior of the system: first, political institutions are durable, and typically, a sufficiently large change in the distribution of political power is necessary to cause a change in political institutions, such as a transition from dictatorship to democracy. Second, when a particular group is rich relative to others, this will increase its de facto political power and enable it to push for economic and political institutions favorable to its interests. This will tend to reproduce the initial relative wealth disparity in the future. Despite these tendencies for persistence, the framework also emphasizes the potential for change. In particular, "shocks", including changes in technologies and the international environment, that modify the balance of (de facto) political power in society and can lead to major changes in political institutions and therefore in economic institutions and economic growth.

A brief example might be useful to clarify these notions before commenting on some of the underlying assumptions and discussing comparative statics. Consider the development of property rights in Europe during the Middle Ages. There is no doubt that lack of property rights for landowners, merchants and proto-industrialists was detrimental to economic growth during this epoch. Since political institutions at the time placed political power in the hands of kings and various types of hereditary monarchies, such rights were largely decided by these monarchs. Unfortunately for economic growth, while monarchs had every incentive to protect their own property rights, they did not generally enforce the property rights of others. On the contrary, monarchs often used their powers to expropriate producers, impose arbitrary taxation, renege on their debts, and allocate the productive resources of society to their allies in return for economic benefits or political support. Consequently, economic institutions during the Middle Ages provided little incentive to invest in land, physical or human capital, or technology, and failed to foster economic growth. These economic institutions also ensured that the monarchs controlled a large fraction of the economic resources in society, solidifying their political power and ensuring the continuation of the political regime.

The seventeenth century, however, witnessed major changes in the economic and political institutions that paved the way for the development of property rights and limits on monarchs' power, especially in England after the Civil War of 1642 and the Glorious Revolution of 1688, and in the Netherlands after the Dutch Revolt against the Hapsburgs. How did these major institutional changes take place? In England, for example, until the sixteenth century the king also possessed a substantial amount of de facto political power, and leaving aside civil wars related to royal succession, no other social group could amass sufficient de facto political power to challenge the king. But changes in the English land market [Tawney (1941)] and the expansion of Atlantic trade in the sixteenth and seventeenth centuries [Acemoglu, Johnson and Robinson (2005)] gradually increased the economic fortunes, and consequently the de facto power of landowners and merchants. These groups were diverse, but contained important elements that perceived themselves as having interests in conflict with those of the king: while the English kings were interested in predating against society to increase their tax incomes, the gentry and merchants were interested in strengthening their property rights.

By the seventeenth century, the growing prosperity of the merchants and the gentry, based both on internal and overseas, especially Atlantic, trade, enabled them to field military forces capable of defeating the king. This de facto power overcame the Stuart monarchs in the Civil War and Glorious Revolution, and led to a change in political institutions that stripped the king of much of his previous power over policy. These changes in the distribution of political power led to major changes in economic institutions, strengthening the property rights of both land and capital owners and spurred a process of financial and commercial expansion. The consequence was rapid economic growth, culminating in the Industrial Revolution, and a very different distribution of economic resources from that in the Middle Ages.

It is worth returning at this point to two critical assumptions in our framework. First, why do the groups with conflicting interests not agree on the set of economic institutions that maximize aggregate growth? So in the case of the conflict between the monarchy and the merchants, why does the monarchy not set up secure property rights to encourage economic growth and tax some of the benefits? Second, why do groups with political power want to change political institutions in their favor? For instance, in the context of the example above, why did the gentry and merchants use their de facto political power to change political institutions rather than simply implement the policies they wanted? The answers to both questions revolve around issues of *commitment* and go to the heart of our framework.

The distribution of resources in society is an inherently conflictual, and therefore political, decision. As mentioned above, this leads to major commitment problems, since groups with political power cannot commit to not using their power to change the distribution of resources in their favor. For example, economic institutions that increased the security of property rights for land and capital owners during the Middle Ages would not have been credible as long as the monarch monopolized political power. He could promise to respect property rights, but then at some point, renege on his promise, as exemplified by the numerous financial defaults by medieval kings [e.g., Veitch (1986)]. Credible secure property rights necessitated a reduction in the political power of the monarch. Although these more secure property rights would foster economic growth, they were not appealing to the monarchs who would lose their rents from predation and expropriation as well as various other privileges associated with their monopoly of political power. This is why the institutional changes in England as a result of the Glorious Revolution were not simply conceded by the Stuart kings. James II had to be deposed for the changes to take place.

The reason why political power is often used to change political institutions is related. In a dynamic world, individuals care not only about economic outcomes today but also in the future. In the example above, the gentry and merchants were interested in their profits and therefore in the security of their property rights, not only in the present but also in the future. Therefore, they would have liked to use their (de facto) political power to secure benefits in the future as well as the present. However, commitment to future allocations (or economic institutions) was not possible because decisions in the future would be decided by those who had political power in the future with little reference to past promises. If the gentry and merchants would have been sure to maintain their de facto political power, this would not have been a problem. However, de facto political power is often transient, for example because the collective action problems that are solved to amass this power are likely to resurface in the future, or other groups, especially those controlling de jure power, can become stronger in the future. Therefore, any change in policies and economic institutions that relies purely on de facto political power is likely to be reversed in the future. In addition, many revolutions are followed by conflict within the revolutionaries. Recognizing this, the English gentry and merchants strove not just to change economic institutions in their favor following their victories against the Stuart monarchy, but also to alter political institutions and the future allocation of de jure power. Using political power to change political institutions then emerges as a useful strategy to make gains more durable. The framework that we propose, therefore, emphasizes the importance of political institutions, and changes in political institutions, as a way of manipulating future political power, and thus indirectly shaping future, as well as present, economic institutions and outcomes.

This framework, though abstract and highly simple, enables us to provide some preliminary answers to our main question: why do some societies choose "good economic institutions"? At this point, we need to be more specific about what good economic institutions are. A danger we would like to avoid is that we define good economic institutions as those that generate economic growth, potentially leading to a tautology. This danger arises because a given set of economic institutions may be relatively good during some periods and bad during others. For example, a set of economic institutions that protects the property rights of a small elite might not be inimical to economic growth when all major investment opportunities are in the hands of this elite, but could be very harmful when investments and participation by other groups are important for economic growth [see Acemoglu (2003b)]. To avoid such a tautology and to simplify and focus the discussion, throughout we think of good economic institutions as those that provide security of property rights and relatively equal access to economic resources to a broad cross-section of society. Although this definition is far from requiring equality of opportunity in society, it implies that societies where only a very small fraction of the population have well-enforced property rights do not have good economic institutions. Consequently, as we will see in some of the historical cases discussed below, a given set of economic institutions may have very different implications for economic growth depending on the technological possibilities and opportunities.

Given this definition of good economic institutions as providing secure property rights for a broad cross-section of society, our framework leads to a number of important comparative statics, and thus to an answer to our basic question. First, political institutions that place checks on those who hold political power, for example, by creating a balance of power in society, are useful for the emergence of good economic institutions. This result is intuitive; without checks on political power, power holders are more likely to opt for a set of economic institutions that are beneficial for themselves and detrimental for the rest of society, which will typically fail to protect property rights of a broad cross-section of people. Second, good economic institutions are more likely to arise when political power is in the hands of a relatively broad group with significant investment opportunities. The reason for this result is that, everything else equal, in this case power holders will themselves benefit from secure property rights.[2] Third, good economic institutions are more likely to arise and persist when there are only limited rents that power holders can extract from the rest of society, since such rents would

[2] The reason why we inserted the caveat of "a relatively broad group" is that when a small group with significant investment opportunities holds power, they may sometimes opt for an oligarchic system where their own property rights are protected, but those of others are not [see Acemoglu (2003b)].

encourage them to opt for a set of economic institutions that make the expropriation of others possible. These comparative statics therefore place political institutions at the center of the story, as emphasized by our term "hierarchy of institutions" above. Political institutions are essential both because they determine the constraints on the use of (de facto and de jure) political power and also which groups hold de jure political power in society. We will see below how these comparative statics help us understand institutional differences across countries and over time in a number of important historical examples.

1.3. Outline

In the next section we discuss how economic institutions constitute the basis for a fundamental theory of growth, and we contrast this with other potential fundamental theories. In Section 3 we consider some empirical evidence that suggests a key role for economic institutions in determining long-run growth. We also emphasize some of the key problems involved in establishing a causal relationship between economic institutions and growth. We then show in Section 4 how the experience of European colonialism can be used as a 'natural experiment' which can address these problems. Having established the central causal role of economic institutions and their importance relative to other factors in cross-country differences in economic performance, the rest of the paper focuses on developing a theory of economic institutions. Section 5 discusses four types of explanation for why countries have different institutions, and argues that the most plausible is the *social conflict view*. According to this theory, bad institutions arise because the groups with political power benefit from bad institutions. The emphasis on social conflict arises naturally from our observation above that economic institutions influence the distribution of resources as well as efficiency. Different groups or individuals will therefore prefer different institutions and conflict will arise as each tries to get their own way. Section 6 delves deeper into questions of efficiency and asks why a political version of the Coase Theorem does not hold. We emphasize the idea that commitment problems are intrinsic to the exercise of political power. In Section 7 we argue that a series of historical examples of diverging economic institutions are best explained by the social conflict view. These examples illustrate how economic institutions are determined by the distribution of political power, and how this distribution is influenced by political institutions. Section 8 puts these ideas together to build our theory of institutions. In Section 9 we then consider two more extended examples of the theory in action, the rise of constitutional rule in early modern Europe, and the creation of mass democracy, particularly in Britain, in the nineteenth and twentieth centuries. Section 10 concludes with a discussion of where this research program can go next.

2. Fundamental causes of income differences

We begin by taking a step back. The presumption in the introduction was that economic institutions matter, and should in fact be thought of as one of the key fundamental causes of economic growth and cross-country differences in economic performance. How do we know this?

2.1. Three fundamental causes

If standard economic models of factor accumulation and endogenous technical change only provide proximate explanations of comparative growth, what types of explanations would constitute fundamental ones? Though there is no conventional wisdom on this, we can distinguish three such theories: the first set of theories, our main focus in this chapter, emphasize the importance of economic institutions, which influence economic outcomes by shaping economic incentives; the second emphasize geography, and the third emphasize the importance of culture (a fourth possibility is that differences are due to "luck", some societies were just lucky; however we do not believe that differences in luck by themselves constitute a sufficient fundamental causes of cross-country income differences).

2.1.1. Economic institutions

At its core, the hypothesis that differences in economic institutions are the fundamental cause of different patterns of economic growth is based on the notion that it is the way that humans themselves decide to organize their societies that determines whether or not they prosper. Some ways of organizing societies encourage people to innovate, to take risks, to save for the future, to find better ways of doing things, to learn and educate themselves, solve problems of collective action and provide public goods. Others do not.

The idea that the prosperity of a society depends on its economic institutions goes back at least to Adam Smith, for example in his discussions of mercantilism and the role of markets, and was prominent in the work of many nineteenth century scholars such as John Stuart Mill [see the discussion in Jones (1981)]: societies are economically successful when they have 'good' economic institutions and it is these institutions that are the *cause* of prosperity. We can think of these good economic institutions as consisting of an inter-related cluster of things. There must be enforcement of property rights for a broad cross-section of society so that all individuals have an incentive to invest, innovate and take part in economic activity. There must also be some degree of equality of opportunity in society, including such things as equality before the law, so that those with good investment opportunities can take advantage of them.[3]

One could think of other types of economic institutions and many explanations for growth and development have moved beyond models based on preferences, technology and factor endowments to focus on what might loosely be called institutions. One

[3] In Acemoglu, Johnson and Robinson (2001), we coined the term *institutions of private property* for a cluster of good economic institutions, including the rule of law and the enforcement of property rights, and the term *extractive institutions* to designate institutions under which the rule of law and property rights are absent for large majorities of the population.

set of ideas, important for our work, has emphasized that conflict over resources and predation, as well as production, are fundamental forces in society. Scholars such as Skaperdas (1992), Grossman and Kim (1995, 1996), Hirshleifer (2001) and Dixit (2004) have examined how stable property rights can emerge in such circumstances. These scholars have studied almost institution free models and asked how the type of social order that underlies standard economic models might emerge endogenously. Closely related to this work is the research that shows how rent-seeking and redistributional conflict more generally has important implications for growth [e.g., Tornell and Velasco (1992), Murphy, Shleifer and Vishny (1991) Acemoglu (1995), Alesina and Perotti (1996), Benhabib and Rustichini (1996)].

Another literature, following in the footsteps of traditional accounts of economic growth by historians, following the lead of Adam Smith, has emphasized the perfection and spread of markets, clearly a key economic institution [Pirenne (1937), Hicks (1969)]. Problems of the imperfection or absence of markets can clearly have important ramifications for resource allocation, incentives and growth. A central role here has been played by capital markets. For example, Banerjee and Newman (1993) and Galor and Zeira (1993) propose canonical models of how imperfect financial markets can impede growth and development. Models of poverty traps in the tradition of Rosenstein-Rodan (1943), Murphy, Vishny and Shleifer (1989a, 1989b) and Acemoglu (1995, 1997), are based on the idea that market imperfections can lead to the existence of multiple Pareto-ranked equilibria. As a consequence a country can get stuck in a Pareto inferior equilibrium, associated with poverty, but getting out of such a trap necessitates coordinated activities that the market cannot deliver. Other mechanisms, such as increasing returns to scale, can lead to similar situations [e.g., Durlauf (1993), Krugman and Venables (1995), see Azariadis and Stachurski (2005), for other mechanisms and examples]. The implications of many other types of market imperfections have been considered, for example in the labor market [Aghion and Howitt (1994), Pissarides (2000)] and other scholars have examined the implications of industrial organization, market structure and the nature of competition [e.g., Acemoglu and Zilibotti (1997), Aghion et al. (2001), Aghion and Howitt (2005)].

The idea that market imperfections and economic institutions play a central role in development has also been important in the academic literature on development economics since its initiation. Both Adam Smith and Alfred Marshall argued that sharecropping was an inefficient way of organizing agriculture because it gave incorrect incentives to tenants. This argument has been formalized, and at the heart of a large literature on development are imperfections in tenancy, labor, land and credit markets [see Ray (1998), Bardhan and Udry (1999), Banerjee and Duflo (2005)].

Finally, the literature that one might broadly class as institutional has extensively discussed political economy models. Most influential is the early work of Perotti (1993), Saint-Paul and Verdier (1993), Alesina and Rodrik (1994) and Persson and Tabellini (1994) who developed dynamic models to examine the effect of redistributive taxation on growth. There are now many models where political mechanisms and outcomes can have important influences on the growth rate [see Ades and Verdier (1996), Krusell and

Ríos-Rull (1999), Bourguignon and Verdier (2000) and other contributions which we discuss in the body of the paper].

At some level then there is a bewildering array of ideas connecting institutions, both economic and political, to growth and development. In this chapter however, as will already be apparent, we do not attempt to survey all of these theories. Rather, we attempt to develop a perspective on this topic which revolves around what we see as the key issues. From the empirical side this entails really establishing the causal role of institutions in development. From the theoretical side this involves emphasizing the importance of understanding why institutions differ across countries. From the perspective of this chapter the main problem with most of the existing research is the lack of comparative statics and the absence of a truly comparative focus. For instance, in the model of Grossman and Kim (1995) stable property rights may emerge as an equilibrium, but whether they do so or not depends on parameters in the fighting technology which are hard to interpret in reality. Most models of imperfect markets and multiple equilibria fail to provide explanations either for why markets are incomplete or imperfect, or for how some societies manage to get into good equilibria while others do not. To the extent that imperfect market are grounded in imperfections in information or possibilities for opportunism, one would like to know how and why these vary across countries in ways which are consistent with the basic facts about relative economic outcomes. We believe that the structure of markets is endogenous, and partly determined by property rights. Once individuals have secure property rights and there is equality of opportunity, the incentives will exist to create and improve markets (even though achieving perfect markets would be typically impossible). Thus we expect differences in markets to be an outcome of differing systems of property rights and political institutions, not unalterable characteristics responsible for cross-country differences in economic performance. This motivates our focus on economic institutions related to the enforcement of the property rights of a broad cross-section of society.

There are some genuinely comparative studies in the literature. For example, Banerjee and Newman (1993), Alesina and Rodrik (1994) and Persson and Tabellini (1994) all point to differences in wealth distribution as the key to success or failure. We will discuss other such theories, for example those connected to legal origins [e.g., La Porta et al. (1998)] later. Nevertheless, these studies are very different from the approach we propose in this chapter.

2.1.2. Geography

While institutional theories emphasize the importance of man-made factors shaping incentives, an alternative is to focus on the role of "nature", that is, on the physical and geographical environment. In the context of understanding cross-country differences in economic performance, this approach emphasizes differences in geography, climate and ecology that determine both the preferences and the opportunity set of individual economic agents in different societies. We refer to this broad approach as the "geography hypothesis". There are at least three main versions of the geography hypothesis, each emphasizing a different mechanism for how geography affects prosperity.

First, climate may be an important determinant of work effort, incentives, or even productivity. This idea dates back at least to the famous French philosopher, Montesquieu (1748), who wrote in his classic book The Spirit of the Laws: "The heat of the climate can be so excessive that the body there will be absolutely without strength. So, prostration will pass even to the spirit; no curiosity, no noble enterprise, no generous sentiment; inclinations will all be passive there; laziness there will be happiness", and "People are ... more vigorous in cold climates. The inhabitants of warm countries are, like old men, timorous; the people in cold countries are, like young men, brave." One of the founders of modern economics Marshall is another prominent figure who emphasized the importance of climate, arguing: "vigor depends partly on race qualities: but these, so far as they can be explained at all, seem to be chiefly due to climate" [Marshall (1890, p. 195)].

Second, geography may determine the technology available to a society, especially in agriculture. This view is developed by an early Nobel Prize winner in economics, Myrdal, who wrote "serious study of the problems of underdevelopment ... should take into account the climate and its impacts on soil, vegetation, animals, humans and physical assets – in short, on living conditions in economic development" [Myrdal (1968, vol. 3, p. 2121)]. More recently, Diamond espouses this view, "... proximate factors behind Europe's conquest of the Americas were the differences in all aspects of technology. These differences stemmed ultimately from Eurasia's much longer history of densely populated ... [societies dependent on food production]", which was in turn determined by geographical differences between Europe and the Americas [Diamond (1997, p. 358)]. The economist Sachs has been a recent and forceful proponent of the importance of geography in agricultural productivity, stating that "By the start of the era of modern economic growth, if not much earlier, temperate-zone technologies were more productive than tropical-zone technologies ..." [Sachs (2001, p. 2)].

The third variant of the geography hypothesis, especially popular over the past decade, links poverty in many areas of the world to their "disease burden", emphasizing that: "The burden of infectious disease is similarly higher in the tropics than in the temperate zones" [Sachs (2000, p. 32)]. Bloom and Sachs (1998) claim that the prevalence of malaria, a disease which kills millions of children every year in sub-Saharan Africa, reduces the annual growth rate of sub-Saharan African economies by more than 1.3 percent a year (this is a large effect, implying that had malaria been eradicated in 1950, income per capita in sub-Saharan Africa would be double what it is today).

2.1.3. Culture

The final fundamental explanation for economic growth emphasizes the idea that different societies (or perhaps different races or ethnic groups) have different cultures, because of different shared experiences or different religions. Culture is viewed as a key determinant of the values, preferences and beliefs of individuals and societies and, the argument goes, these differences play a key role in shaping economic performance.

At some level, culture can be thought to influence equilibrium outcomes for a given set of institutions. Possibly there are multiple equilibria connected with any set of institutions and differences in culture mean that different societies will coordinate on different equilibria. Alternatively, as argued by Greif (1994), different cultures generate different sets of beliefs about how people behave and this can alter the set of equilibria for a given specification of institutions (for example, some beliefs will allow punishment strategies to be used whereas others will not).

The most famous link between culture and economic development is that proposed by Weber (1930) who argued that the origins of industrialization in western Europe could be traced to the Protestant reformation and particularly the rise of Calvinism. In his view, the set of beliefs about the world that was intrinsic to Protestantism were crucial to the development of capitalism. Protestantism emphasized the idea of predestination in the sense that some individuals were 'chosen' while others were not. "We know that a part of humanity is saved, the rest damned. To assume that human merit or guilt play a part in determining this destiny would be to think of God's absolutely free decrees, which have been settled from eternity, as subject to change by human influence, an impossible contradiction" [Weber (1930, p. 60)].

But who had been chosen and who not? Calvin did not explain this. Weber (1930, p. 66) notes "Quite naturally this attitude was impossible for his followers ... for the broad mass of ordinary men ... So wherever the doctrine of predestination was held, the question could not be suppressed whether there was any infallible criteria by which membership of the *electi* could be known". Practical solutions to this problem were quickly developed, "... in order to attain that self-confidence intense worldly activity is recommended as the most suitable means. It and it alone disperses religious doubts and gives the certainly of grace" [Weber (1930, pp. 66–67)].

Thus "however useless good works might be as a means of attaining salvation ... nevertheless, they are indispensable as a sign of election. They are the technical means, not of purchasing salvation, but of getting rid of the fear of damnation" (p. 69). Though economic activity was encouraged, enjoying the fruits of such activity was not. "Waste of time is ... the first and in principle the deadliest of sins. The span of human life is infinitely short and precious to make sure of one's own election. Loss of time through sociability, idle talk, luxury, even more sleep than is necessary for health ... is worthy of absolute moral condemnation ... Unwillingness to work is symptomatic of the lack of grace" (pp. 104–105).

Thus Protestantism led to a set of beliefs which emphasized hard work, thrift, saving, and where economic success was interpreted as consistent with (if not actually signaling) being chosen by God. Weber contrasted these characteristics of Protestantism with those of other religions, such as Catholicism, which he argued did not promote capitalism. For instance on his book on Indian religion he argued that the caste system blocked capitalist development [Weber (1958, p. 112)].

More recently, scholars, such as Landes (1998), have also argued that the origins of Western economic dominance are due to a particular set of beliefs about the world and how it could be transformed by human endeavor, which is again linked to religious differences. Although Barro and McCleary (2003) provide evidence of a positive correlation between the prevalence of religious beliefs, notably about hell and heaven, and economic growth, this evidence does not show a causal effect of religion on economic growth, since religious beliefs are endogenous both to economic outcomes and to other fundamental causes of income differences [points made by Tawney (1926), and Hill (1961b), in the context of Weber's thesis].

Ideas about how culture may influence growth are not restricted to the role of religion. Within the literature trying to explain comparative development there have been arguments that there is something special about particular cultural endowments, usually linked to particular nation states. For instance, Latin America may be poor because of its Iberian heritage, while North America is prosperous because of its Anglo-Saxon heritage [Véliz (1994)]. In addition, a large literature in anthropology argues that societies may become 'dysfunctional' or 'maladapted' in the sense that they adopt a system of beliefs or ways or operating which do not promote the success or prosperity of the society [see Edgerton (1992), for a survey of this literature]. The most famous version of such an argument is due to Banfield (1958) who argued that the poverty of Southern Italy was due to the fact that people had adopted a culture of "amoral familiarism" where they only trusted individuals of their own families and refused to cooperate or trust anyone else. This argument was revived in the extensive empirical study of Putnam, Leonardi and Nanetti (1993) who characterized such societies as lacking "social capital". Although Putnam and others, for example, Knack and Keefer (1997) and Durlauf and Fafchamps (2004), document positive correlations between measures of social capital and various economic outcomes, there is no evidence of a causal effect, since, as with religious beliefs discussed above, measures of social capital are potentially endogenous.

3. Institutions matter

We now argue that there is convincing empirical support for the hypothesis that differences in economic institutions, rather than geography or culture, *cause* differences in incomes per-capita. Consider first Figure 1.

This shows the cross-country bivariate relationship between the log of GDP per-capita in 1995 and a broad measure of property rights, "protection against expropriation risk", averaged over the period 1985 to 1995. The data on economic institutions come from Political Risk Services, a private company which assesses the risk that investments will be expropriated in different countries. These data, first used by Knack and Keefer (1995) and subsequently by Hall and Jones (1999) and Acemoglu, Johnson and Robinson (2001, 2002) are imperfect as a measure of economic institutions, but the findings are robust to using other available measures of economic institutions. The scatter plot

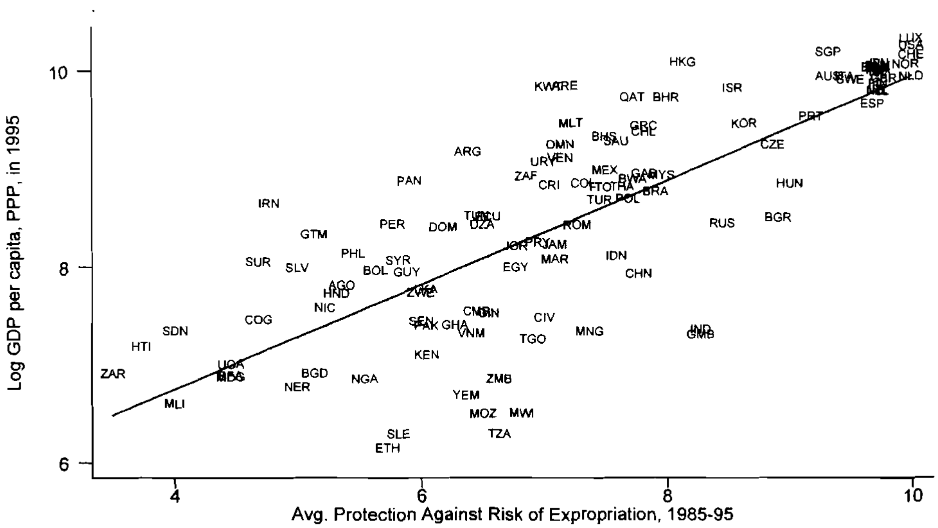

Figure 1. Average protection against risk of expropriation 1985–95 and log GDP per capita 1995.

shows that countries with more secure property rights, i.e., better economic institutions, have higher average incomes.

It is tempting to interpret Figure 1 as depicting a causal relationship (i.e., as establishing that secure property rights cause prosperity). Nevertheless, there are well-known problems with making such an inference. First, there could be reverse causation – perhaps only countries that are sufficiently wealthy can afford to enforce property rights. More importantly, there might be a problem of omitted variable bias. It could be something else, e.g., geography, that explains both why countries are poor and why they have insecure property rights. Thus if omitted factors determine institutions and incomes, we would spuriously infer the existence of a causal relationship between economic institutions and incomes when in fact no such relationship exists. Trying to estimate the relationship between institutions and prosperity using Ordinary Least Squares, as was done by Knack and Keefer (1995) and Barro (1997) could therefore result in biased regression coefficients.

To further illustrate these potential *identification* problems, suppose that climate, or geography more generally, matters for economic performance. In fact, a simple scatterplot shows a positive association between latitude (the absolute value of distance from the equator) and income per capita. Montesquieu, however, not only claimed that warm climate makes people lazy and thus unproductive, but also unfit to be governed by democracy. He argued that despotism would be the political system in warm climates. Therefore, a potential explanation for the patterns we see in Figure 1 is that there is an omitted factor, geography, which explains both economic institutions and economic performance. Ignoring this potential third factor would lead to mistaken conclusions.

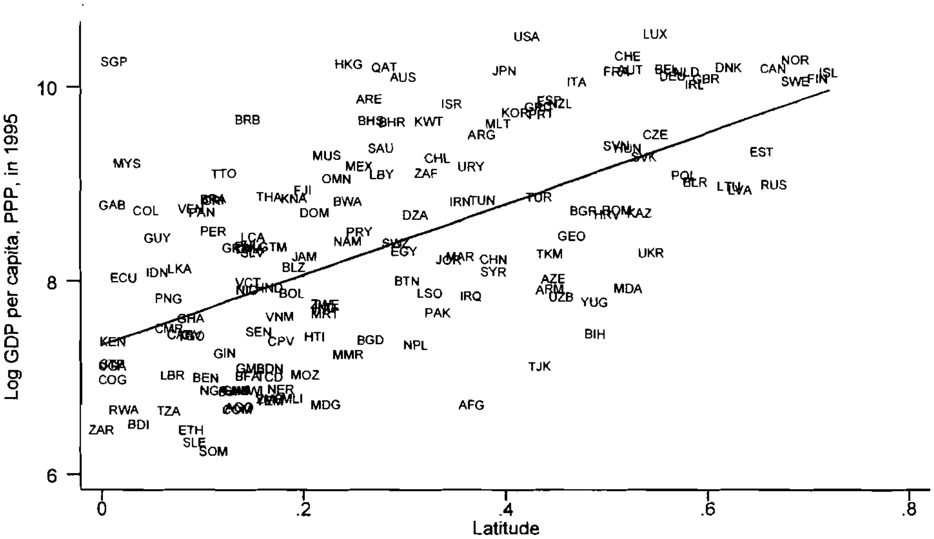

Figure 2. Latitude and log GDP per capita 1995.

Even if Montesquieu's story appears both unrealistic and condescending to our modern sensibilities, the general point should be taken seriously: the relationship shown in Figure 1, and for that matter that shown in Figure 2, is not causal. As we pointed out in the context of the effect of religion or social capital on economic performance, these types of scatterplots, correlations, or their multidimensional version in OLS regressions, *cannot* establish causality.

What can we do? The solution to these problems of inference is familiar in microeconometrics: find a source of variation in economic institutions that should have no effect on economic outcomes, or depending on the context, look for a natural experiment. As an example, consider first one of the clearest natural experiments for institutions.

3.1. *The Korean experiment*

Until the end of World War II, Korea was under Japanese occupation. Korean independence came shortly after the Japanese Emperor Hirohito announced the Japanese surrender on August 15, 1945. After this date, Soviet forces entered Manchuria and North Korea and took over the control of these provinces from the Japanese. The major fear of the United States during this time period was the takeover of the entire Korean peninsular either by the Soviet Union or by communist forces under the control of the former guerrilla fighter, Kim Il Sung. U.S. authorities therefore supported the influential nationalist leader Syngman Rhee, who was in favor of separation rather than a united communist Korea. Elections in the South were held in May 1948, amidst a widespread boycott by Koreans opposed to separation. The newly elected representatives proceeded to draft a new constitution and established the Republic of Korea to the south of the 38th parallel. The North became the Democratic People's Republic of Korea, under the control of Kim Il Sung. These two independent countries organized themselves in very different ways and adopted completely different sets of institutions. The North followed the model of Soviet socialism and the Chinese Revolution in abolishing private property of land and capital. Economic decisions were not mediated by the market, but by the communist state. The South instead maintained a system of private property and the government, especially after the rise to power of Park Chung Hee in 1961, attempted to use markets and private incentives in order to develop the economy.

Before this "natural experiment" in institutional change, North and South Korea shared the same history and cultural roots. In fact, Korea exhibited an unparalleled degree of ethnic, linguistic, cultural, geographic and economic homogeneity. There are few geographic distinctions between the North and South, and both share the same disease environment. For example, the CIA Factbook describes the climate of North Korea as "temperate with rainfall concentrated in summer" and that of South Korea as "temperate, with rainfall heavier in summer than winter". In terms of terrain North Korea is characterized as consisting of "mostly hills and mountains separated by deep, narrow valleys; coastal plains wide in west, discontinuous in east", while South Korea is "mostly hills and mountains; wide coastal plains in west and south". In terms of natural resources North Korea is better endowed with significant reserves of coal, lead, tungsten, zinc, graphite, magnesite, iron ore, copper, gold, pyrites, salt, fluorspar, hydropower. South Korea's natural resources are "coal, tungsten, graphite, molybdenum, lead, hydropower potential". Both countries share the same geographic possibilities in terms of access to markets and the cost of transportation.

Other man-made initial economic conditions were also similar, and if anything, advantaged the North. For example, there was significant industrialization during the colonial period with the expansion of both Japanese and indigenous firms. Yet this development was concentrated more in the North than the South. For instance, the large Japanese zaibatsu of Noguchi, which accounted for one third of Japanese investment in Korea, was centered in the North. It built large hydroelectric plants, including the Suiho dam on the Yalu river, second in the world only to the Boulder dam on the Colorado river. It also created Nippon Chisso, the second largest chemical complex in the world that was taken over by the North Korean state. Finally, in Ch'ongjin North Korea also had the largest port on the Sea of Japan. All in all, despite some potential advantages for the North,[4] Maddison (2001) estimates that at the time of separation, North and South Korea had approximately the same income per capita.

We can therefore think of the splitting on the Koreas 50 years ago as a natural experiment that we can use to identify the causal influence of a particular dimension of

[4] Such initial differences were probably eradicated by the intensive bombing campaign that the United States unleashed in the early 1950's on North Korea [see Cumings (2004, Chapter 1)].

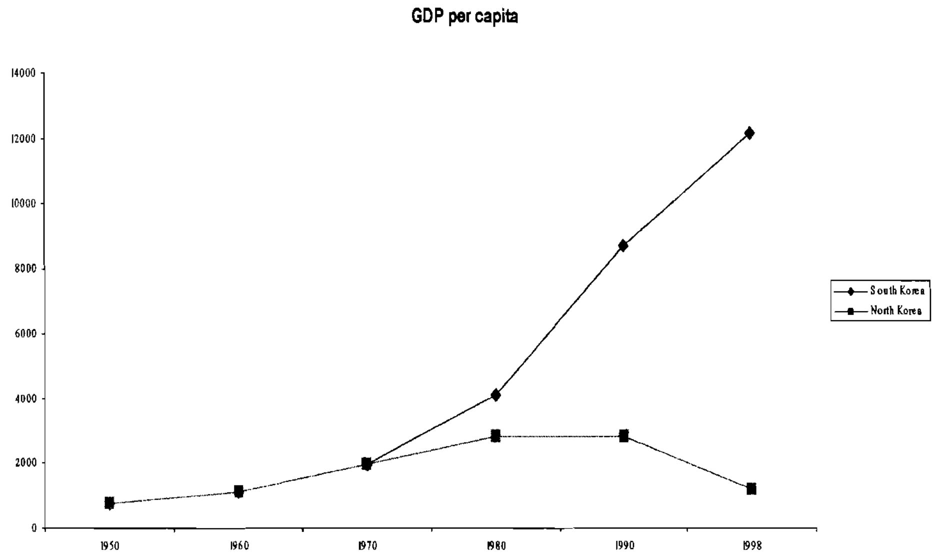

Figure 3. GDP per capita in North and South Korea, 1950–98.

institutions on prosperity. Korea was split into two, with the two halves organized in radically different ways, and with geography, culture and many other potential determinants of economic prosperity held fixed. Thus any differences in economic performance can plausibly be attributed to differences in institutions.

Consistent with the hypothesis that it is institutional differences that drive comparative development, since separation, the two Koreas have experienced dramatically diverging paths of economic development (Figure 3).

By the late 1960's South Korea was transformed into one of the Asian "miracle" economies, experiencing one of the most rapid surges of economic prosperity in history while North Korea stagnated. By 2000 the level of income in South Korea was \$16,100 while in North Korea it was only \$1,000. By 2000 the South had become a member of the Organization of Economic Cooperation and Development, the rich nations club, while the North had a level of per-capita income about the same as a typical sub-Saharan African country. There is only one plausible explanation for the radically different economic experiences on the two Koreas after 1950: their very different institutions led to divergent economic outcomes. In this context, it is noteworthy that the two Koreas not only shared the same geography, but also the same culture.

It is possible that Kim Il Sung and Communist Party members in the North believed that communist policies would be better for the country and the economy in the late 1940s. However, by the 1980s it was clear that the communist economic policies in the North were not working. The continued efforts of the leadership to cling to these policies and to power can only be explained by those leaders wishing to look after their own interests at the expense of the population at large. Bad institutions are therefore kept in place, clearly not for the benefit of society as a whole, but for the benefit of the ruling elite, and this is a pattern we encounter in most cases of institutional failure that we discuss in detail below.

However convincing on its own terms, the evidence from this natural experiment is not sufficient for the purposes of establishing the importance of economic institutions as the primary factor shaping cross-country differences in economic prosperity. First, this is only one case, and in the better-controlled experiments in the natural sciences, a relatively large sample is essential. Second, here we have an example of an extreme case, the difference between a market-oriented economy and a communist one. Few social scientists today would deny that a lengthy period of totalitarian centrally planned rule has significant economic costs. And yet, many might argue that differences in economic institutions among capitalist economies or among democracies are not the major factor leading to differences in their economic trajectories. To establish the major role of economic institutions in the prosperity and poverty of nations we need to look at a larger scale "natural experiment" in institutional divergence.

3.2. The colonial experiment

The colonization of much of the world by Europeans provides such a large scale natural experiment. Beginning in the early fifteenth century and massively intensifying after 1492, Europeans conquered many other nations. The colonization experience transformed the institutions in many diverse lands conquered or controlled by Europeans. Most importantly, Europeans imposed very different sets of institutions in different parts of their global empire, as exemplified most sharply by the contrast to the economic institutions in the northeast of America to those in the plantation societies of the Caribbean. As a result, while geography was held constant, Europeans initiated large changes in economic institutions, in the social organization of different societies. We will now show that this experience provides evidence which conclusively establishes the central role of economic institutions in development. Given the importance of this material and the details we need to provide, we discuss the colonial experience in the next section.

4. The Reversal of Fortune

The impact of European colonialism on economic institutions is perhaps most dramatically conveyed by a single fact – historical evidence shows that there has been a remarkable Reversal of Fortune in economic prosperity within former European colonies. Societies like the Mughals in India, and the Aztecs and the Incas in the Americas were among the richest civilizations in 1500, yet the nation states that now coincide with the boundaries of these empires are among the poorer societies of today. In contrast, countries occupying the territories of the less-developed civilizations in North America, New Zealand and Australia are now much richer than those in the lands of the Mughals, Aztecs and Incas.

4.1. The reversal among the former colonies

The Reversal of Fortune is not confined to such comparisons. Using reasonable proxies for prosperity before modern times, we can show that it is a much more systematic phenomenon. Our proxies for income per capita in pre-industrial societies are urbanization rates and population density. Only societies with a certain level of productivity in agriculture and a relatively developed system of transport and commerce can sustain large urban centers and a dense population. Figure 4 shows the relationship between income per capita and urbanization (fraction of the population living in urban centers with greater than 5000 inhabitants) today, and demonstrates that in the current era there is a significant relationship between urbanization and prosperity.

Naturally, high rates of urbanization do not mean that the majority of the population lived in prosperity. In fact, before the twentieth century urban areas were centers of poverty and ill health. Nevertheless, urbanization is a good proxy for average income per capita in society, which closely corresponds to the measure we are using to look at prosperity.

Figures 5 and 6 show the relationship between income per capita today and urbanization rates and (log) population density in 1500 for the sample of European colonies.[5]

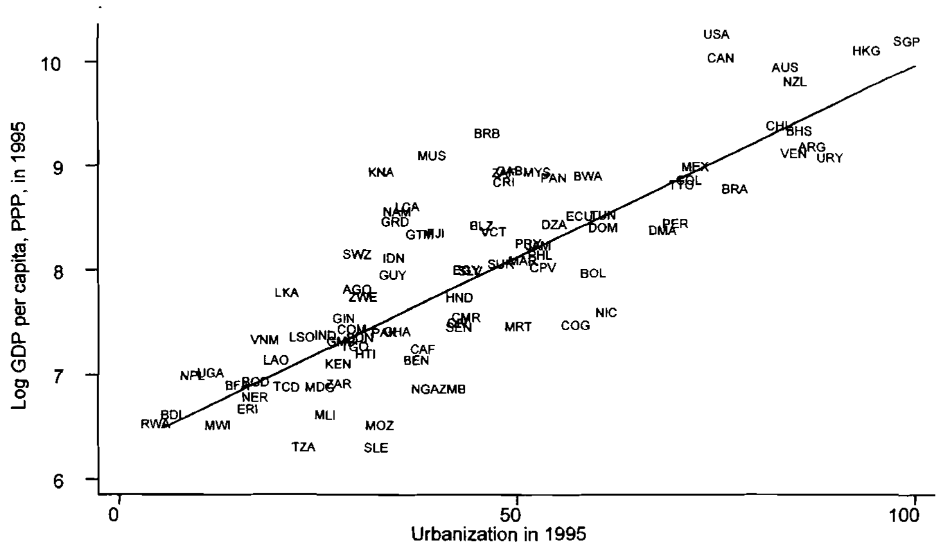

Figure 4. Urbanization in 1995 and log GDP per capita in 1995.

[5] The sample includes the countries colonized by the Europeans between the 15th and the 19th centuries as part of their overseas expansion after the discovery of the New World and the rounding of the Cape of Good Hope. It therefore excludes Ireland, parts of the Russian Empire and also the Middle East and countries briefly controlled by European powers as U.N. Mondays during the 20th century.

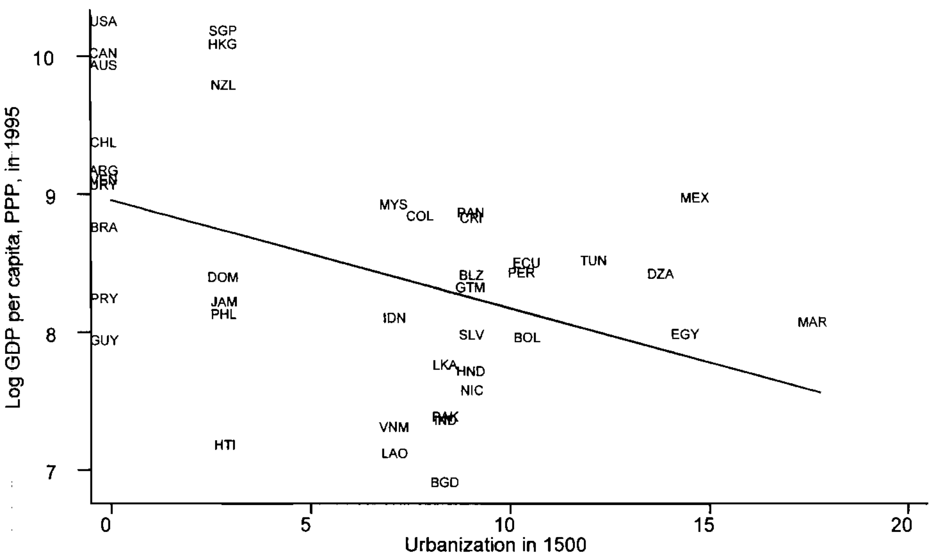

Figure 5. Urbanization in 1500 and log GDP per capita in 1995, among former European colonies.

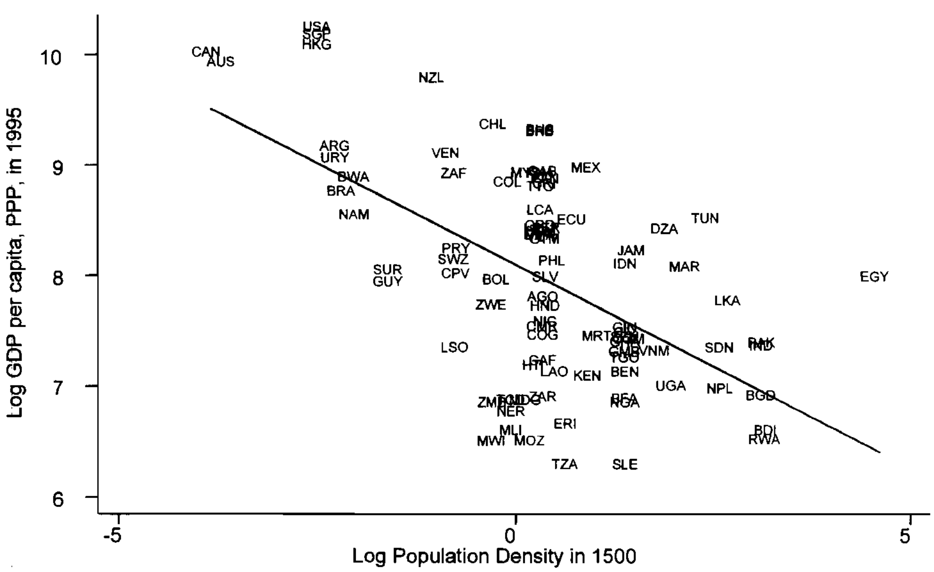

Figure 6. Log population density in 1500 and log GDP per capita in 1995, among former European colonies.

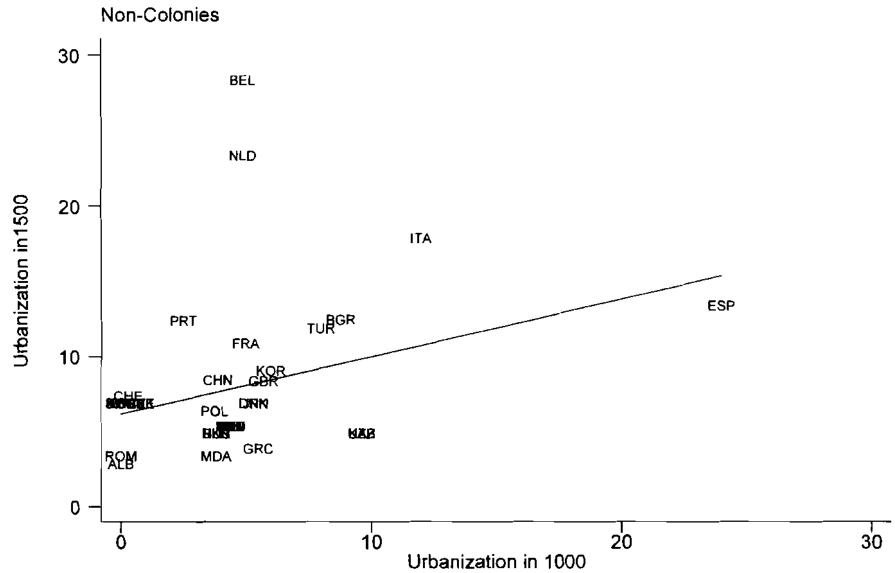

Figure 7. Urbanization in 1000 and 1500, among non-colonies.

We pick 1500 since it is before European colonization had an effect on any of these societies. A strong negative relationship, indicating a reversal in the rankings in terms of economic prosperity between 1500 and today, is clear in both figures. In fact, the figures show that in 1500 the temperate areas were generally less prosperous than the tropical areas, but this pattern too was reversed by the twentieth century.

The urbanization data for these figures come from Bairoch (1988), Bairoch, Batou and Chèvre (1988), Chandler (1987), and Eggimann (1999). The data on population density are from McEvedy and Jones (1978). Details and further results are in Acemoglu, Johnson and Robinson (2002).

There is something extraordinary about this reversal. For example, after the initial spread of agriculture there was remarkable persistence in urbanization and population density for all countries, including those which were to be subsequently colonized by Europeans. In Figures 7 and 8 we show the relationships for urbanization plotting separately the relationship between urbanization in 1000 and in 1500 for the samples of colonies and all other countries. Both figures show persistence, not reversal. Although Ancient Egypt, Athens, Rome, Carthage and other empires rose and fell, what these pictures show is that there was remarkable persistence in the prosperity of regions.

Moreover, reversal was not the general pattern in the world after 1500. Figure 9 shows that within countries not colonized by Europeans in the early modern and modern period, there was no reversal between 1500 and 1995.

There is therefore no reason to think that what is going on in Figures 5 and 6 is some sort of natural reversion to the mean.

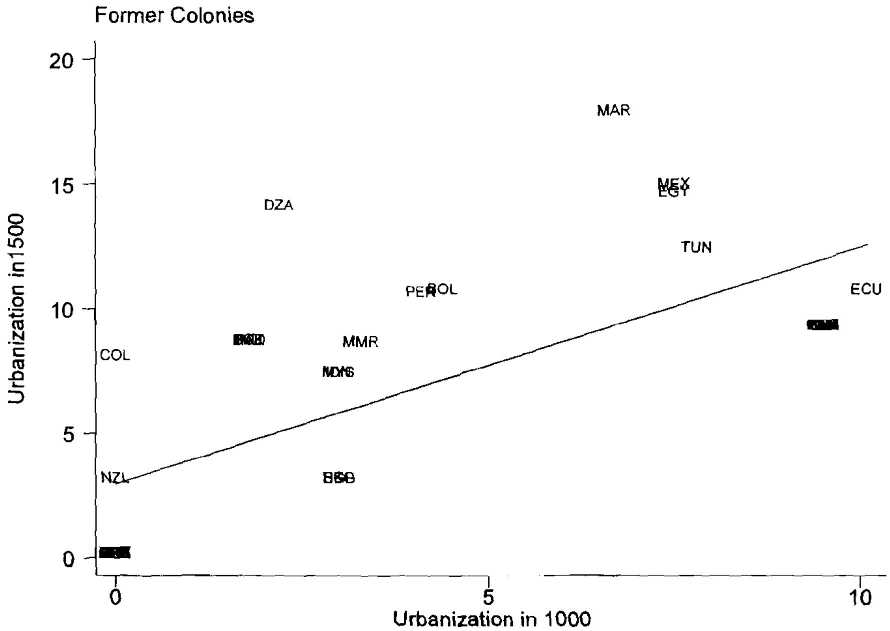

Figure 8. Urbanization in 1000 and 1500, among former European colonies.

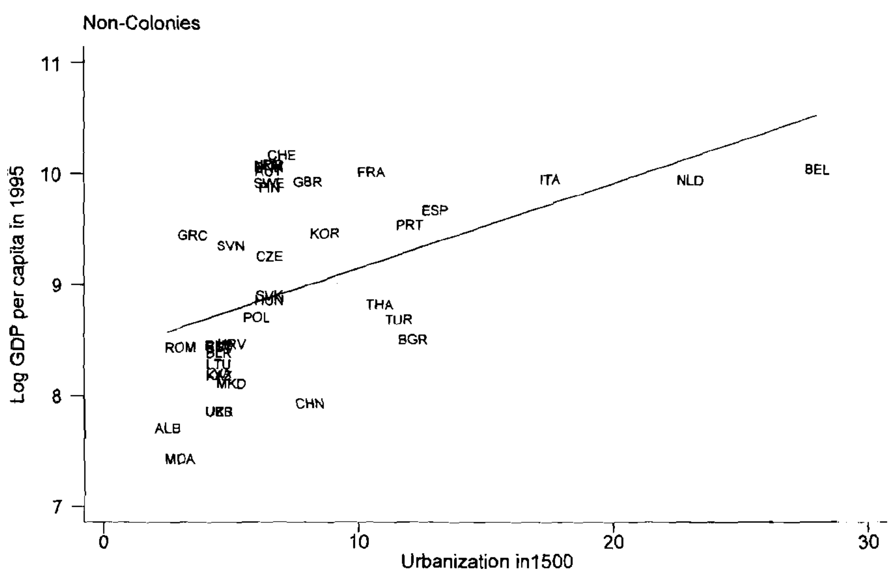

Figure 9. Urbanization in 1500 and log GDP per capita in 1995, among non-colonies.

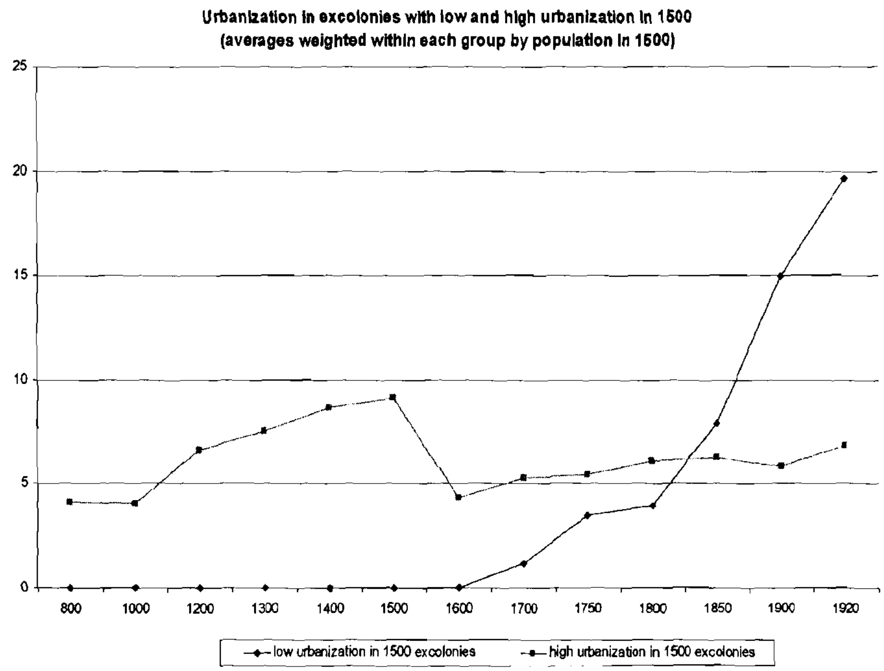

Figure 10. Evolution of urbanization among former European colonies.

4.2. *Timing of the reversal*

When did the reversal occur? One possibility is that it arose shortly after the conquest of societies by Europeans but Figures 10 and 11 show that the previously-poor colonies surpassed the former highly-urbanized colonies starting in the late eighteenth and early nineteenth centuries, and this went hand in hand with industrialization. Figure 10 shows average urbanization in colonies with relatively low and high urbanization in 1500. The initially high-urbanization countries have higher levels of urbanization and prosperity until around 1800. At that time the initially low-urbanization countries start to grow much more rapidly and a prolonged period of divergence begins. Figure 11 shows industrial production per capita in a number of countries. Although not easy to see in the figure, there was more industry (per capita and total) in India in 1750 than in the United States. By 1860, the United States and British colonies with relatively good economic institutions, such as Australia and New Zealand, began to move ahead rapidly, and by 1953, a huge gap had opened up.

4.3. *Interpreting the reversal*

Which of the three broad hypotheses about the sources of cross-country income differences are consistent with the reversal and its timing? These patterns are clearly

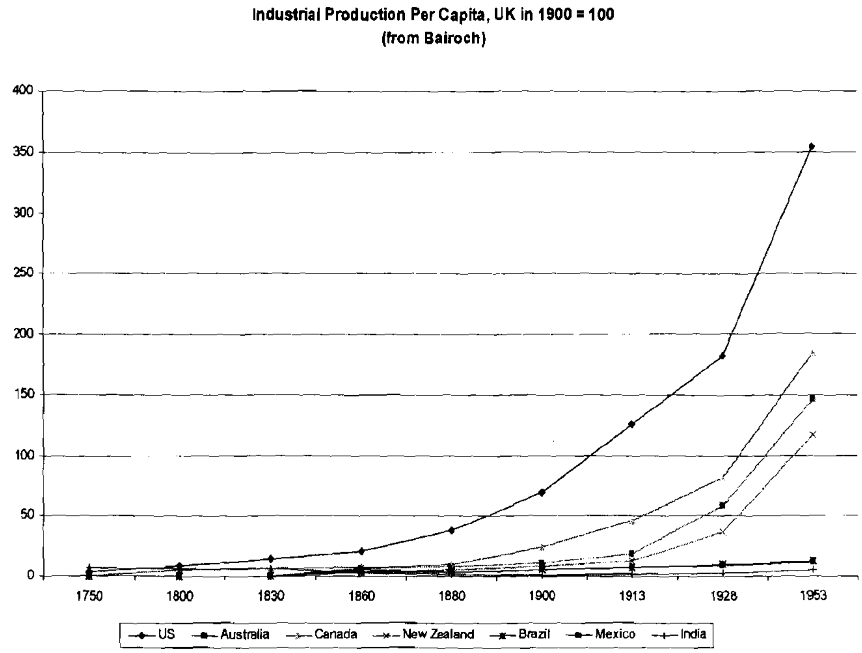

Figure 11. Evolution of industrial production per capita among former European colonies.

inconsistent with simple geography based views of relative prosperity. In 1500 it was the countries in the tropics which were relatively prosperous, in 2003 it is the reverse. This makes it implausible to base a theory of relative prosperity today, as Sachs (2000, 2001) does, on the intrinsic poverty of the tropics. This argument is inconsistent with the historical evidence.

Nevertheless, following Diamond (1997), one could propose what Acemoglu, Johnson and Robinson (2002) call a "sophisticated geography hypothesis" which claims that geography matters but in a time varying way. For example, Europeans created "latitude specific" technology, such as heavy metal ploughs, that only worked in temperate latitudes and not with tropical soils. Thus when Europe conquered most of the world after 1492, they introduced specific technologies that functioned in some places (the United States, Argentina, Australia) but not others (Peru, Mexico, West Africa). However, the timing of the reversal, coming as it does in the nineteenth century, is inconsistent with the most natural types of sophisticated geography hypotheses. Europeans may have had latitude specific technologies, but the timing implies that these technologies must have been industrial, not agricultural, and it is difficult to see why industrial technologies do not function in the tropics (and in fact, they have functioned quite successfully in tropical Singapore and Hong Kong).[6]

[6] A possible link is that proposed by Lewis (1978) who argued that tropical agriculture is less productive than temperate agriculture, and that an 'agricultural revolution' is a prerequisite to an industrial revolution because

Similar considerations weigh against the culture hypothesis. Although culture is slow-changing the colonial experiment was sufficiently radical to have caused major changes in the cultures of many countries that fell under European rule. In addition, the destruction of many indigenous populations and immigration from Europe are likely to have created new cultures or at least modified existing cultures in major ways [see Vargas Llosa (1989), for a fictionalized account of just such a cultural change]. Nevertheless, the culture hypothesis does not provide a natural explanation for the reversal, and has nothing to say on the timing of the reversal. Moreover, we discuss below how econometric models that control for the effect of institutions on income do not find any evidence of an effect of religion or culture on prosperity.

The most natural explanation for the reversal comes from the institutions hypothesis, which we discuss next.

4.4. *Economic institutions and the reversal*

Is the Reversal of Fortune consistent with a dominant role for economic institutions in comparative development? The answer is yes. In fact, once we recognize the variation in economic institutions created by colonization, we see that the Reversal of Fortune is exactly what the institutions hypothesis predicts.

In Acemoglu, Johnson and Robinson (2002) we tested the connection between initial population density, urbanization, and the creation of good economic institutions. We showed that, others things equal, the higher the initial population density or the greater initial urbanization, the worse were subsequent institutions, including both institutions right after independence and today. Figures 12 and 13 show these relationships using the same measure of current economic institutions used in Figure 1, protection against expropriation risk today. They document that the relatively densely settled and highly urbanized colonies ended up with worse (or 'extractive') institutions, while sparsely-settled and non-urbanized areas received an influx of European migrants and developed institutions protecting the property rights of a broad cross-section of society. European colonialism therefore led to an institutional reversal, in the sense that the previously-richer and more-densely settled places ended up with worse institutions.[7]

To be fair, it is possible that the Europeans did not actively introduce institutions discouraging economic progress in many of these places, but inherited them from previous civilizations there. The structure of the Mughal, Aztec and Inca empires were already very hierarchical with power concentrated in the hands of narrowly based ruling elites and structured to extract resources from the majority for the benefit of a minority. Often

high agricultural productivity is needed to stimulate the demand for industrial goods. Though obviously such an explanation is not relevant for explaining industrialization in Singapore or Hong Kong, it may be relevant in other places.

[7] The institutional reversal does not mean that institutions were necessarily better in the previously more densely-settled areas (see next paragraph). It only implies a tendency for the relatively poorer and less densely-settled areas to end up with better institutions than previously-rich and more densely-settled areas.

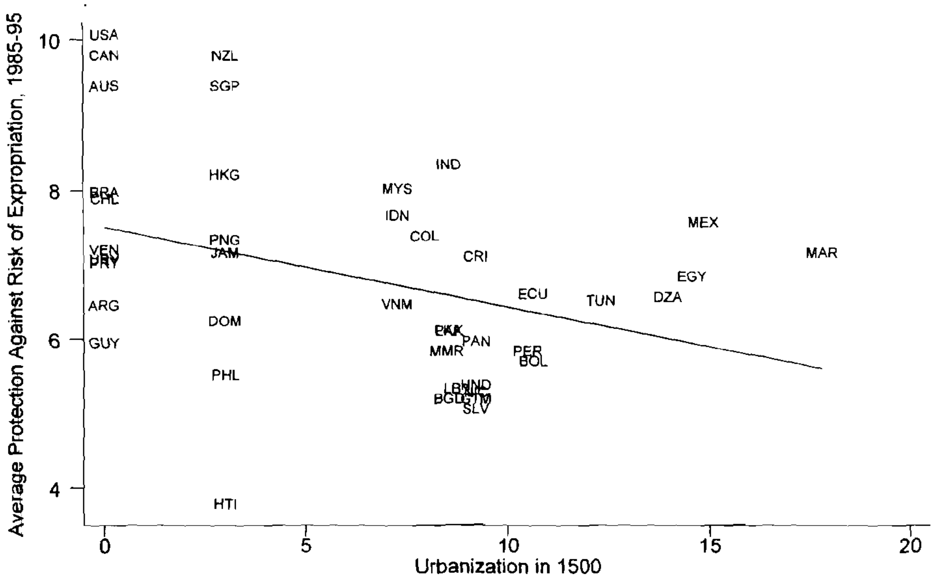

Figure 12. Urbanization in 1500 and average protection against risk of expropriation 1985–95.

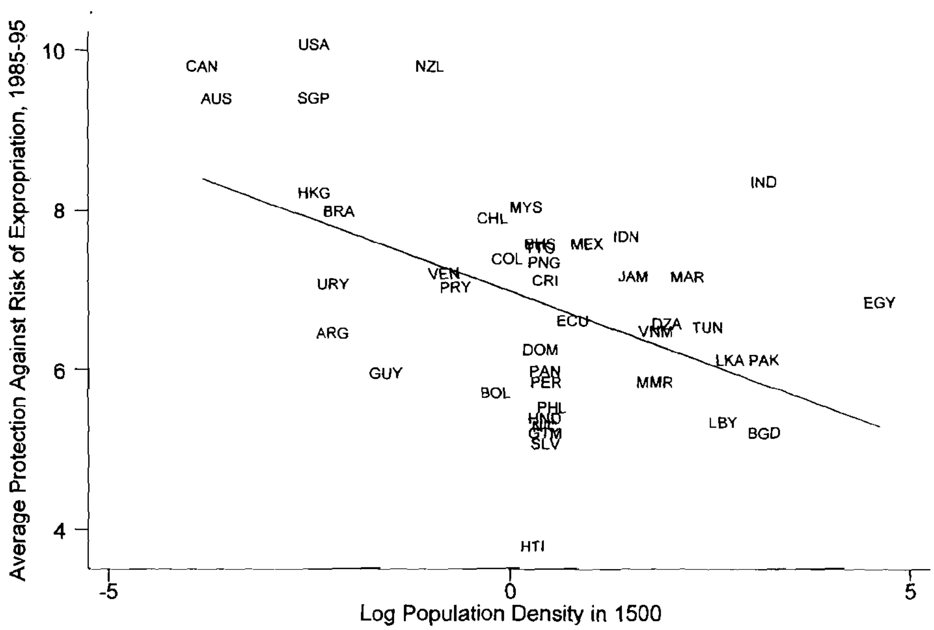

Figure 13. Log population density in 1500 and average protection against risk of expropriation 1985–95.

Europeans simply took over these existing institutions. Whether this is so is secondary for our focus, however. What matters is that in densely-settled and relatively-developed places it was in the interests of Europeans to have institutions facilitating the extraction of resources thus not respecting the property rights of the majority, while in the sparsely-settled areas it was in their interests to develop institutions protecting property rights. These incentives led to an institutional reversal.

The institutional reversal, combined with the institutions hypothesis, predicts the Reversal of Fortune: relatively rich places got relatively worse economic institutions, and if these institutions are important, we should see them become relatively poor over time. This is exactly what we find with the Reversal of Fortune.

Moreover, the institutions hypothesis is consistent with the timing of the reversal. Recall that the institutions hypothesis links incentives to invest in physical and human capital and in technology to economic institutions, and argues that economic prosperity results from these investments. Therefore, economic institutions should become more important when there are major new investment opportunities. The opportunity to industrialize was the major investment opportunity of the nineteenth century. Countries that are rich today, both among the former European colonies and other countries, are those that industrialized successfully during this critical period.

4.5. *Understanding the colonial experience*

The explanation for the reversal that emerges from our discussion so far is one in which the economic institutions in various colonies were shaped by Europeans to benefit themselves. Moreover, because conditions and endowments differed between colonies, Europeans consciously created different economic institutions, which persisted and continue to shape economic performance. Why did Europeans introduce better institutions in previously-poor and unsettled areas than in previously-rich and densely-settled areas? The answer to this question relates to the comparative statics of our theoretical framework. Leaving a full discussion to later, we can note a couple of obvious ideas.

Europeans were more likely to introduce or maintain economic institutions facilitating the extraction of resources in areas where they would benefit from the extraction of resources. This typically meant areas controlled by a small group of Europeans, and areas offering resources to be extracted. These resources included gold and silver, valuable agricultural commodities such as sugar, but most importantly people. In places with a large indigenous population, Europeans could exploit the population, be it in the form of taxes, tributes or employment as forced labor in mines or plantations. This type of colonization was incompatible with institutions providing economic or civil rights to the majority of the population. Consequently, a more developed civilization and a denser population structure made it more profitable for the Europeans to introduce worse economic institutions.

In contrast, in places with little to extract, and in sparsely-settled places where the Europeans themselves became the majority of the population, it was in their interests to introduce economic institutions protecting their own property rights.

4.6. Settlements, mortality and development

The initial conditions we have emphasized so far refer to indigenous population density and urbanization. In addition, the disease environments differed markedly among the colonies, with obvious consequences on the attractiveness of European settlement. As we noted above, when Europeans settled, they established institutions that they themselves had to live under. Therefore, whether Europeans could settle or not had an exogenous effect on the subsequent path of institutional development. In other words, if the disease environment 200 or more years ago affects outcomes today only through its effect on institutions today, then we can use this historical disease environment as an exogenous source of variation in current institutions. From an econometric point of view we have a valid instrument which will enable us to pin down the casual effect of economic institutions on prosperity.[8]

We developed this argument in Acemoglu, Johnson and Robinson (2001) and investigated it empirically. We used initial conditions in the European colonies, particularly data from Curtin (1989, 1998) and Gutierrez (1986) on the mortality rates faced by Europeans (primarily soldiers, sailors, and bishops), as instruments for current economic institutions. The justification for this is that, outside of its effect on economic institutions during the colonial period, historical European mortality has no impact on current income levels. Figures 14 and 15 give scatter plots of this data against contemporaneous economic institutions and GDP per-capita. The sample is countries which were colonized by Europeans in the early modern and modern periods and thus excludes, among others, China, Japan, Korea, Thailand.

Figure 14 shows the very strong relationship between the historical mortality risk faced by Europeans and the current extent to which property rights are enforced. A bivariate regression has an R^2 of 0.26. It also shows that there were very large differences in European mortality. Countries such as Australia, New Zealand and the United States were very healthy with life expectancy typically greater than in Britain. On the other hand mortality was extremely high in Africa, India and South-East Asia. Differential mortality was largely due to tropical diseases such as malaria and yellow fever and at the time it was not understood how these diseases arose nor how they could be prevented or cured.

In Acemoglu, Johnson and Robinson (2001) we showed, using European mortality as an instrument for the current enforcement of property rights, that most of the gap between rich and poor countries today is due to differences in economic institutions. More precisely, we showed (p. 1387) that if one took two typical – in the sense that they both lie on the regression line – countries with high and low expropriation risk, like Nigeria and Chile, then almost the entire difference in incomes per-capita between

[8] Although European mortality is potentially correlated with indigenous mortality, which may determine income today, in practice local populations have developed much greater immunity to malaria and yellow fever. Thus the historical experience of European mortality is a valid instrument for institutional development. See Acemoglu, Johnson and Robinson (2001).

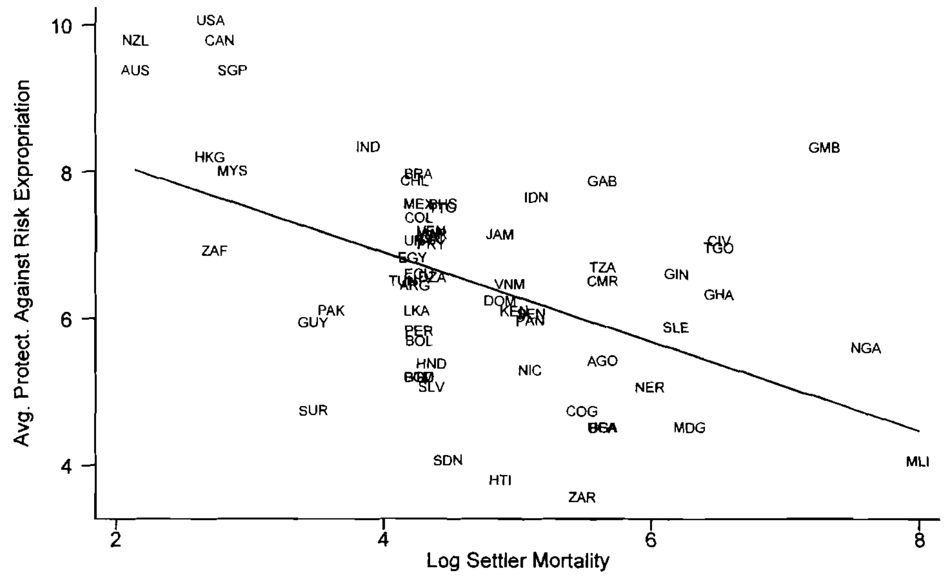

Figure 14. Log mortality of potential European settlers and average protection against risk of expropriation 1985–95.

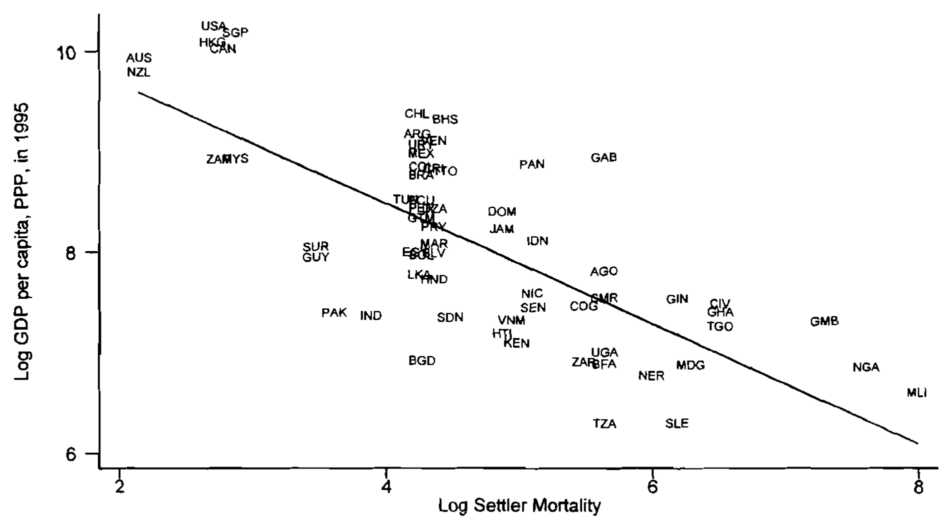

Figure 15. Log mortality of potential European settlers and log GDP per capita in 1995.

them could be explained by the differences in the security of property rights. We also presented regression evidence that showed that once the effect of economic institutions on GDP per-capita was properly controlled for, geographical variables, such as latitude, whether or not a country is land-locked and the current disease environment, have no explanatory power for current prosperity.

These ideas and results provide an interpretation of why there are strong correlations between geographical variables such as latitude and income per-capita. Basically this is because Europeans did not have immunity to tropical diseases during the colonial period and thus settler colonies tended, other things equal, to be created in temperate latitudes. Thus the historical creation of economic institutions was correlated with latitude. Without considering the role of economic institutions it is easy to find a spurious relationship between latitude and income per-capita. However, once economic institutions are properly controlled for, these relationships go away. There is no causal effect of geography on prosperity today, though geography may have been important historically in shaping economic institutions.

What about the role of culture? On the face of it, the Reversal of Fortune is consistent with cultural explanations of comparative growth. The Europeans not only brought new institutions, they also brought their own cultures. There seem to be three main ways to test this idea. First, cultures may be systematically related to the national identity of the colonizing power. For example, the British may have implanted a 'good' Anglo-Saxon culture into colonies such as Australia and the United States, while the Spanish may have condemned Latin America by endowing it with a Hispanic or Iberian culture [the academic literature is full of ideas like this, for recent versions see Véliz (1994), North, Summerhill and Weingast (2000) and Wiarda (2001)]. Second, following Landes (1998), Europeans may have had a culture, for example a work ethic or set of beliefs, which was uniquely propitious to prosperity. Finally, following Weber (1930), Europeans also brought different religions with different implications for prosperity. Such a hypothesis could explain why Latin America is relatively poor since its citizens are primarily Roman Catholic, while North America is relatively rich because its citizens are mostly Protestant.

However, the econometric evidence in Acemoglu, Johnson and Robinson (2001) is not consistent with any these views. Once we control properly for the effects of economic institutions, neither the identity of the colonial power, nor the contemporary fraction of Europeans in the population, nor the proportions of the populations of various religions, are significant determinants of income per capita.

These econometric results are supported by historical examples. For instance, with respect to the identity of the colonizing power, in the 17th century the Dutch had perhaps the best domestic economic institutions in the world but the colonies they created in South-East Asia ended up with institutions designed for the extraction of resources, providing little economic or civil rights to the indigenous population.

It is also be clear that the British in no way simply re-created British institutions in their colonies. For example, by 1619 the North American colony of Virginia had a representative assembly with universal male suffrage, something that did not arrive in Britain

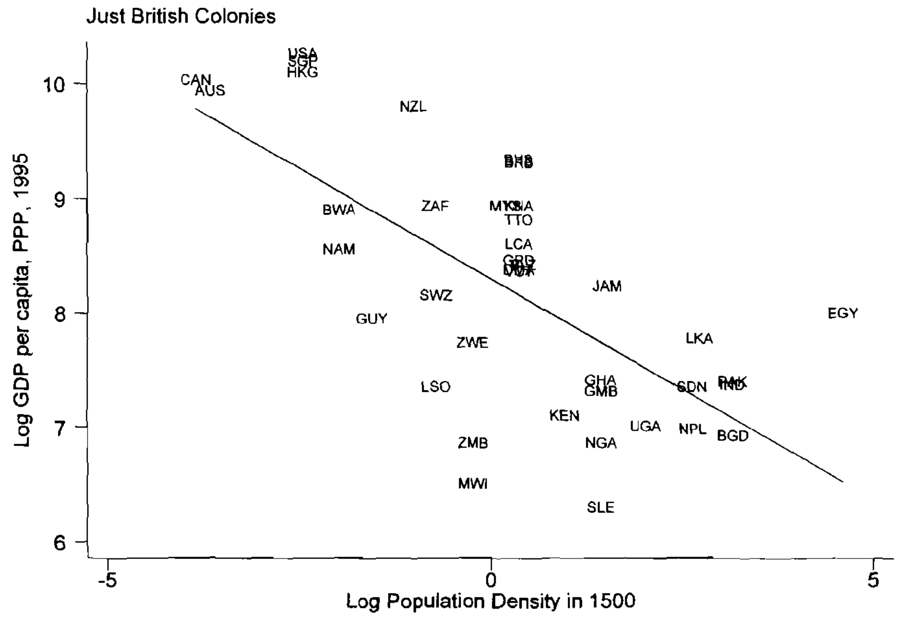

Figure 16. Log population density in 1500 and log GDP per capita in 1995, among former British colonies.

itself until 1919. Another telling example is due to Newton (1914) and Kupperman (1993), who showed that the Puritan colony in Providence Island in the Caribbean quickly became just like any other Caribbean slave colony despite the Puritanical inheritance. Although no Spanish colony has been as successful economically as British colonies such as the United States, it is also important to note that Britain had many unsuccessful colonies (in terms of per capita income), such as in Africa, India and Bangladesh.

To emphasize that the culture or the religion of the colonizer was not at the root of the divergent economic performances of the colonies, Figure 16 shows the reversal among the British colonies (with population density in 1500 on the horizontal axis). Just as in Figure 6, there is a strong negative relationship between population density in 1500 and income per capita today.

With respect to the role of Europeans, Singapore and Hong Kong are now two of the richest countries in the world, despite having negligible numbers of Europeans. Moreover, Argentina and Uruguay have higher proportions of people of European descent than the United States and Canada, but are much less rich. To further document this, Figure 17 shows a similar reversal of fortune for countries where the fraction of those with European descent in 1975 is less than 5 percent of the population.

Overall, the evidence is not consistent with a major role of geography, religion or culture transmitted by the identity of the colonizer or the presence of Europeans. Instead,

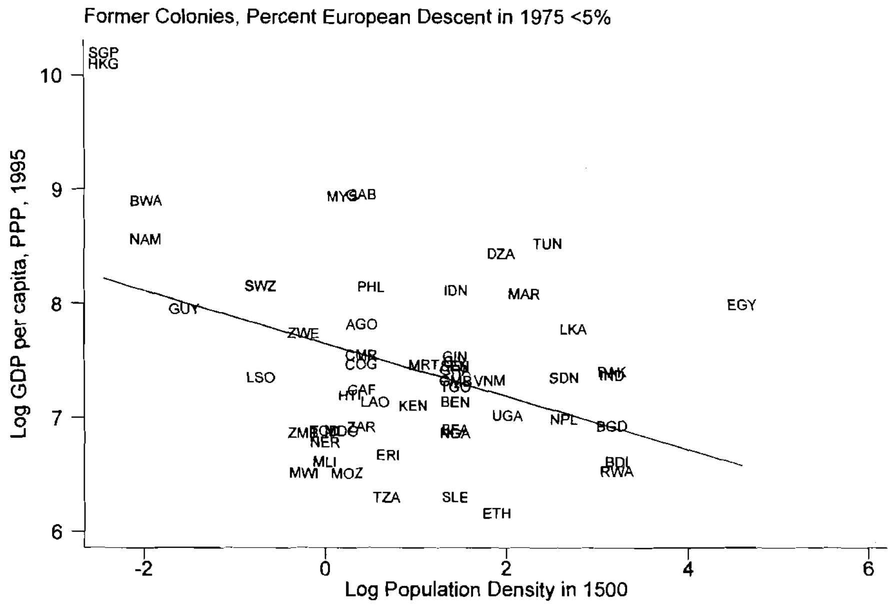

Figure 17. Log population density in 1500 and log GDP per capita in 1995, among former European colonies with current population less than 5% of European descent.

differences in economic institutions appear to be the robust causal factor underlying the differences in income per capita across countries. Institutions are therefore the fundamental cause of income differences and long-run growth.

5. Why do institutions differ?

We saw that economic institutions matter, indeed are central in determining relative prosperity. In terms of the different fundamental theories that we discussed, there is overwhelming support for the emphasis of North and Thomas on institutions, as opposed to alternative candidate explanations which emphasize geography or culture. Yet, as we discussed in the introduction, finding that differences in economic institutions can account for the preponderance of differences in per-capita income between countries creates as many questions as it answers. For example, why do countries have different economic institutions? If poor countries are poor because they have bad economic institutions why do they not change them to better institutions? In short, to explain the evidence presented in the last two sections we need a theory of economic institutions. The theory will help to explain the equilibrium set of economic institutions in a particular country and the comparative statics of this theory will help to explain why economic institutions differ across countries.

In the Introduction (Section 1.2), we began to develop such a theory based on social conflict over economic institutions. We have now substantiated the first point we made there, that economic institutions determine prosperity. We must now move to substantiate our second point, that economic institutions must be treated as endogenous and what which economic institutions emerge depends on the distribution of political power in society. This is a key step towards our theory of economic institutions. In the process of substantiating this point however it is useful to step back and discuss other alternative approaches to developing a theory of economic institutions. Broadly speaking, there are four main approaches to the question of why institutions differ across countries, one of which coincides with the approach we are proposing, the social conflict view. We next discuss each of these separately and our assessment as to whether they provide a satisfactory framework for thinking about differences in economic institutions [see Acemoglu (2003a) and Robinson (1998), for related surveys of some of these approaches]. We shall conclude that the approach we sketched in Section 1.2 is by far the most promising one.

5.1. The efficient institutions view – the Political Coase Theorem

According to this view, societies will choose the economic institutions that are socially efficient. How this surplus will be distributed among different groups or agents does not affect the choice of economic institutions. We stress here that the concept of efficiency is stronger than simply Pareto Optimality; it is associated with surplus, wealth or output maximization.

The underlying reasoning of this view comes from the Coase Theorem. Coase (1960) argued that when different economic parties could negotiate costlessly, they will be able to bargain to internalize potential externalities. A farmer, who suffers from the pollution created by a nearby factory, can pay the factory owner to reduce pollution. Similarly, if the current economic institutions benefit a certain group while creating a disproportionate cost for another, these two groups can negotiate to change the institutions. By doing so they will increase the size of the total surplus that they can divide between themselves, and they can then bargain over the distribution of this additional surplus.

Many different versions of the efficient economic institutions view have been proposed. Indeed, assuming that existing economic institutions are efficient is a standard methodological approach of economists, i.e., observing an institution, one tries to understand what are the circumstances that lead it to be efficient. For instance, Demsetz (1967) argued that private property emerged from common property when land became sufficiently scarce and valuable that it was efficient to privatize it. More recently, Williamson's (1985) research, as well as Coase's (1937) earlier work and the more formal analysis by Grossman and Hart (1986), argues that the governance of firms or markets is such as to guarantee efficiency (given the underlying informational and contractual constraints). Williamson argued that firms emerged as an efficient response to contractual problems that plague markets, particularly the fact that there may be ex-post opportunism when individuals make relationship specific investments. Another famous application of the efficient institutions view is due to North and Thomas (1973) who argued that feudal economic institutions, such as serfdom, were an efficient contract between serfs and lords. The lords provided a public good, protection, in exchange for the labor of the serfs on their lands. In this view, without a modern fiscal system this was an efficient way to organize this exchange. [See Townsend (1993), for a recent version of the idea that other economic institutions of Medieval Europe, such as the open field system, were efficient.]

Williamson and North and Thomas do not specify how different parties will reach agreement to achieve efficient economic institutions, and this may be problematical in the sense that many economic institutions relevant for development are collective choices not individual bargains. There may therefore be free riding problems inherent in the creation of efficient economic institutions. Nevertheless, the underlying idea, articulated by Becker (1958) and Wittman (1989), is that, at least in democracies, competition among pressure groups and political parties will lead to efficient policies and collective choices. In their view, an inefficient economic institution cannot be stable because a political entrepreneur has an incentive to propose a better economic institution and with the extra surplus generated will be able to make himself more attractive to voters. The efficient institutions view regards the structure of political institutions or power as irrelevant. This may matter for the distribution of total surplus, but it will not matter for efficiency itself. The 'efficient' set of political institutions is therefore indeterminate.

The notion that a Coasian logic applies in political life as well as in economics is referred to by Acemoglu (2003a) as the Political Coase Theorem. Although the intuition that individuals and groups will strive towards efficient economic outcomes is appealing, there are both theoretical and empirical limits to the Political Coase Theorem. First, as argued by Acemoglu (2003a) and further discussed below, in politics there is an inherent commitment problem, often making the Political Coase Theorem inapplicable.

Second, the Political Coase Theorem does not take us very far in understanding the effect of economic (or indeed political) institutions on economic outcomes – in this view, economic institutions are chosen efficiently, and all societies have the best possible economic institutions given their needs and underlying structures; hence, with the Political Coase Theorem, economic institutions cannot be the fundamental cause of income differences. However, the empirical results we discussed above suggest a major role for such institutional differences.

The only way to understand these patterns is to think of economic institutions varying for reasons other than the underlying needs of societies. In fact, the instrumental variables and natural experiment strategies we exploited above make use precisely of a source of variation unrelated to the underlying needs of societies. For example, South and North Korea did not adopt very different economic systems because they had different needs, but because different systems were imposed on them for other exogenous reasons. In sum, we need a framework for understanding why certain societies consistently end up with economic institutions that are not, from a social point of view, in their best interests. We need a framework other than the Political Coase Theorem.

5.2. *The ideology view*

A second view is that economic institutions differ across countries because of ideological differences – because of the similarity between this and the previous view, Acemoglu (2003a) calls this the Modified Political Coase Theorem. According to this view, societies may choose different economic institutions, with very different implications, because they – or their leaders – disagree about what would be good for the society. According to this approach, there is sufficient uncertainty about the right economic institutions that well-meaning political actors differ about what's good for their own people. Societies where the leaders or the electorate turn out to be right ex post are those that prosper. The important point is that, just as with the efficient institutions view, there are strong forces preventing the implementation of policies that are *known* to be bad for the society at large.

Several theoretical models have developed related ideas. For example, Piketty (1995) examined a model where different people have different beliefs about how much effort is rewarded in society. If effort is not rewarded then taxation generates few distortions and agents with such beliefs prefer a high tax rate. On the other hand if one believes that effort is rewarded then low taxes are preferable. Piketty showed that dispersion of beliefs could create dispersion of preferences over tax rates, even if all agents had the same objective. Moreover, incorrect beliefs could be self-fulfilling and persist over time because different beliefs tend to generate information consistent with those beliefs. Romer (2003) also presents a model where voters have different beliefs and showed that if mistakes are correlated, then society can choose a socially inefficient outcome. These models show that if different societies have different beliefs about what is socially efficient they can rationally choose different economic institutions.

Belief differences clearly do play a role in shaping policies and institutions. Several interesting examples of this come from the early experience of independence in former British colonies. For example, it is difficult to explain Julius Nyerere's policies in Tanzania without some reference to his and other leading politicians' beliefs about the desirability of a socialist society. It also appears true that in India the Fabian socialist beliefs of Jawaharlal Nehru were important in governing the initial direction that Indian economic policies took.

Nevertheless, the scope of a theory of institutional divergence and comparative development based on ideology seems highly limited. Can we interpret the differences in institutional development across the European colonies or the divergence in the economic institutions and policies between the North and South of Korea as resulting from differences in beliefs? For example, could it be the case that while Rhee, Park, and other South Korean leaders believed in the superiority of capitalist institutions and private property rights enforcement, Kim Il Sung and Communist Party members in the North believed that communist policies would be better for the country?

In the case of South versus North Korea, this is certainly a possibility. However, even if differences in beliefs could explain the divergence in economic institutions in the immediate aftermath of separation, by the 1980s it was clear that the communist economic policies in the North were not working. The continued effort of the leadership to cling to these policies, and to power, can only be explained by leaders looking after their own interests at the expense of the population at large. Most likely, North Korean leaders, the Communist Party, and bureaucratic elites are prolonging the current system, which gives them greater economic and political returns than the alternative, even though they fully understand the costs that the system imposes on the North Korean people.

Differences in colonial policies are even harder to explain on the basis of differences in ideology. British colonists established different economic institutions in very different parts of the world: in the Caribbean they set up plantation societies based on slavery, supported by highly oppressive economic institutions. In contrast, the economic institutions that developed in areas where the British settled, and where there was no large population of indigenous people to be captured and put to work, and where slavery could not be profitably used, such as northeastern United States, Canada, Australia and New Zealand, were very different. Moreover, differences in the incentives of the colonists in various colonies are easy to understand: when they did not settle, they were choosing economic institutions simply to extract resources from the native population. When they settled in large numbers, economic institutions and policies emerged in order to protect them in the future and encourage investment and prosperity.

These considerations make us tend towards a view which emphasizes the actions of key economic and political agents that are taken rationally and in recognition of their consequences, not simply differences in beliefs. We do not deny that belief differences and ideology often play important roles but we do not believe that a satisfactory theory of institutional differences can be founded on differences in ideology.

5.3. The incidental institutions view

The efficient institutions view is explicitly based on economic reasoning: the social costs and benefits of different economic institutions are weighed against each other to determine which economic institutions should prevail. Efficiency arises because individuals ultimately calculate according to social costs and benefits. Institutions are therefore choices. A different approach, popular among many political scientists and sociologists, but also some economists, is to downplay choices and to think of institutions, both economic and political, as the by-product or unintended consequence of other social interactions or historical accidents. In other words, historical accidents at critical junctures determine institutions, and these institutions persist for a long time, with significant consequences.

Here, we discuss two such theories. The first is the theory of political institutions developed by Moore (1966) in his *Social Origins of Dictatorship and Democracy*, the second is the recent emphasis in the economics literature on legal origins, for example as in the work of Shleifer and his co-authors [La Porta et al. (1998, 1999), Djankov et al. (2002, 2003), Glaeser and Shleifer (2002)].

Moore attempted to explain the different paths of political development in Britain, Germany and Russia. In particular, he investigated why Britain evolved into a democracy, while Germany succumbed to fascism and Russia had a communist revolution. Moore stressed the extent of commercialization of agriculture and resulting labor relations in the countryside, the strength of the 'bourgeoisie', and the nature of class coalitions. In his theory, democracy emerged when there was a strong, politically assertive, commercial middle class, and when agriculture had commercialized so that there were no feudal labor relations in the countryside. Fascism arose when the middle classes were weak and entered into a political coalition with landowners. Finally, a communist revolution resulted when the middle classes were non-existent, agriculture was not commercialized, and rural labor was repressed through feudal regulations. In Moore's theory, therefore, class coalitions and the way agriculture is organized determine which political institutions will emerge. However, the organization of agriculture is not chosen with an eye to its effects on political institutions, so these institutions are an unintended consequence. Although Moore was not explicitly concerned with economic development, it is a direct implication of his analysis that societies may end up with institutions that do not maximize income or growth, for example, when they take the path to communist revolution.

Beginning with the work on shareholder rights [La Porta et al. (1998)], continuing to the efficiency of government [La Porta et al. (1999)] and more recently the efficiency of the legal system [Djankov et al. (2003)], Shleifer and his co-authors have argued that a central source of variation in many critical economic institutions is the origin of the legal system. For example, "Civil laws give investors weaker legal rights than common laws do, independent of the level of per-capita income. Common-law countries give both shareholders and creditors – relatively speaking – the strongest, and French-civil-law countries the weakest, protection" [La Porta et al. (1998, p. 1116)].

These differences have important implications for resource allocation. For example, when shareholders have poor protection of their rights, ownership of shares tends to be more highly concentrated. Djankov et al. (2003) collected a cross-national dataset on how different countries legal systems dealt with the issue of evicting a tenant for nonpayment of rent and collecting on a bounced check. They used these data to construct an index of procedural formalism of dispute resolution for each country and showed that such formalism was systematically greater in civil than in common law countries, and is associated with higher expected duration of judicial proceedings, less consistency, less honesty, less fairness in judicial decisions, and more corruption. Legal origins therefore seems to matter for important institutional outcomes.

Where do legal origins come from? The main argument is that they are historical accidents, mostly related to the incidence of European colonialism. For example, Latin American countries adopted the Napoleonic codes in the nineteenth century because these were more compatible with their Spanish legal heritage. Importantly, the fact that Latin American countries therefore have 'French legal origin' is due to a historical accident and can be treated as exogenous with respect to current institutional outcomes. What about the difference between common law and civil law? Glaeser and Shleifer (2002) argue that the divergence between these systems stems from the medieval period and reflects the balance of power between the lords and the king in England and France. Once these systems established, they persisted long after the initial rationale vanished.

Although we believe that historical accidents and persistence are important, in reality the aspect of choice over institutions seems too important to be denied. Even if institutions have a tendency to persist, their persistence is still a choice, in the sense that if the agents decided to change institutions, change would be possible. There are important examples from history of countries radically changing their legal systems such as in Japan after the Meiji restoration, Russia after the Crimean War, and Turkey under Mustafa Kemal in the 1920's. Another example might be central planning of the economy. Though many countries adopted this way of organizing the economy some abandoned it while others, such as North Korea and Cuba, still maintain it. The point here is that though institutions may in some circumstances be the incidental outcome of history, at some point people will start to ask why society has the institutions that it does and to consider other alternatives. At this point we are back in the realm of choice.

5.4. *The social conflict view*

According to this view, economic (and political) institutions are not always chosen by the whole society (and not for the benefit of the whole society), but by the groups that control political power at the time (perhaps as a result of conflict with other groups). These groups will choose the economic institutions that maximize their own rents, and the economic institutions that result may not coincide with those that maximize total surplus, wealth or income. For example, economic institutions that enforce property rights by restricting state predation may not be in the interest of a ruler who wants to appropriate assets in the future. By establishing property rights, this ruler would be reducing his own future rents, so may well prefer economic institutions other than enforced private property. Therefore, equilibrium economic institutions will not be those that maximize the size of the overall pie, but the slice of the pie taken by the powerful groups.

The first systematic development of this point of view in the economics literature is North (1981), who argued in the chapter on "A Neoclassical Theory of the State" that agents who controlled the state should be modeled as self-interested. He then argued that the set of property rights that they would choose for society would be those that maximized their payoff and because of 'transactions costs', these would not necessarily be the set that maximized social welfare. One problem with North's analysis is that he does not clarify what the transactions costs creating a divergence between the interests of the state and the citizens are. Here, we will argue that commitment problems are at the root of this divergence.

The notion that elites, i.e., the politically powerful, may opt for economic institutions which increase their incomes, often at the expense of society, is of course also present in much of the Marxist and dependency theory literature. For example, Dobb (1948), Brenner (1976, 1982) and Hilton (1981) saw feudalism, contrary to North and Thomas's (1973) model, as a set of institutions designed to extract rents from the peasants at the expense of social welfare.[9] Dependency theorists such as Williams (1944), Wallerstein (1974–1980), Rodney (1972), Frank (1978) and Cardosso and Faletto (1979) argued that the international trading system was designed to extract rents from developing countries to the benefit of developed countries.

The social conflict view includes situations where economic institutions may initially be efficient for a set of circumstances but are no longer efficient once the environment changes. For example, Acemoglu, Aghion and Zilibotti (2002) show that though certain sorts of organizations may be useful for countries a long way from the technological frontier, it may be socially efficient to change them subsequently. This may not happen however because it is not privately rational. An interesting example may be the large business enterprises (the *chaebol*) of South Korea. In the context of political institutions, one might then develop a similar thesis. Certain sets of institutions are efficient for very poor countries but they continue to exist even after they cease to be the efficient institutional arrangement.

In stark contrast to the efficient institutions view, political institutions play a crucial role in the social conflict view. Which economic institutions arise depends on who has political power to create or block different economic institutions. Since political institutions play a central role in the allocation of such power they will be an intimate part of a social conflict theory of economic institutions.

What distinguishes the social conflict view from the ideological view is that social conflict can lead to choices of economic institutions which cause underdevelopment even when all agents have common knowledge that this is so. What distinguishes it from the incidental view is that it emphasizes that institutional choices which cause underdevelopment are conscious choices, rather than the result of some historical accident. The aspect that distinguishes the social conflict view from the efficient institutions view is that it does not assume that institutions are always efficient. This is one possible outcome but it is not the only one or indeed the most likely. Why is this? Why cannot efficiency be separated from distribution? We discuss this issue in the next section.

6. Sources of inefficiencies

Having motivated our first two assertions in Section 1.2, we are now in a position to discuss the third, related to the importance of commitment problems. The inability to commit to how political power will be used in the future means that the impact of economic institutions on efficiency cannot be separated from their effects on distribution.[10]

In any market situation where economic exchange takes place, and the quid is separated from the pro quo, issues of commitment will arise. That these issues are of crucial

[9] Postan (1966, pp. 603–604) famously estimated that lords extracted about 50% of the entire production of peasants.

[10] An alternative approach would be to stress informational asymmetries [Farrell (1987)].

importance has been recognized in the literatures on incomplete contracts and renegotiation [e.g., Hart (1995)]. Nevertheless, if the legal system functions properly, there is an array of enforceable contracts that owners can sign with managers, workers with employers, borrowers with lenders, etc. These contracts can be enforced because there is an authority, a third party, with the power to enforce contracts. Although the authority that is delegated to enforce contracts and to resolve disputes varies depending on the exact situation, all such power ultimately emanates from the state, which, in modern society, has a near-monopoly on the use of legitimate coercion. An owner and manager can write a contract because they believe that the state, and its agents the courts, would be impartial enforcers of the contract.

In contrast, if, for example, a manager believed that the state would be aligned with the interests of the owner and refuse to punish the owner if and when he failed to make a payment stipulated by the contract, then the contract would have little value. Therefore, the presence of an impartial enforcer is important for contracting. The problem when it comes to institutional choices is that there is no such impartial third party that can be trusted to enforce contracts. This is the origin of the commitment problem in politics.[11]

To elaborate on this point, let us consider a situation where society can be governed as a dictatorship or as a democracy. Imagine that the dictator does not relinquish his power, but instead he promises that he will obey the rules of democracy, so that individuals can undertake the same investments as they would in democracy. This promise would not necessarily be credible. As long as the political system remains a dictatorship, there is no higher authority to make the dictator stick to his promise. There is no equivalent of a contract that can be enforced by an impartial third-party. After all, the dictator has the monopoly of military and political power, so he is the final arbiter of conflicting interests. There is no other authority to force the dictator to abide by his promises.

A similar problem plagues the reverse solution, whereby the dictator agrees to a voluntary transition to democracy in return for some transfers in the future to compensate him for the lost income and privileges. Those who will benefit from a transition to democracy would be willing to make such promises, but once the dictator relinquishes his political power, there is no guarantee that citizens would agree to tax themselves in order to make payments to this former dictator. Promises of compensation to a former dictator are typically not credible.

The essence of the problem is commitment. Neither party can commit to compensate the other nor can they commit to take actions that would not be in their interests ex post. The reason why commitment problems are severe in these examples is because we are

[11] Many scholars have emphasized the fact that a key feature of political economy is that there is no third party which can enforce the promises made by the state and that this leads to problems of commitment and endemic inefficiencies. This idea is discussed by North (1990) and Olson (1993), is central to the work of Weingast (1997, 1998) and is implicit in many other studies. See also Grossman and Noh (1994), Dixit (1996), Dixit and Londregan (1995), Besley and Coate (1998) and Powell (2004) for discussions of how inability to commit generates inefficiencies in political outcomes.

dealing with political power. Different institutions are associated with different distributions of political power, and there is no outside impartial party with the will and the power to enforce agreements. In some cases, there may be self-enforcing promises that maintain an agreement. Acemoglu (2003a) discusses such possibilities, but in general, there are limits to such self-enforcing agreements, because they require the participants to be sufficiently patient, and when it comes to matters of political power, the future is uncertain enough that no party would behave in a highly patient manner.

Based on this reasoning, we can now discuss three different channels via which the presence of commitment problems will lead to the choice and persistence of inefficient institutions.

6.1. Hold-up

Imagine a situation in which an individual or a group holds unconstrained political power. Also suppose that productive investments can be undertaken by a group of citizens or producers that are distinct from the "political elites", i.e., the current power holders. The producers will only undertake the productive investments if they expect to receive the benefits from their investments. Therefore, a set of economic institutions protecting their property rights are necessary for investment. Can the society opt for a set of economic institutions ensuring such secure property rights? The answer is often no (even assuming that "society" wants to do so).

The problem is that the political elites – those in control of political power – cannot commit to respect the property rights of the producers once the investment are undertaken. Naturally, ex ante, before investments are undertaken, they would like to promise secure property rights. But the fact that the monopoly of political power in their hands implies that they cannot commit to not hold-up producers once the investments are sunk.

This is an obvious parallel to the hold-up problem in the theory of the firm, where once one of the parties in a relationship has undertaken investments specific to the relationship, other parties can hold her up, and capture some of the returns from her investments. As in the theory of the firm, the prospect of hold-up discourages investment. But now the problem is much more severe, since it is not only investments that are specific to a relationship that are subject to hold-up, but all investments.

This is therefore an example of how inefficient economic institutions arise because of a monopoly of political power. Those with political power cannot commit not to use their political power ex post, and this translates directly into a set of economic institutions that do not provide secure property rights to groups without political power. The consequence is clear: without such protection, productive investments are not undertaken, and opportunities for economic growth go unexploited.

The reason why these inefficient economic institutions persist (or may be the equilibrium institutions of the society) is related to commitment problems. Parallel to our above example of inducing the dictator to relinquish power, there are two ways to introduce secure property rights. First, in principle, political elites could promise to respect property rights. However, mere promises would not be credible, unless backed up by the political elites relinquishing power, and this would mean relinquishing their rents and privileges. Second, political elites can be bought off by the beneficiaries of a system of more secure property rights. This would typically be achieved by a promise of future payments. For example, after investments are undertaken and output is produced, a share can be given to the political elites. But, as pointed out above, there is another, reverse commitment problem here; the beneficiaries of the new regime cannot commit to make the promised payments to the previous political elites.

Many real world examples illustrate the commitment problems involved in limiting the use of political power. In practice, although buying off dictators and persuading them to leave power is difficult, there have been many attempts to do so, usually by trying to guarantee that they will not be persecuted subsequently. One way of doing this is to give them asylum in another country. Nevertheless, such attempts rarely succeed, most likely again because of commitment problems (the new regime cannot commit to abide by its promises). An illustrative example of this is the attempts by the Reagan administration to persuade Jean-Claude ('Baby Doc') Duvalier to relinquish power in Haiti in 1986. In the face of a popular uprising and rising social and economic chaos, the Reagan administration, via the intermediation of the Jamaican Prime Minster Edward Seaga, tried to persuade Duvalier to go into exile. He at first agreed and the White House announced his departure on January 30th, but the next day he changed his mind, unsure that he would really be protected, and stayed in Haiti. One month later he was forced into exile in France by the military.

A more common, and in many ways more interesting strategy to induce dictators to relinquish power is to try to structure political institutions so as to guarantee that they will not be punished. Such institutional changes are sometimes important in transitions to democracy. For example, President Pinochet was willing to abide by the results of the 1989 plebiscite he lost in Chile because as a senator the Constitution protected him from prosecution. It was only when he left the country that he was vulnerable.

Although Pinochet's experience illustrates an example of structuring political institutions to achieve commitment, to create durable institutions constraining future use of political power is difficult in practice. These difficulties are well illustrated by the transition from white rule in Rhodesia to majority rule in Zimbabwe. Facing an unwinable guerrilla war, the white elite in Rhodesia sought to negotiate a transition of majority rule, but with enough institutional safeguards that their rents would be protected. These safeguards included the electoral system they wanted, which was used for the first post-independence elections, and massive over-representation in parliament [Reynolds (1999, p. 163)]. Whites were guaranteed 20% of the seats in the legislature for seven years despite making up only 2–3% of the population and were guaranteed 10 seats of the 40 seat senate. Clauses of the 1980 Constitution were also aimed at directly guaranteeing the property rights of the whites. In particular land reform was outlawed for 10 years after which it could only take place if compensated.

The white negotiators at the Lancaster House talks in 1979 that produced these agreements understood that any promises made by the black majority negotiators about what would happen after independence could not be believed. They sought therefore to find a set of rules that would get around this problem [Herbst (1990, pp. 13–36)]. Nevertheless, these guarantees were not enough to protect the property rights (and rents) of the whites in anything other than the short run. The Mugabe regime quickly absorbed the other factions from among the African guerrilla opposition, and more moderate relatively pro-white groups, such as Abel Muzorewa's United African National Council, crumbled. In 1985 the Mugabe regime switched back to the electoral system it preferred [Reynolds (1999, p. 164)] and in 1987, at the first possible opportunity, it removed the guaranteed representation for whites. Though in 1987 Mugabe nominated white candidates for these seats [Horowitz (1991, pp. 135–136)], this did not last for long. In 1990 the senate was abolished. Finally, in 1990 the Constitution was amended to allow for the redistribution of land. Since this time the Mugabe government has begun a sustained policy of land redistribution away from whites through legal and extra-legal means.

6.2. *Political losers*

Another related source of inefficient economic institutions arises from the desire of political elites to protect their political power. Political power is the source of the incomes, rents, and privileges of the elite. If their political power were eroded, their rents would decline. Consequently, the political elite should evaluate every potential economic change not only according to its economic consequences, such as its effects on economic growth and income distribution, but also according to its political consequences. Any economic change that will erode the elite's political power is likely to reduce their economic rents in the long run.

As an example, imagine a change in economic institutions that will increase economic growth, but in doing so, will also enrich groups that could potentially contest political power in the future. Everything else equal, greater economic growth is good for those holding political power. It will create greater returns on the assets that they possess, and also greater incomes that they can tax or expropriate. However, if their potential enemies are enriched, this also means greater threats against their power in the future. Fearing these potential threats to their political power, the elites may oppose changes in economic institutions that would stimulate economic growth.

That the threat of becoming a political loser impedes the adoption of better institutions is again due to a commitment problem. If those who gained political power from institutional change could promise to compensate those who lost power then there would be no incentive to block better institutions.

There are many historical examples illustrating how the fear of losing political power has led various groups of political and economic elites to oppose institutional change and also the introduction of new technologies. Perhaps the best documented examples come from the attitude of the elites to industrialization during the nineteenth century [see Acemoglu and Robinson (2000b, 2002)]. There were large differences between the rates at which countries caught up with British industrialization with many countries completely failing to take advantage of the new technologies and opportunities. In most of these cases, the attitudes of political elites towards industrialization, new technology and institutional change appear to have been the decisive factor, and these attitudes were driven by their fears of becoming political losers. These issues are best illustrated by the experiences of Russia and Austria-Hungary.

In both Russia and Austria-Hungary, absolutist monarchies feared that promoting industrialization would undermine their political power. In Russia, during the reign of Nikolai I between 1825 and 1855 only one railway line was built in Russia, and this was simply to allow the court to travel between Moscow and St. Petersburg. Economic growth and the set of institutions that would have facilitated it were opposed since, as Mosse (1992, p. 19) puts it "it was understood that industrial development might lead to social and political change". In a similar vein, Gregory (1991, p. 74) argues: "Prior to the about face in the 1850s, the Russian state feared that industrialization and modernization would concentrate revolution minded workers in cities, railways would give them mobility, and education would create opposition to the monarchy".

It was only after the defeat in the Crimean War that Nikolai's successor, Alexandr II, initiated a large scale project of railway building and an attempt to modernize the economy by introducing a western legal system, decentralizing government, and ending feudalism by freeing the serfs. This period of industrialization witnessed heightened political tensions, consistent with the fears of the elites that times of rapid change would destabilize the political status quo and strengthen their opposition [McDaniel (1991) gives a detailed account of these events, see also Mosse (1958)].

The consensus view amongst historians also appears to be that the main explanation for the slow growth of Austria-Hungary in the nineteenth century was lack of technology adoption and institutional change, again driven by the opposition of the state to economic change. This view was proposed by Gerschenkron who argued that the state not only failed to promote industrialization, but rather, "economic progress began to be viewed with great suspicion and the railroads came to be regarded, not as welcome carriers of goods and persons, but as carriers of the dreaded revolution. Then the state clearly became an obstacle to the economic development of the country" (1970, p. 89). See also Gross (1973).

The analysis of Freudenberger (1967, pp. 498–499) is similar. As with the Tsar, the Hapsburg emperors opposed the building of railways and infrastructure and there was no attempt to develop an effective educational system. Blum (1943) pointed to the pre-modern institutional inheritance as the major blockage to industrialization arguing (p. 26) that

> "these living forces of the traditional economic system were the greatest barrier to development. Their chief supporter was ... Emperor Francis. He knew that the advances in the techniques of production threatened the life of the old order of which he was so determined a protector. Because of his unique position as final arbiter of all proposals for change he could stem the flood for a time. Thus when plans for the construction of a steam railroad were put before him, he refused to give consent to their execution 'lest revolution might come into the country'."

6.3. Economic losers

A distinct but related source of inefficiency stems from the basic supposition of the social conflict view that different economic institutions imply different distributions of incomes. This implies that a move from a bad to a better set of economic institutions will make some people or groups worse off (and will not be Pareto improving). This in turn implies that such groups will have an incentive to block or impede such institutional changes even if they benefit the whole of society in some aggregate sense.

The idea that economic losers impede the choice of efficient economic institutions and economic policies is widespread in economics and was seen earliest in the literature on international trade. Even though free trade may be socially desirable, individuals invested in sectors in which an economy does not enjoy comparative advantage will lose economically from free trade. Since at least the work of Schattschneider (1935) the role of economic losers has been central in understanding why free trade is not adopted. In the context of development economics, this idea was first discussed by Kuznets (1968), developed at length by Olson (1982, 2000) and Mokyr (1990), and formalized by Krusell and Ríos-Rull (1996) and Parente and Prescott (1999, 2005). Most of the examples discussed in the development literature on economic losers are about technological change – people with specific investments in obsolete technology try to block the introduction of better technology. The most celebrated example is the case of the Luddites, skilled weavers in early nineteenth century England who smashed new mechanized looms which threatened to lead to massive cuts in their wages [see Thomis (1970), Randall (1991)]. Scott (2000, p. 200) relates a similar example from modern Malaysia, "When, in 1976, combine harvesters began to make serious inroads into the wages of poor villagers, the entire region experienced a rash of machine-breaking and sabotage reminiscent of the 1830s in England".

That better economic institutions are blocked by individuals whose incomes are threatened by such change is again due to a problem of commitment. If those whose incomes rose when economic institutions changed could promise to compensate those whose incomes fell then there would be no incentive to block better economic institutions. Nevertheless, it is difficult to commit to such transfers. To consider again the example of the Luddites, the factory owners could have promised to pay the weavers high wages in the future even though their skills were redundant. Once the new technology was in place however, owners would have a clear incentive to fire the weavers and hire much cheaper unskilled workers.[12]

Although the problem of economic losers is appealing at first sight, has received some attention in the economics literature, and fits into our framework by emphasizing the importance of commitment problems, we view it both theoretically and empirically

[12] One possible way round this problem would be for the owners, if they could afford it, to compensate the weavers in advance for their lower future wages. But this would raise the reverse commitment problem: the weavers would have an incentive to take the money and still break the machines – i.e., they could not commit to not blocking the innovations that would reduce their wages even after they had taken the money.

less important than the holdup and the political loser problems. First, as pointed out in Acemoglu and Robinson (2000b), in theories emphasizing issues of economic losers, there are implicit assumptions about politics, which, when spelled out, imply that political concerns must be important whenever issues of economic losers are present. The idea of economic losers is that certain groups, fearing that they will lose their economic rents, prevent adoption of beneficial economic institutions or technologies. The assumption in this scenario is that these groups have the political power to block socially beneficial changes. But then, if they have the political power to block change, why would not they allow the change to take place and then use their political power to redistribute some of the gains to themselves? The implicit assumption must therefore be that groups losing economically also experience a reduction in their political power, making it impossible for them to redistribute the gains to themselves after to change takes place. This reasoning therefore suggests that whether certain groups will lose economically or not is not as essential to their attitudes towards change as *whether their political power will be eroded.* Problems of political losers therefore seem much more important than problems of economic losers.

Possibly for this reason, advocates of the economic losers view have been unable to come up with any well documented examples where the economic losers hypothesis can actually explain first-order patterns of development. For instance, while it is true that the Luddites tried to break machines, they singularly failed to halt the progress of agricultural technology in nineteenth century Britain. The same is true for Malaysia in the 1970s, one of the fastest growing economies in the world at that time. Neither set of workers had sufficient political power to stop change. Indeed, when political powerful groups became economic losers, such as landowners in nineteenth century England who saw land prices and agricultural rents fall rapidly after 1870, they did nothing to block change because their political power allowed them to benefit from efficient economic institutions [Acemoglu and Robinson (2002)].

Perhaps the most interesting failure of economic losers to halt progress in English economic history comes from the impact of the enclosure of common lands. Land has not always been privately owned as property. In much of Africa land is still owned communally, rather than individually, and this was true in Medieval Britain. Starting around 1550 however an 'enclosure movement' gathered pace where 'common land' was divided between cultivators and privatized. By 1850 this process of enclosures had made practically all of Britain private property.

Enclosure was a heterogenous process [Overton (1996, p. 147)] and it also took place at different times in different places. Nevertheless, most of it was in two waves, the so called 'Tudor enclosures' between 1550 and 1700 and the 'parliamentary enclosures' in the century after 1750.

> "From the mid-eighteenth century the most usual way in which common rights were removed was through a specific act of parliament for the enclosure of a particular locality. Such acts ... made the process easier because enclosure could be secured provided the owners of a majority (four fifths) of the land, the lord of the manor, and the owner of the tithe agreed it should take place. Thus the law of

parliament (statue law) only took account of the wishes of those *owning* land as opposed to the common law which took account of all those who had both ownership rights and *use* rights to land. Moreover ... in some parishes the ... majority could be held by a single landowner ... parliamentary enclosure often resulted in a minority of owners imposing their will on the majority of farmers" [Overton (1996, p. 158), italics in original].

The historical evidence is unanimous that the incentive to enclose was because "enclosed land was worth more than open common field land ... the general consensus has been that rents doubled" [Overton (1996, p. 162)]. More controversial is the source of this increase in rent. Overton continues (pp. 162–163) "The proportion of profits taken as rents from tenants by landlords is the outcome of a power struggle between the two groups, and the increase in rent with enclosure may simply reflect an increase in landlord power". Allen (1982, 1992) showed, in his seminal study of the enclosure movement in the South Midlands, that the main impact was a large increase in agricultural rents and a redistribution of income away from those cultivators who had previously used the commons.

The enclosure of common land thus led to a huge increase in inequality in early modern England. Many peasants and rural dwellers had their traditional property rights expropriated. In protest, groups of citizens dispossessed by enclosure attempted to oppose it through collective action and riots – attempting to influence the exercise of political power. These groups were no match for the British state, however. Kett's rebellion of 1549, the Oxfordshire rebellion of 1596, the Midland Revolt of 1607, and others up to the Swing Riots of 1830–1831 were all defeated [see Charlesworth (1983)]. The presence of economic losers did not prevent this huge change in economic institutions and income distribution.

6.4. *The inseparability of efficiency and distribution*

Commitment problems in the use and the allocation of political power therefore introduce a basic trade-off between efficiency and distribution. For example, when lack of commitment causes hold-ups, those who hold political power know that people will not have the right incentives to invest so growth will be low. In response to this, they might voluntarily give away their power or try to create political institutions that restricted their power. Such a change in political institutions would create better investment incentives. Though this situation is hypothetically possible and has formed the basis for some theories of institutional change [e.g., Barzel (2001)] it appears to be insignificant in reality. Even faced with severe underinvestment, political elites are reluctant to give away their power because of its *distributional* implications, i.e., because this would reduce their ability to extract rents from the rest of society. Thus poor economic institutions, here lack of property rights and hold-up, persist in equilibrium because to solve the problem, holders of political power have to voluntarily constrain their power or give it away. This may increase the security of property in society and increase incentives to invest, but it also undermines the ability of rulers to extract rents. They may be better off with a large slice of a small pie.

Similar phenomena are at work when there are either political or economic losers. In the first case, namely a situation where political power holders anticipate being political losers, promoting good institutions directly reduces the political power and rents of incumbents and a similar trade-off emerges. Adopting efficient economic institutions will stimulate growth, but when the political status quo is simultaneously eroded the amount of rent accruing to the initially powerful may fall. In the second case, the incomes of those with political power to determine economic institutions falls directly when better economic institutions are introduced. In the absence of credible commitments to side-payments, those whose incomes fall when better economic institutions are introduced have an incentive to block such institutions.

Because commitment problems seem so endemic in collective choice and politics, it seems natural to believe that institutional change has significant distributional consequences and as a result there will be conflict over the set of institutions in society.

6.5. Comparative statics

Our analysis so far has made some progress towards our theory of differences in economic institutions. Although our full theory is yet to be developed in the later sections, the different mechanisms discussed in this section already point out the major comparative static implications of our approach regarding when economic institutions protecting the property rights of a broad cross-section of society are likely to be adopted, and when they are likely to be opposed and blocked. We now briefly discuss these comparative statics.

Hold-up, political loser and economic loser considerations lead to some interesting comparative static results which can be derived by considering the political institutions that lie behind these phenomena.

1. First, the perspective of hold-ups immediately suggests that situations in which there are constraints on the use of political power, for example, because there is a balance of political power in society or a form of separation of powers between different power-holders, are more likely to engender an environment protecting the property rights of a broad cross-section of society. When political elites cannot use their political power to expropriate the incomes and assets of others, even groups outside the elite may have relatively secure property rights. Therefore, constraints and checks on the use of political power by the elite are typically conducive to the emergence of better economic institutions
2. Second, a similar reasoning implies that economic institutions protecting the rights of a broad cross-section are more likely to arise when political power is in the hands of a relatively broad group containing those with access to the most important investment opportunities. When groups holding political power are narrower, they may protect their own property rights, and this might encourage their own

investments, but the groups outside the political elites are less likely to receive adequate protection for their investments [see Acemoglu (2003b)].

3. Third, good economic institutions are more likely to arise and persist when there are only limited rents that power holders can extract from the rest of society, since such rents would encourage them to opt for a set of economic institutions that make the expropriation of others possible.
4. Finally, considerations related to issues of political losers suggest that institutional reforms that do not threaten the power of incumbents are more likely to succeed. Therefore, institutional changes that do not strengthen strong opposition groups or destabilize the political situation are more likely to be adopted.

6.6. *The colonial experience in light of the comparative statics*

We now briefly return to the colonial experience, and discuss how the comparative statics discussed here shed light on the differences in economic institutions across the former colonies and the institutional reversal.

The second comparative static result above suggests a reason why better economic institutions developed in places where Europeans settled. In these societies, a relatively broad-based group of Europeans came to dominate political power, and they opted for a set of economic institutions protecting their own property rights. In contrast, in places where Europeans did not settle, especially where they were a small minority relative to a large indigenous population, they did not have the incentives to develop good economic institutions because such institutions would have made it considerably more difficult for them to extract resources from the rest of society.

The third comparative static suggests an important reason why in places with more wealth, resources and also a high density of indigenous population to be exploited, Europeans were more likely to opt for worse institutions, without any protection for the majority of the population, again because such institutions facilitated the extraction of resources by the Europeans.

The first comparative static result, in turn, is related to the persistence of the different types of economic institutions that Europeans established, or maintained, in different colonies. In colonies where Europeans settled in large numbers, they also developed political institutions placing effective checks on economic and political elites. In contrast, the political institutions in colonies with high population density, extractive systems of production, and few Europeans, concentrated power in the hands of the elite, and built a state apparatus designed to use coercion against the majority of the population. These different political institutions naturally implied different constraints on political and economic elites. In the former set of colonies, there were constraints on the development of economic institutions that would favor a few at the expense of the majority. Such constraints were entirely absent in the latter set of colonies.

Finally, the fourth comparative static is useful in thinking about why many colonies did not attempt to change their economic institutions during the nineteenth century when new economic opportunities made their previous system based on forced labor, slavery, or tribute-taking much less beneficial relative to one encouraging investment in industry and commerce. Part of the answer appears to lie in the fact that the political power of the elites, for example of the plantation owners in the Caribbean, was intimately linked to the existing economic system. A change in the economic system would turn them into political losers, an outcome they very much wanted to avoid.

6.7. Reassessment of the social conflict view

So far we have shown that the econometric evidence is convincing that differences in economic institutions are the root cause of differences in prosperity. We then argued that although there are different approaches which can account for variation in economic institutions, the most plausible approach is the social conflict view. Though we believe that there certainly are instances where history and ideology matter for the institutional structure of society, and clearly institutions are highly persistent, the most promising approach to understanding why different countries have different institutions is to focus on choices and their subsequent consequences. The social conflict view emphasizes the distributional implication of economic institutions and how commitment problems imply that efficiency and distribution cannot be separated. Hence the fundamental conflict within society over the nature of economic institutions has important implications for economic performance. Some economic institutions will promote growth, but they will not necessarily benefit all groups in society. Alternative economic institutions may induce economic stagnation, but may nevertheless enrich some groups. Which set of institutions results and whether or not a society prospers will be determined by which of these groups has the political power to get the institutions that differentially benefit them. At this point we have therefore substantiated the first three points we made in the introduction. To develop our theory of economic institutions further we need to be more specific about political power – where it comes from and why some people have it and not others. We undertake this task in Section 8. Before doing this however the next section discusses three important historical examples of the evolution of economic institutions. We use these examples to show the explanatory power of the social conflict view and to begin to illustrate in concrete settings how political power works.

7. The social conflict view in action

We now discuss three important examples to bring out the fact that conflict over economic institutions is critical to the functioning of the economy and that this conflict stems, not from differences in beliefs, ideology or historical accidents, but from the impact of economic institutions on distribution. The examples also show that those with political power have a disproportionate effect on economic institutions and they illustrate how the distribution of political power is influenced by different factors. These factors include the allocation of de jure political power through the structure of political institutions and the ability of groups to solve the collective action problem, or exercise what we called de facto political power. With these examples in mind in Section 8 we move to discuss in more detail the nature and sources of political power.

7.1. Labor markets

A market – an opportunity for individuals to exchange a commodity or service – is obviously a fundamental economic institution relevant for development. As Adam Smith (1776) argued, markets allow individuals to take advantage of the benefits of specialization and the division of labor, and scholars such as Pirenne (1937) and Hicks (1969) argued that the expansion of markets was perhaps *the* driving forces in long-run development.

In the history of Europe a key transformation was from feudal labor market institutions towards modern notions of a free labor market where individuals were able to decide who to work for and where to live. This process of institutional change was intimately connected to the transition from a whole set of feudal economic institutions to the economic institutions we think of as 'capitalist'. Most historians see this as key to the economic take-off that began in the nineteenth century. It was the countries which had made the transition away from feudalism most completely, such as England, the Netherlands and France, thanks to the revolution of 1789, which developed most rapidly. It was those where feudalism was still in operation, such as Russia and Austria-Hungary, which lagged far behind.

What can account for this differential evolution of feudalism? Scholars beginning with Postan (1937) saw the demographic collapse caused by the black death in the 1340's as demolishing feudalism in Western Europe. By dramatically altering the land/labor ratio as approximately 40% of the population of Europe died [e.g., Cantor (2001)], the Black Death greatly increased the bargaining power of peasants and allowed them to negotiate a free status ending feudal obligations, particularly with respect to labor. Therefore, Postan's demographic theory implicitly emphasizes the role of political power in the decline of feudalism: this set of economic institutions started to disappear when the political power of the peasants increased and that of lords declined.

In fact, the distribution of power may be even more important in the whole story than Postan's theory suggests. As first pointed out by Brenner (1976), the demographic theory of the decline feudalism is not consistent with the comparative evidence. Although demographic trends were similar all over Europe and

> "it is true that ... in most of Western Europe serfdom was dead by the early sixteenth century. On the other hand, in Eastern Europe, in particular Pomerania, Brandenburg, East Prussia and Poland, decline in population from the late fourteenth century was accompanied by an ultimately successful movement towards imposing extra-economic controls, that is serfdom, over what had been, until then, one of Europe's freest peasantries. By 1500 the same Europe-wide trends had gone a long way towards establishing one of the great divides in European history, the emergence of an almost totally free peasant population in Western Europe, the

debasement of the peasantry to unfreedom in Eastern Europe." [Brenner (1976, p. 41)].

What can explain these divergent outcomes? Brenner notes (p. 51): "It was the logic of the peasant to try to use his apparently improved bargaining position to get his freedom. It was the logic of the landlord to protect his position by reducing the peasants' freedom". The outcome "obviously came down to a question of power" (p. 51); whether the peasants or the lords had more political power determined whether serfdom declined or became stronger.

Although we are far from an understanding of the determinants of the relative structure of political power in different parts of Europe, Brenner suggests that an important element was the "patterns of the development of the contending agrarian classes and their relative strength in the different European societies: their relative levels of internal solidarity, their self-consciousness and organization, and their general political resources – especially their relationships to the non-agricultural classes (in particular, potential urban class allies) and to the state" (p. 52). To substantiate this view, Brenner studies how villages tended to be organized differently in Eastern Europe, there was "more of a tendency to individualistic farming; less developed organization of collaborative agricultural practices at the level of the village or between villages; and little of the tradition of the 'struggle for commons rights' against the lords which was so characteristic of western development" (p. 57). This differential organization was due to the process of initial occupation of these Eastern lands.

Although many parts of Brenner's analysis remain controversial, there is general agreement that the decline of feudalism and the transformation of European labor markets were intimately related to the political power of the key groups with opposing interests, the peasants and the lords [see, for example, Aston and Philpin (1985) on reactions to Brenner's interpretation]. Feudal institutions, by restricting labor mobility and by removing the role of the labor market in allocating labor to jobs, undermined incentives and resulted in underdevelopment. But these same economic institutions created large rents for the aristocracy. As a consequence, aristocracies all over Europe attempted to maintain them. It was when their political power weakened that the process of transformation got underway.

7.2. *Financial markets*

Much recent work on growth and development has focused on capital markets. Growth requires investment, so poor agents without access to financial markets will not have the resources to invest. Empirically many scholars have found correlations between the depth of financial markets and growth [see Levine (2005)] and absence of financial markets is at the heart of ambitious theories of comparative development by Banerjee and Newman (1993) and Galor and Zeira (1993).

If the stress on financial markets and financial intermediation is correct, a central issue is to understand why financial systems differ. For example, studies of the development of banking in the United States in the nineteenth century demonstrate a rapid expansion of financial intermediation which most scholars see as a crucial facilitator of the rapid growth and industrialization that the economy experienced. In his recent study Haber (2001, p. 9) found that in the United States, "In 1818 there were 338 banks in operation, with a total capital of $160 million-roughly three times as many banks and bank capital as in 1810. Circa 1860, the United States had 1,579 banks, with a total capital of $422.5 million. Circa 1914 there were 27,864 banks in the United States. Total bank assets totaled $27.3 billion".

One might see this rapid expansion of banking and financial services as a natural feature. Yet Haber (2001) shows that the situation was very different in Mexico (p. 24). "Mexico had a series of segmented monopolies that were awarded to a group of insiders. The outcome, circa 1910 could not have been more different: the United States had roughly 25,000 banks and a highly competitive market structure; Mexico had 42 banks, two of which controlled 60 percent of total banking assets, and virtually none of which actually competed with another bank."

The explanation for this huge difference is not obvious. The relevant technology was certainly readily available everywhere and it is difficult to see why the various types of moral hazards or adverse selection issues connected with financial intermediation should have limited the expansion of banks in Mexico but not the United States. Haber then shows that (p. 9), "at the time that the U.S. Constitution was put into effect in 1789, ... [U.S. banking] was characterized by a series of segmented monopolies that shared rents with state governments via taxes or state ownership of bank stock. In some cases, banks also shared rents directly with the legislators who regulated them."

This structure, which looked remarkably like that which arose subsequently in Mexico, emerged because state governments had been stripped of revenues by the Constitution. In response, states started banks as a way to generate tax revenues. State governments restricted entry "in order to maximize the amount of rent earned by banks, rent which would then be shared with the state government in the form of dividends, stock distributions, or taxes of various types".

Thus in the early nineteenth century, U.S. banks evolved as monopolies with regulations aimed at maximizing revenues for the state governments. Yet this system did not last because states began competing among themselves for investment and migrants.

> "The pressure to hold population and business in the state was reinforced by a second, related, factor: the broadening of the suffrage. By the 1840s, most states had dropped all property and literacy requirements, and by 1850 virtually all states (with some minor exceptions) had done so. The broadening of the suffrage, however, served to undermine the political coalitions that supported restrictions on the number of bank charters. That is, it created a second source of political competition-competition within states over who would hold office and the policies they would enact."

The situation was very different in Mexico. After 50 years of endemic political instability the country unified under the highly centralized 40 year dictatorship of Porfirio Diaz until the Revolution in 1910.

In Haber's argument political institutions in the United States allocated political power to people who wanted access to credit and loans. As a result they forced state governments to allow free competitive entry into banking. In Mexico political institutions were very different. There were no competing federal states and the suffrage was highly restrictive. As a result the central government granted monopoly rights to banks who restricted credit to maximize profits. The granting of monopolies turned out to be a rational way for the government to raise revenue and redistribute rents to political supporters [see North (1981, Chapter 3)].

A priori, it is possible that the sort of market regulation Haber found in Mexico might have been socially desirable. Markets never function in a vacuum, but rather within sets of rules and regulations which help them to function. Yet it is hard to believe that this argument applies to Mexico [see also Maurer (2002)]. Haber (2001) documents that market regulation was aimed not at solving market failures and it is precisely during this period that the huge economic gap between the United States and Mexico opened up [on which see Coatsworth (1993), Engerman and Sokoloff (1997)]. Indeed, Haber and Maurer (2004) examined in detail how the structure of banking influenced the Mexican textile industry between 1880 and 1913. They showed that only firms with personal contacts with banks were able to get loans. They conclude (p. 5):

> "Our analysis demonstrates that textile mills that were related to banks were less profitable and less technically efficient than their competitors. Nevertheless, access to bank credit allowed them to grow faster, become larger, and survive longer than their more productive competitors. The implication for growth is clear: relatively productive firms lost market share to relatively unproductive (but bank-related) competitors."

Despite the fact that economic efficiency was hurt by regulations, those with the political power were able to sustain these regulations.

7.3. Regulation of prices

As our final example we turn to the regulation of prices in agricultural markets (which is intimately related to the set of agricultural policies adopted by governments). The seminal study of agricultural price regulation in Africa and Latin America is by Bates (1981, 1989, 1997). Bates (1981) demonstrated that poor agricultural performance in Ghana, Nigeria and Zambia was due to government controlled marketing boards systematically paying farmers prices for their crops much below world levels.

> "Most African states possess publicly sanctioned monopsonies for the purchase and export of agricultural goods ... These agencies, bequeathed to the governments of the independent states by their colonial predecessors, purchase cash crops

for export at administratively determined domestic prices, and then sell them at the prevailing world market prices. By using their market power to keep the price paid to the farmer below the price set by the world market, they accumulate funds from the agricultural sector" [Bates (1981, p. 12)].

The marketing boards made surpluses which were given to the government as a form of taxation. Bates (1981, p. 15) notes

> "A major test of the intentions of the newly independent governments occurred ... [when] between 1959–60 and 1961–62, the world price of cocoa fell approximately £50 a ton. If the resources generated by the marketing agencies were to be used to stabilize prices, then surely this was the time to use the funds for that purpose. Instead ... the governments of both Ghana and Nigeria passed on the full burden of the drop in price to the producers."

Bates continues "Using the price setting power of the monopsonistic marketing agencies, the states have therefore made the producers of cash crops a significant part of their tax base, and have taken resources from them without compensation in the form of interest payments or of goods and services returned" (pp. 181–189). As a result of this pernicious taxation, reaching up to 70% of the value of the crop in Ghana in the 1970s, investment in agriculture collapsed as did output of cocoa and other crops. In poor countries with comparative advantage in agriculture such a situation mapped into negative rates of economic growth.

Why were resources extracted in this way? Though part of the motivation was to promote industrialization, the main one is to generate resources that could be either expropriated or redistributed to maintain power

> "governments face a dilemma: urban unrest, which they cannot successfully eradicate through co-optation or repression, poses a serious challenge to their interests ... Their response has been to try to appease urban interests not by offering higher money wages but by advocating policies aimed at reducing the cost of living, and in particular the cost of food. Agricultural policy thus becomes a by-product of political relations between governments and urban constituents." [Bates (1981, p. 33)].

In contrast to the situation in Ghana, Zambia and Nigeria, Bates (1981, 1989, 1997) showed that agricultural policy in Kenya and Colombia over this period was much more pro-farmer. The difference was due to who controlled the marketing board. In Kenya, farmers were not smallholders, as they were in Ghana, Nigeria and Zambia, and concentrated landownership made it much easier to solve the collective action problem. Moreover, farming was important in the Kikuyu areas, an ethnic group closely related to the ruling political party, KANU, under Jomo Kenyatta [Bates (1981, p. 122)]. Farmers in Kenya therefore formed a powerful lobby and were able to guarantee themselves high prices. Even though the government of Kenya engaged in land reform after independence

> "80% of the former white highlands were left intact and ... the government took elaborate measures to preserve the integrity of the large-scale farms ... [which] readily combine in defense of their interests. One of the most important collective efforts is the Kenya National Farmer's Union (KNFU) ... The organization ... is dominated by the large-scale farmers ... [but] it can be argued that the KNFU helps to create a framework of public policies that provides an economic environment favorable to all farmers." [Bates (1981, pp. 93–94)].

Bates concludes (p. 95) that in Kenya

> "large farmers ... have secured public policies that are highly favorable by comparison to those in other nations. Elsewhere the agrarian sector is better blessed by the relative absence of inequality. But is also deprived of the collective benefits which inequality, ironically, can bring."

In Colombia, farmers were favored because of competition for their votes from the two main political parties. Bates (1997, p. 54) notes

> "Being numerous and small, Colombia's coffee producers, like peasants elsewhere, encountered formidable costs of collective action. In most similar instances such difficulties have rendered smallholders politically powerless. And yet ... Colombia's peasants elicited favorable policies from politicians, who at key moments themselves bore the costs of collective action, provisioning the coffee sector with economic institutions and delegating public power to coffee interests."

How could the coffee growers gain such leverage over national policy?

> "A major reason they could do so ... is because the structure of political institutions, and in particular the structure of party competition, rendered them pivotal, giving them the power over the political fortunes of those with ambition for office and enabling them to make or break governments. They thereby gained the power to defeat government officials who sought to orchestrate or constrain their behavior." [Bates (1997, pp. 51, 54)].

A telling piece of evidence in favor of this thesis is that during the 1950s when a civil war broke out between the two parties, there was five years of military rule and policy turned decisively again the coffee growers, only to switch back again with the peaceful resumption of democracy in 1958.

7.4. Political power and economic institutions

These three examples of the creation of economic institutions have certain features in common. All these institutions, labor market regulation/feudalism, the rules governing financial market development, and agricultural price regulation, clearly reflect the outcome of conscious choices. Feudalism did not end in England for incidental or ideological reasons, but because those who were controlled and impoverished by feudal regulations struggled to abolish them. In Eastern Europe the same struggle took place but with a different outcome. Similarly, Mexico did not end up with different financial institutions than the United States by accident, because of different beliefs about what an efficient banking system looked like, or because of some historical factor independent of the outcome. The same is true for differences in economic policies in Kenya and Ghana. Moreover, different sets of economic institutions arising in different places cannot be argued to be efficient adaptations to different environments. Most historians believe that the persistence of feudal institutions in Eastern Europe well into the nineteenth century explains why it lagged far behind Western Europe in economic development. The difference between the financial institutions of Mexico and the United States also plausibly played a role in explaining why they diverged economically in the nineteenth century. The same holds with respect to agricultural price regulation.

The driving force behind all three examples is that economic institutions are chosen for their distributional consequences. Which specific economic institutions emerge depends on who is able to get their way – who has political power. In England, peasant communities had developed relatively strong local political institutions and were able to consolidate on the shock of the Black Death to put an end to feudal regulations. In Eastern Europe it was the lords who had relatively more power and they were able to intensify feudalism in the face of the same demographic shock [as Domar (1970) pointed out, the Black Death actually made serfdom more attractive to the lords even if at the same time it increased the bargaining power of the peasants]. In the case of banking in the nineteenth century, Haber's research shows while the authoritarian regime in Mexico had the political power to freely create monopolies and create rents in the banking industry, the United States was different because it was federal and much more democratic. The political institutions of the United States prevented politicians from appropriating the rents that could flow from the creation of monopolies. Finally, in Bates's analysis, distortionary price regulations arose in Ghana and Zambia, but not in Kenya and Colombia, because in the latter countries agricultural producers had more political power and so could prevent the distortionary policies that would harm their interests.

It is also useful to consider in the context of these examples the mechanisms we discussed in Section 6 which underlie the adoption of inefficient economic institutions. Why could not the peasants and lords of feudal Europe negotiate and allow the introduction of a set of economic institutions that would have given peasants incentives to innovate and would have allowed for the efficient allocation of labor? Why could not either the lords have promised not to expropriate any benefits that accrued from innovation, or alternatively the peasants agreed to compensate the lords if feudal labor institutions were abolished? Though it is difficult to find direct evidence on such counterfactuals from the Medieval period, the most plausible explanation is that such 'deals' were impossible to make credible. The political power of the lords was intimately connected to feudal institutions and thus dismantling these would not only have increased peasant incentives to innovate, but would also have dramatically altered the balance of political power and the distribution of rents in society. Moreover, under feudal regulations peasants were tied to the land. The introduction of free labor mobility would have given workers an exit option, thus increasing their bargaining power with the lords over the division of output. Thus lords might anticipate being both political and economic losers from the ending of feudalism, even if total output would have increased.

In the case of agricultural price regulation, similar arguments are plausible. Cocoa farmers in Ghana would not have believed promises by governments that they would not expropriate the fruits of higher investment, and the governments themselves would not have believed promises by the farmers to compensate them if they left office. Moreover, efficient sets of economic institutions in Ghana or Nigeria would have strengthened the economic base of the rural sector at the expense of the political power of the then dominant urban sector. Indeed, for Ghana in the 1960s, we have direct evidence from the urban economy that the threat of being a political loser led to inefficient economic institutions. This emerges in the analysis of Killick (1978, p. 37) of the attempt by the government of Kwame Nkrumah to promote industrialization. Killick notes:

> "Even had there been the possibility [of creating an indigenous entrepreneurial class] it is doubtful that Nkrumah would have wanted to create such a class, for reasons of ideology and political power. He was very explicit about this saying 'we would be hampering our advance to socialism if we were to encourage the growth of Ghanian private capitalism in our midst'. There is evidence that he also feared the threat that a wealthy class of Ghanaian businessmen might pose to his own political power."

Further evidence on the importance of political loser considerations comes from E. Ayeh-Kumi one of Nkrumah's main economic advisers who noted after the coup that ousted Nkrumah in 1966 that Nkrumah: "informed me that if he permitted African business to grow, it will grow to becoming a rival power to his and the party's prestige, and he would do everything to stop it, which he actually did" [Killick (1978, p. 60)].

In this context, it is interesting that Nkrumah's solution to consolidate his power was to limit the size of businesses that Ghanaians could own. This caused problems for his industrialization policy which he got round by allowing foreign businessmen to enter Ghana. Though this was inconsistent with his aggressively nationalistic and anti-imperialistic rhetoric, these businessmen did not pose a domestic political threat. Killick notes "Given Nkrumah's desire to keep Ghanaian private businesses small, his argument that 'Capital investment must be sought from abroad since there is no bourgeois class amongst us to carry on the necessary investment' was disingenuous" (p. 37). He goes on to add that Nkrumah "had no love of foreign capitalists but he preferred to encourage them rather than local entrepreneurs, whom he wished to restrict" (p. 40).

All these examples show that the distribution of political power in society is crucial for explaining when economic institutions are good and when they are bad. But where does political power come from and who has political power? In addressing these questions we will develop our theory of economic institutions. In a theory based on social conflict where economic institutions are endogenous, it will be to differences in political institutions and the distribution of political power that we must look to explain variation in economic institutions.

8. A theory of institutions

8.1. Sources of political power

Who has political power and where does it come from? As we noted in the Introduction (Section 1.2, point 4), political power comes from two sources. First, an individual or group can be allocated *de jure* power by *political institutions*. But institutions are not the only source of power. A second type of political power accrues to individuals or groups if they can solve the collective action problem, create riots, revolts, or demonstrations, own guns, etc. We call this type of power *de facto* political power [see Acemoglu and Robinson(2003, Chapter 5)].

Actual political power is the composition, the joint outcome, of de jure and de facto power. To see how this works out in practice, consider the situation in Chile in the early 1970's. Salvador Allende was elected President with a plurality of the popular vote. The formal political institutions of democracy in Chile allocated power to him to propose legislation, issue decrees, etc. Consequently, even though he did not have an absolute majority in congress, Allende had a great deal of de jure political power. Political power is not just de jure however; it does not simply stem from political institutions. Allende, despite being empowered under the Chilean Constitution, was overthrown by a military coup in 1973. Here, the military under the leadership of General Pinochet, were able to use brute force and guns to over-ride the formal political institutions. The ability to use force is one example of de facto political power.

As we suggested in the introduction, the relationship between political power and economic and political institutions is complex and dynamic. Consider the example we discussed in Section 7.2, the research by Haber on the comparative financial evolution of Mexico and the United States in the nineteenth century. Haber traced the different evolution of economic institutions to differences in initial political institutions. These political institutions led to different distributions of power and this was critical for the emergence of good financial institutions in the United States, whereby those who benefited from a competitive banking industry were able to force politicians to provide the rules which would guarantee it. But where did these differences in political institutions come from? These differences were partly a result of political events in the nineteenth century, and partially a result of different colonial political institutions. In the United States, during the initial phase of colonization in the early seventeenth century very low population density and lack of easily exploitable resources forced colonizing companies and the British state to make both economic and political concessions; they granted the settlers access to land and accepted the formation of representative democratic institutions [Morgan (1975)]. Consequently, even at independence the United States had relatively democratic political institutions [Keyssar (2000)]. Moreover, the initial egalitarian distribution of assets and the high degree of social mobility made for a situation where, at least in the northern states, the distribution of economic resources, and thus de facto power, was relatively equal. The relatively representative political institutions therefore persisted and were supported by the balance of de facto power in society.

In Mexico there were very different initial conditions during the colonial period with a large indigenous population and rich silver mines to exploit. This led to a much more hierarchical and authoritarian balance of political power and very different colonial economic institutions [see Engerman and Sokoloff (1997)]. These conditions fed into the different institutional structures at independence, the United States with its constitution, checks and balances and federalism, Mexico with its much more centralized, unchecked, unbalanced and absolutist state. These different political institutions then led to very different economic institutions and economic outcomes after independence. Thus, in some ultimate sense, the source of different political institutions were different initial conditions during the colonial period.

Consider now the evidence presented by Bates. Agricultural policies were better in Kenya because large farmers could solve the collective action problem and exercise de facto political power. But the main reason for the existence of large farms was that British settlers expropriated the land from Africans during the expansion of colonialism [see Berman and Lonsdale (1992)]. Thus previous combinations of formal political institutions (colonial institutions) and de facto power (the military might of the British Empire) determined economic institutions, feeding into the future distribution of de facto power even after the nature of de jure power changed dramatically with independence.

We can now see that these examples substantiate the dynamic model that we sketched in Section 1.2. There we showed that at any date, political power is shaped by political institutions, which determine de jure power, and the inherited distribution of resources, which affect the balance of de facto power. Political power then determines economic institutions and economic performance. It also influences the future evolution of political power and prosperity. Economic institutions determine the distribution of resources at that point, which, in turn, influences the distribution of de facto power in the future. Similarly, the distribution of power at any point determines not just the economic institutions then, but also the future political institutions. Thus the allocation of political power at one date, because of the way it influences the distribution of resources and future political institutions, has a crucial effect on the future allocation of both de facto and de jure political power.

Both the comparison Haber made between Mexico and the United States, and that which Bates made between Ghana, Zambia, Kenya and Colombia illustrate this diagram in action. They show how political institutions and de facto power combine to generate different set of economic institutions, how these institutions determine both the distribution of resources and the growth rate of the economy, and how power and institutions evolve over time, often in ways that tend to reinforce particular initial conditions.

8.2. *Political power and political institutions*

The examples we discussed above showed how political power depends on political institutions and de facto power, and how this determines economic institutions. Moreover, we saw that at any time the pre-existing economic institutions will be an important determinant of the distribution of de facto power. The final element to emphasize is how political institutions evolve over time and how they influence the distribution of political power.

To see why political institutions are so important as a source of political power think of a situation where a group, say the Chilean army in the early 1970s, has a great deal of de facto power. Indeed, it has so much de facto power that it can overrule the Chilean Constitution, making the political institutions largely irrelevant. In fact in Chile the de facto power of the military was able to overthrow the legitimate government and completely reverse the economic policies and economic institutions chosen by the Allende government (including land reform and mass nationalization of industry). Not only did the military reverse the economic institutions preferred by Allende and the groups who elected him, they then implemented their own preferred set of economic institutions, in particularly deregulating the trade regime and the economy. Yet the Pinochet regime was heavily concerned with political institutions, and in 1980 Pinochet re-wrote the constitution.

If de facto power was decisive in Chile what is the role for political institutions? If the constitution can be overthrown, why bother to re-write it? The secret to this lies in the intrinsically transitory nature of de facto power.[13] Yes, the military were able to organize a coup in 1973 but this was only because times were uniquely propitious. There was a world-wide economic crisis, and factions of the military that opposed the coup could be marginalized. Moreover, the United States government at the time was happy to encourage and endorse the overthrow of a socialist government, even if it had been democratically elected. The coming together of such circumstances could not be expected to happen continually, hence once Chilean society re-democratized, as it did after 1990, the military would not be able to continually threaten a coup. In response to this Pinochet changed the political institutions in order to attempt to lock in the power of the military, and thus the economic institutions that he/they preferred. Therefore, the important role for political institutions is that they influence the future allocation of political power. This dynamic role is crucial because it explains the major desire of agents to change political institutions when they get the chance – this is how they can attempt to enduringly alter the balance of political power in their favor [see Acemoglu and Robinson (2003)].

[13] The empirical literature on the collective action problem has recognized that the difficulty of solving the collective action problems lead collective action to typically be transitory. Lichbach (1995, p. 17) notes "collective action, if undertaken on a short-term basis, may indeed occur; collective action that requires long periods to time does not... Given that most people's commitments to particular causes face inevitable decline, most dissident groups are ephemeral, most dissident campaigns brief". This transitory nature of collective action is echoed by Tarrow (1991, p. 15) who notes "the exhaustion of mass political involvement", while Ross and Gurr (1989, p. 414) discuss political "burnout". Similarly, Hardin (1995, p. 18) argues that "the extensive political participation of civil society receives enthusiastic expression only in moments of state collapse or great crisis. It cannot be maintained at a perpetually high level."

8.3. A theory of political institutions

We now have in place the outlines of our theory of institutions. There are seven points to emphasize, paralleling the discussion in Section 1.2 and our diagrammatic exposition there. First, individuals have preferences over economic institutions because of the allocation of resources that these institutions induce.

Second, people's preferences typically do not agree because efficiency and distribution cannot be separated. Different economic institutions will benefit different groups, and this will determine the preferences of these individuals and groups with respect to economic institutions.

Third, the problem of commitment explains why efficiency and distribution are inseparable. Economic institutions are collective choices, and they are chosen and sustained by the state. Since there is no third party to enforce the decisions of the state, problems of commitment are particularly severe in the political realm.

Fourth, the equilibrium structure of economic institutions will therefore be determined by who has the power to get their way, i.e., who can create and sustain economic institutions that benefit themselves. The distribution of political power thus determines economic institutions, the allocation of resources and the rate of economic growth.

Fifth, political power has two forms: de jure power determined by the political institutions, such as the constitution and the electoral rules, and de facto power, which stems from the ability to solve the collective action problem, mobilize weapons, etc. De facto power can influence political outcomes independently of the political institutions, and its distribution often critically determines how a given set of institutions works in practice and whether or not they are actually obeyed.

Sixth, the distribution of de facto political power at any date is influenced to a large degree by the distribution of resources in society, since those with greater resources can command more power both through legitimate and intimate means, and perhaps can also solve the collective action problem more efficiently. Naturally, the distribution of resources at this point is influenced by economic institutions and economic outcomes in the past.

Finally, political institutions are also endogenous; the current balance of political power, incorporating both de jure and de facto elements, also determines future political institutions. Political institutions are important because they allocate, at least within the limits defined by the exercise of future de facto power, the allocation of future de jure political power. Since de facto power, because of the nature of the collective action problem, is intrinsically transitory and difficult to wield, political institutions are often crucial in creating a source of durable political power. This makes it very attractive for groups to use their de facto political power to change political institutions so as to modify the distribution of future political power in their favor.

9. The theory in action

We now consider two more examples that further demonstrate our theory of institutions in action. Like the examples discussed in Section 7, these examples contain all the elements of our theory laid out in a skeletal way in Section 1.2. They show the role of political power in determining economic institutions, they demonstrate the different factors, both de facto and de jure, that determine political power, and they illustrate how de facto political power is often used to change political institutions in order to influenced the future distribution of de jure political power.

9.1. Rise of constitutional monarchy and economic growth in early modern Europe

Our first example is the rise of constitutional monarchy in Europe. In the medieval period most European nations were governed by hereditary monarchies. However, as the feudal world changed, various groups struggled to gain political rights and reduce the autocratic powers of monarchies. In England, this process began as early as 1215 when King John was forced by his barons to sign the Magna Carta, a document which increased the powers of the barons, introduced the concept of equality before the law, and forced subsequent kings to consult with them. Many other European nations also developed 'parliaments' which kings could summon to discuss taxation or warfare [see Graves (2001), Ertman (1997)]. Nevertheless, the movement towards limited, constitutional monarchy was not linear or simple. Indeed, in France, certainly from the beginning of Louis XIV's reign in 1638, a more powerful absolutist monarchy appeared with very few controls. Indeed the feudal French parliaments, the Estates General, were not summoned between 1614 and 1788, just before the Revolution.

In England, the Tudor monarchs, particularly Henry VIII and then Elizabeth I, followed by the first Stuart kings, James I and Charles I, also attempted to build an absolutist monarchy. They failed, however, mostly because of Parliament, which blocked attempts to concentrate power. The constitutional outcome in England was settled by the Civil War from 1642–1651 and the Glorious Revolution in 1688. In the first of these conflicts the forces of Parliament defeated those loyal to Charles I and the king was beheaded. In 1660 the monarchy was restored when Charles II became king, but his brother James II was deposed in 1688 and Parliament invited William of Orange to become king.

Other places in Europe, particularly the Netherlands, saw similar developments to those in England. Under the Dukes of Burgundy, the Netherlands had won a considerable amount of political and economic freedom, particularly under the Grand Privilege of 1477 which gave the States General of the Burgundian Netherlands the right to gather on their own initiative and curbed the right of the ruler to raise taxes. However, the Netherlands were inherited by the Hapsburgs through marriage, and by 1493 Maximilian of Hapsburg had reversed the Grand Privilege. After 1552, war with France increased the Hapsburgs' fiscal needs and led them to impose a large tax burden on the Netherlands, already a prosperous agricultural and mercantile area. Growing fiscal and religious resentment in 1572 led to a series of uprisings against the Hapsburgs, mostly orchestrated by commercial interests. These culminated in the War of Independence which was finally won in 1648.

While England and the Netherlands were developing limited constitutional governments, Spain and Portugal were moving in the same direction as France, towards greater absolutism. Davis (1973a, p. 66) notes [in Castille] "the king ruled subject only to weak constitutional restraints. In the first decades of the sixteenth century the crown had reduced the pretensions of the Castillian nobility and towns, so that the representative body, the Cortes, could obstruct but not in the last resort prevent royal tax raising."

These differential institutional trajectories were of enormous consequence. The economies of the Netherlands and England moved ahead of the rest of Europe precisely because these countries developed limited, constitutional government. This form of government led to secure property rights, a favorable investment climate and had rapid multiplier effects on other economic institutions, particularly financial markets [see, e.g., North and Weingast (1989), de Vries and van der Woude (1997)]. While the Netherlands and Britain prospered, France was convulsed by the French Revolution, and by the nineteenth century Spain and Portugal were impoverished backward nations. How can we account for these diverging paths in the early modern period? Why did England and the Netherlands develop limited constitutional rule, while France, Spain and Portugal did not?

We proposed an explanation in Acemoglu, Johnson and Robinson (2005) related to the differential responses of these countries to the opportunities of 'Atlantic trade', that is, overseas trade and colonial activity unleashed by the discovery of the New World and the rounding of the Cape of Good Hope at the end of the fifteenth century. All five nations engaged in Atlantic trade, but they did so in different ways, with very different implications for the organization of society, political institutions and subsequent economic growth.

In England "most trade was carried on by individuals and small partnerships, and not by the Company of Merchant Adventurers, the Levant Company ... or others of their kind" [Davis (1973b, p. 41)]. At least by 1600 there was quite free entry into the English merchant class. The same was true in the Netherlands. In contrast, Cameron (1993, p. 127) describes the Portuguese situation as follows: "The spice trade in the East Indies of the Portuguese Empire was a crown monopoly; the Portuguese navy doubled as a merchant fleet, and all spices had to be sold through the *Casa da India* (India House) in Lisbon ... no commerce existed between Portugal and the East except that organized and controlled by the state". In Spain, similarly, colonial trade was a monopoly of the Crown of Castille, which they delegated to the *Casa de Contratación* (House of Trade) in Seville. This merchants guild was closely monitored by the government [Parry (1966, Chapter 2)]. The main aim of these regulations was to make sure that all of the gold and silver from the Americas flowed back to Spain, creating a source of direct tax revenues for the crown. As a result, Latin American colonies were forbidden to buy manufactured goods from anywhere other than Spain, and all exports and imports had to pass through controlled channels. For example, until the Bourbon reforms of the mid eighteenth century, nothing could be exported directly from Buenos Aires, and if somebody produced anything for export on the Pampas, it had to be carried over the Andes and exported from Lima in Peru!

The source of the differences in the organization of trade, in turn, reflected the different political institutions of these countries. At the time, the granting of trade monopolies was a key fiscal instrument to raise revenues; the more powerful monarchs could increase their revenues by granting trade monopolies or by directly controlling overseas trade, while weaker monarchs could not. At the turn of the fifteenth century, the crown was much stronger in France, Spain and Portugal than in Britain and the Netherlands, and this was the most important factor in the differences in the organization of overseas trade. In fact, when both Tudor and Stuart monarchs attempted to create monopolies similar to those in Spain and Portugal, this was successfully blocked by the English Parliament [see, for example, Hill (1969)]. Consequently, as world trade expanded in the sixteenth and early seventeenth centuries, in England and the Netherlands it enriched merchants engaged in overseas trade, but in France, Spain and Portugal it enriched the crown and groups allied with it. In England and the Netherlands, but not in France, Spain and Portugal, a new class of merchants (and gentry in England) arose with interests directly opposed to those of the Stuarts and the Hapsburgs, and this group was to play a central part in subsequent political changes.

In the case of the Netherlands, de Vries and van der Woude (1997) argue that "urban economic interests ultimately believed it advantageous to escape the Hapsburg imperial framework" (p. 369), and that it was "the traditional pillars of the maritime economy ... that supported and strengthened the young Republic in its hour of need" (p. 366). Moreover, in the case of Amsterdam, "[Hapsburgs'] opponents included most of the city's international merchants ... In 1578 a new Amsterdam city council threw the city's lot in with the Prince of Orange ... among the merchants returning from ... exile were [those whose families] and several generations of their descendants would long dominate the city" (1997, p. 365). The expansion of world trade enriched and expanded precisely those groups within Dutch society most opposed to Hapsburg rule. Israel (1995, pp. 241–242) writes: "From 1590, there was a dramatic improvement in the Republic's economic circumstances. Commerce and shipping expanded enormously, as did the towns. As a result, the financial power of the states rapidly grew, and it was possible to improve the army vastly, both qualitatively, and quantitatively, within a short space of time. The army increased from 20,000 men in 1588 to 32,000 by 1595, and its artillery, methods of transportation, and training were transformed" [see also Israel (1989, Chapter 3)]. By 1629, the Dutch were able to field an army of 77,000 men, 50% larger than the Spanish army of Flanders [Israel (1995, p. 507)]. As a consequence of the Dutch revolt, the Netherlands developed a republican form of government closely attuned to mercantile interests. De Vries and van der Woude (1997, p. 587) describe the new political elite following the Dutch Revolt as: "6 to 8% of urban households with incomes in excess of 1,000 guilders per year. This was the *grote burgerij* from whom was drawn the political and commercial leadership of the country. Here we find, first and foremost, the merchants", and point out how merchants dominated the governments of Leiden, Rotterdam and the cities in two largest states, Zeeland and Holland.

In England, the Civil War and Glorious Revolution coincided with the great expansion of English mercantile groups into the Atlantic. The East India Company was founded in 1600 as the culmination of a series of efforts to develop trade routes with Asia. The 1620s saw the great expansion of tobacco cultivation in Virginia and this was shortly followed by the development of the highly profitable English sugar colonies in the Caribbean. Finally, in the 1650s the English began to take over the Atlantic slave trade. Both the Civil War and the Glorious Revolution were at root battles over the rights and prerogatives of the monarchy. In both cases new merchant interests predominantly sided with those in the gentry demanding restrictions on the powers of the monarchy in order to protect their property and commerce.

The majority of merchants trading with the Americas and in Asia supported Parliament during the Civil War. Brunton and Pennington (1954, p. 62) also note "in the country as a whole there was probably a preponderance of Parliamentarian feeling among merchants". Detailed analyses of the initial members of the Long Parliament in 1640 show that a significant majority of merchants supported the Parliamentarian cause [see Brenner (1973, 1993), Keeler (1954) and Brunton and Pennington (1954)]. Members of the Commons from the City of London (the main center of mercantile activity), as well as many non-London commercial constituencies, such as Southampton, Newcastle and Liverpool, supported Parliament against the King. These men included both professional merchants and aristocrats who invested in colonizing the Americas. These new merchants also provided the financial support needed by Parliament in the difficult early days of the war. They became the customs farmers for the new regime and therefore advanced tens of thousands of pounds that were essential in building up the army [Brenner (1973, p. 82)].

Pincus (1998, 2001, 2002) further documents the critical role of mercantile interests in the Glorious Revolution. He concludes (2002, p. 34) "England's merchant community actively supported William's plan for invasion, and provided a key financial prop to the regime in the critical early months". He notes that James II favored the East India Company and granted various monopoly privileges, alienating the merchant class. Thus, "no wonder the merchant community poured money into William of Orange's coffers in 1688" [Pincus (2002, pp. 32–33)].

The changes in the distribution of political power, political institutions and thus economic institutions that took place in England and the Netherlands had no counterparts in countries with relatively absolutist institutions, like Spain and Portugal, where the crown was able to closely control the expansion of trade. In these countries it was the monarchy and groups allied with it that were the main beneficiaries of the early profits from Atlantic trade, and groups favoring political and economic change did not become strong enough to induce such change. As a result, only in the Netherlands and England did constitutional rule emerge, and only in these two countries were property rights secure. As a result it was these same two countries that prospered.

Why could the monarchies of Spain and Portugal not negotiate a more efficient set of institutions? Alternatively why did the Stuart monarchs in England have to be beheaded or forced from power before better economic institutions could emerge?

It seems quite clear that a change to a more efficient set of institutions in Spain and Portugal would not have been possible under the auspices of the absolutist state, and a reduction in the power of the state was certainly inimical to the interest of the crown. In the case of England, Hill (1961a, p. 22) argues directly that the reason that the Tudor and Stuart monarchs were not in favor of efficient economic institutions is because they feared that this would undermine their political power. He notes:

> "in general the official attitude to industrial advance was hostile, or at best indifferent. It was suspicious of social change and social mobility, the rapid enrichment of capitalists, afraid of the fluctuations of the market and of unemployment, of vagabondage, and social unrest ... the Elizabethan codes aimed at stabilizing the existing class structure, the location of industry and the flow of labor supply by granting privileges and by putting hindrances in the way of the mobility and the freedom of contract."

The account so far explains why a change in the balance of (de facto) political power in England and the Netherlands led to a set of economic institutions favoring the interests of merchants. But in fact much more happened during the seventeenth century; an entirely new set of political institutions, constitutional regimes, restricting the power of the monarchy, were introduced. The reason why the merchants and the gentry in England (and the merchants in the Netherlands) used their newfound powers for political reform illustrates the dynamics of political power emphasized by our theoretical framework.

For example in the case of England, although in 1688 the Parliament might have been strong, it could not be sure that this power would endure. Indeed, the ability to solve the collective action problem and wield de facto power is intrinsically transitory. For instance, the Parliament vanquished James II with the help of a Dutch army, after which they invited William of Orange to take the throne. But how could they anticipate whether or not William would try to assert the absolutist prerogatives that James II had demanded?

The way to make transitory power permanent is to embody it into the rules of the game which is exactly what the English Parliament did after 1688. The changes in institutions after 1688 had large and important effects. For instance, in the eighteenth century the English monarchy was able to borrow huge amounts of money because the fiscal control of Parliament guaranteed that it would not default [see Brewer (1988), Stasavage (2003)]. This borrowing has been seen as crucial to the success of the English war machine. Moreover, with Parliament in control of fiscal policy, the crown was no longer able either to raise money through arbitrary taxation, or to grant monopoly rights in exchange for money – issues which had previously been constant sources of friction between the crown and Parliament. Similarly, after 1688, the greater security of property rights in England led to a huge expansion of financial institutions and markets [Neal

(1990)], which, North and Weingast (1989) argue, laid the institutional foundations for the Industrial Revolution.

Of course the English crown was not without some residual power and might have attempted to mount a coup against Parliament to change political institutions back in its favor. This certainly happened in some places, such as in France after 1849 when Louis Napoleon mounted a successful coup to restore absolutist privileges lost in 1848. Nevertheless, changes in political institutions altered the nature of the status quo in significant ways, and therefore, influenced the future distribution of de jure political power. Political institutions are not cast stone, and they can change, but they still create a source of political power more durable than mere de facto power.

9.2. Summary

The emergence of constitutional rule in some societies of early modern Europe therefore provides a nice example of how economic institutions, which shape economic outcomes, are determined by political power, which is in turn determined by political institutions and the distribution of resources in society. The Netherlands and England prospered in this period because they had good economic institutions, particularly secure property rights and well developed financial markets. They had these economic institutions because their governments were controlled by groups with a strong vested interest in such economic institutions. These groups wielded political power because of the structure of political institutions, i.e., they received de jure power in the Netherlands after the Dutch Revolt and in England after the Civil War and Glorious Revolution.

Moving one step back, we see that political institutions allocated more de jure political power to commercial interests in England and the Netherlands than in France, Spain and Portugal because of major changes in political institutions during the 1600s. These changes took place because commercial interests in England and the Netherlands acquired significant de facto political power as a result of their improving economic fortunes, itself a consequence of the interaction of Atlantic trade and the organization of overseas trade in these countries. Crucially for our framework, these commercial interests used their de facto power to reform (or revolutionize) political institutions so as to acquire de jure political power and solidify their gains.

These events, therefore, illustrate the various elements of our theoretical framework. In particular, they show how it is useful to think of political institutions and the distribution of economic resources as the state variables of the dynamic system, which determine the distribution of political power, and via this channel, economic institutions and economic outcomes. Political institutions and the distribution of economic resources are, themselves, endogenous, determined by political power and economic institutions, as exemplified by the fact that the distribution of economic resources changed significantly during the sixteenth century as a result of the new economic opportunities presented by the rise of Atlantic trade, and these changes were crucially influenced by the existing economic institutions (the organization of overseas trade). Furthermore, the change in the balance of political power led to the changes in political institutions through the English Civil War, the Glorious Revolution and the Dutch Revolt.

9.3. *Rise of electoral democracy in Britain*

Our second example, based on Acemoglu and Robinson (2000a, 2001, 2003), is the rise of mass democracy. In the early nineteenth century, European countries were run by small elites. Most had elected legislatures, often descendants of medieval parliaments, but the franchise was highly restricted to males with relatively large amounts of assets, incomes or wealth. However, as the century and the Industrial Revolution progressed, this political monopoly was challenged by the disenfranchised who engaged in collective action to force political change.

In response to these developments, the elites responded in three ways. First by using the military to repress the opposition, as in the responses to the revolutions of 1848. Second, by making economic concessions to buy off opposition – this is the standard explanation for the beginnings of the welfare state in Germany under Bismarck. Finally, if neither repression nor concessions were attractive or effective, elites expanded the franchise and gave political power to the previously disenfranchised – they created the precedents of modern democracy.

The history of the rise of democracy in Britain is in many ways representative of the experiences of many other European countries. The first important move towards democracy in Britain came with the First Reform Act of 1832. This act removed many of the worst inequities under the old electoral system, in particular the 'rotten boroughs' where several members of parliament were elected by very few voters. The 1832 reform also established the right to vote based uniformly on the basis of property and income. The reform was passed in the context of rising popular discontent at the existing political status quo in Britain.

By the 1820s the Industrial Revolution was well under way and the decade prior to 1832 saw continual rioting and popular unrest. Notable were the Luddite Riots from 1811–1816, the Spa Fields Riots of 1816, the Peterloo Massacre in 1819, and the Swing Riots of 1830 [see Stevenson (1979) for an overview]. Another catalyst for the reforms was the July revolution of 1830 in Paris. Much of this was led and orchestrated by the new middle-class groups who were being created by the spread of industry and the rapid expansion of the British economy. For example, under the pre-1832 system neither Manchester nor Sheffield had any members of the House of Commons.

There is little dissent amongst historians that the motive for the 1832 Reform was to avoid social disturbances [e.g., Lang (1999, p. 36)]. The 1832 Reform Act increased the total electorate from 492,700 to 806,000, which represented about 14.5% of the adult male population. Yet, the majority of British people could not vote, and the elite still had considerable scope for patronage, since 123 constituencies still contained less than 1000 voters. There is also evidence of continued corruption and intimidation of voters until the Ballot Act of 1872 and the Corrupt and Illegal Practices Act of 1883. The Reform Act therefore did not create mass democracy, but rather was designed as a strategic concession. In presenting his electoral reform to the British Parliament in 1831, the Prime Minister Earl Grey was well aware that this was a measure necessary to prevent a likely revolution. He argued:

> "The Principal of my reform is to prevent the necessity for revolution ... reforming to preserve and not to overthrow." [Quoted in Evans (1983, p. 212)].

Unsurprisingly therefore, the issue of parliamentary reform was still very much alive after 1832, and it was taken up centrally by the Chartist movement. But as Lee (1994, p. 137) notes "The House of Commons was largely hostile to reform because, at this stage, it saw no need for it". This had changed by 1867, largely due to a juxtaposition of factors, including the sharp business cycle downturn that caused significant economic hardship and the increased threat of violence. Also significant was the founding of the National Reform Union in 1864 and the Reform League in 1865, and the Hyde Park riots of July 1866 provided the most immediate catalyst.

Lang (1999, p. 75) sums up his discussion by saying "The Hyde Park affair, coupled with other violent outbursts, helped to underscore the idea that it would be better to keep the goodwill of the respectable workers than to alienate them". Reform was initially proposed by the Liberal Prime Minister Russell in 1866 but was defeated by the Conservatives and dissident MP's. As a result Russell's government fell, and the Conservatives formed a minority administration with Lord Derby as their leader in the House of Lords, and Disraeli in charge of the House of Commons. It was Disraeli who then constructed a coalition to pass the Second Reform Act in 1867. As a result of these reforms, the total electorate was expanded from 1.36 million to 2.48 million, and working class voters became the majority in all urban constituencies. The electorate was doubled again by the Third Reform Act of 1884, which extended the same voting regulations that already existed in the boroughs (urban constituencies) to the counties (rural constituencies). The Redistribution Act of 1885 removed many remaining inequalities in the distribution of seats and from this point on Britain only had single member electoral constituencies (previously many constituencies had elected two members – the two candidates who gained the most votes). After 1884 about 60% of adult males were enfranchised. Once again social disorder appears to have been an important factor behind the 1884 act.

In Britain, the Reform Acts of 1867–1884 were a turning point in the history of the British state. Economic institutions also began to change. In 1871 Gladstone reformed the civil service, opening it to public examination, making it meritocratic. Liberal and Conservative governments introduced a considerable amount of labor market legislation, fundamentally changing the nature of industrial relations in favor of workers. During 1906–1914, the Liberal Party, under the leadership of Asquith and Lloyd George, introduced the modern redistributive state into Britain, including health and unemployment insurance, government financed pensions, minimum wages, and a commitment to redistributive taxation. As a result of the fiscal changes, taxes as a proportion of National Product more than doubled in the 30 years following 1870, and then doubled again. In the meantime, the progressivity of the tax system also increased [Lindert (2004)]. Finally, there is also a consensus amongst economic historians that inequality in Britain fell after the 1870's [see Lindert (2000, 2004)].

Meanwhile, the education system, which was either primarily for the elite or run by religious denominations during most of the nineteenth century, was opened up to the masses; the Education Act of 1870 committed the government to the systematic provision of universal education for the first time, and this was made free in 1891. The school leaving age was set at 11 in 1893, then in 1899, it increased to 12 and special provisions for the children of needy families were introduced [Mitch (1993)]. As a result of these changes, the proportion of 10-year olds enrolled in school that stood at 40 percent in 1870 increased to 100 percent in 1900 [Ringer (1979, p. 207)]. Finally, a further act in 1902 led to a large expansion in the resources for schools and introduced the grammar schools which subsequently became the foundation of secondary education in Britain.

Following the Great War, the Representation of the People Act of 1918 gave the vote to all adult males over the age of 21, and women over the wage of 30 who were ratepayers or married to ratepayers. Ultimately, all women received the vote on the same terms as men in 1928. The measures of 1918 were negotiated during the war and may reflect to some extent a quid pro quo between the government and the working classes who were needed to fight and produce munitions. Nevertheless, Garrard (2002, p. 69) notes "most assumed that, if the system was to survive and 'contentment and stability prevail', universal citizenship could not be denied men, perceived to have suffered so much and to have noticed Russia's Revolution".

Overall, the picture which emerges from British political history is clear. Beginning in 1832, when Britain was governed by the relatively rich, primarily rural aristocracy, a series of strategic concessions were made over an 86 year period. These concessions were aimed at incorporating the previously disenfranchised into politics since the alternative was seen to be social unrest, chaos and possibly revolution. The concessions were gradual because in 1832, social peace could be purchased by buying off the middle classes. Moreover, the effect of the concessions was diluted by the specific details of political institutions, particularly the continuing unrepresentative nature of the House of Lords. Although challenged during the 1832 reforms, the House of Lords provided an important bulwark for the wealthy against the potential of radical reforms emanating from a democratized House of Commons. Later, as the working classes reorganized through the Chartist movement and later through trade unions, further concessions had to be made. The Great War and the fallout from it sealed the final offer of full democracy. Though the pressure of the disenfranchised played less of a role in some reforms than others, and other factors undoubtedly played a role, the threat of social disorder was the main driving force behind the creation of democracy in Britain.

The story of the rise of mass democracy that emerges from the British evidence is one where economic and social changes connected with industrialization (for example, rising inequality) and urbanization increased the de facto power of the disenfranchised. In response, they demanded political rights, in particular changes in the political institutions which would allocate future political power to them. These changes in political institutions were, in many ways, the direct cause of the changes in economic institutions, in particular, in the labor market, in government policy, in the educational system, with major distributional implications, including the fall in inequality.

Why did elites in Britain create a democracy? Our discussion makes it clear that democracy did not emerge from the voluntary acts of an enlightened elite. Democracy was, in many ways, forced on the elite, because of the threat of revolution. Nevertheless, democratization was not the only potential outcome in the face of pressure from the disenfranchised, or even in the face of the threat of revolution. Many other countries faced the same pressures and political elites decided to repress the disenfranchised rather than make concessions to them. This happened with regularity in Europe in the nineteenth century, though by the turn of the twentieth century most had accepted that democracy was inevitable. Repression lasted much longer as the favorite response of elites in Latin America, and it is still the preferred option for current political elites in China or Burma.

The problem with repression is that it is costly. Faced with demands for democracy political elites face a trade-off. If they grant democracy, then they lose power over policy and face the prospect of, possibly radical, redistribution. On the other hand, repression risks destroying assets and wealth. In the urbanized environment of nineteenth century Europe (Britain was 70% urbanized at the time of the Second Reform Act), the disenfranchised masses were relatively well organized and therefore difficult to repress. Moreover, industrialization had led to an economy based on physical, and increasing human, capital. Such assets are easily destroyed by repression and conflict, making repression an increasingly costly option for elites. In contrast, in predominantly agrarian societies like many parts of Latin America earlier in the century or current-day Burma, physical and human capital are relatively unimportant and repression is easier and cheaper. Moreover, not only is repression cheaper in such environments, democracy is potentially much worse for the elites because of the prospect of radical land reform. Since physical capital is much harder to redistribute, elites in Western Europe found the prospect of democracy much less threatening.

Faced with the threat of revolt and social chaos, political elites may also attempt to avoid giving away their political power by making concessions, such as income redistribution or other policies that favor non-elites and the disenfranchised. The problem with concessions however is their credibility, particularly when de facto power is transitory. For example, if a crisis, such as a harvest failure or business cycle recession creates a window of opportunity to solve the collective action problem and challenge the existing regime, the elites would like to respond with the promise of policy concessions. Yet windows of opportunity disappear and it is difficult to sustain collective action which entails people protesting in the streets and being away from their families and jobs. Thus collective action quickly dissipates and once it does so, the government has an incentive to renege on its promise of concessions. The promise of concessions, which people know to be non-credible is unlikely to defuse collective action. Hence, Acemoglu and Robinson (2000a, 2001, 2003) argue that democratization occurred as a way of making credible commitments to the disenfranchised. Democratization was a credible commitment to future redistribution, because it reallocated de jure political power away from the elites to the masses. In democracy, the poorer segments of the society would be more powerful and could vote, in other words, could use their de jure political power, to implement economic institutions and policies consistent with their interests. Therefore, democratization was a way of transforming the transitory de facto power of the disenfranchised poor into more durable de jure political power.

9.4. *Summary*

The emergence of mass democracy is another example illustrating our theory of institutions. Into the nineteenth century, economic institutions, particularly in the labor market, disadvantaged the poor. For example, trade unions were illegal and as late as the 1850 in Britain workers trying to organize a union could be shipped to the penal colony in Tasmania, Australia. The poor could not alter economic institutions in their favor because, being disenfranchised, they had little de jure political power, and also limited de facto power because they were often unable to solve their collective action problems.

However, changes in the structure of society and the economy during the early nineteenth century altered the balance of political power, in particular making the exercise of de facto power by the politically disenfranchised much easier [Tilly (1995) and Tarrow (1998) document the changing qualitative nature of collective action over this period]. The rise in the de facto political power of the poor necessitated a change in political institutions in their favor to defuse the threat of revolution. This was to tilt the future allocation of de jure political power, and consequently to ensure future economic institutions and policies consistent with their interests.

Whether or not increases in de facto power translated into democracy depended on a number of factors, in particular how difficult and costly it was for elites to use repression to counter the increase in the power of the masses, and how costly the prospect of democracy was. The changes in political institutions that occurred with democracy had profound implications for economic institutions. In the case of Britain, the period after the Second Reform Act of 1867 led the British state to commit itself to providing universal education and it also led to radical changes in labor market institutions allowing trade unions to form legally for the first time and increasing the bargaining power of labor. Hence economic institutions changed radically in favor of those newly endowed with de jure political power, mostly the relatively poor. This is in fact a relatively general result of democratization. Democracy enfranchises the poor, and the poor are able to use democracy to tilt economic institutions and the distribution of income in society in their favor [Li, Squire and Zou (1998), Rodrik (1999)].

The emergence of democracy in the nineteenth-century Europe therefore also illustrates the workings of our theoretical framework. In particular, it shows how political institutions determine economic institutions and policies, and thus the distribution of resources, and it shows how political institutions change, especially in response to an imbalance of de facto political power, as a credible way of influencing the future allocation of de jure political power.

10. Future avenues

In this chapter we have proposed a framework for thinking about why some countries grow faster and are richer than others. We emphasized, following North and Thomas (1973), that most economic growth theory focuses only on proximate determinants of prosperity. Although this body of work has been useful in helping us understand the mechanics of growth, it fails to provide a satisfactory account of why some countries grow while others do not. Even more recent analyses which have emphasized market institutions and imperfections, and even political economy, have not provided convincing explanations for why countries differ in their equilibrium set of institutions. A major research goal must now be to get beyond the neoclassical growth model and its modern extensions, and search for the deeper causes, i.e., the fundamental determinants of growth.

We argued that the available evidence is consistent with the view that whether or not a society grows depends on how its economy is organized – on its economic institutions. We then proposed the outlines of a theory of institutions and illustrated it through a series of historical examples. We emphasized that a theory of why different countries have different economic institutions must be based on politics, on the structure of political power, and the nature of political institutions. Much remains to be done. First, the framework we outlined was largely verbal rather than mathematical, and thus, by its very nature, not fully specified. Constructing formal models incorporating and extending these ideas is the most important task ahead. Although some of our past work [e.g., Acemoglu and Robinson (2000a, 2001), Acemoglu (2003b)] formalizes parts of this framework, the full model has not been developed yet.

There are also many important issues left out of our framework, which appear to offer fruitful areas for future research. First, though we know that institutions, both economic and political, persist for long periods of time, often centuries (and sometimes millennia), we do not as yet have a satisfactory understanding of the mechanisms through which institutions persist.

Second, and closely related, although institutions do generally persist, sometimes they change. We have important examples of societies which have radically changed their political and economic institutions. Some do so for internal reasons, such as France after the Revolution of 1789, and some do because of external pressures such as Japan after the Meiji restoration or Russia after the Crimean War.

The important point here is that both institutional persistence and institutional change are equilibrium outcomes. Approaches positing institutional persistence as a matter of fact, and then thinking of institutional changes as unusual events will not be satisfactory. Both phenomena have to be analyzed as part of the same dynamic equilibrium framework.

One type of institutional change, consistent with the examples we discussed in this chapter, takes place when those who benefit from the existing set of institutions are forced to accept change, either because they are the losers in a process of fighting or because of the threat of internal revolution (another possibility is that they might accept change because of the threat of external invasion). However, institutional change can also take place because the set of economic institutions that is optimal for a particular group with political power may vary over time as the state variables in the system and economic opportunities evolve. One example may be the end of slavery in the British Empire and another may be the economic and political changes introduced by Mikhail Gorbachev in the Soviet Union in the 1980s. We need more research on the dynamic mechanisms at work [see Tornell (1997) for a model of such a process].

Finally, it is important to understand the role of policy and interventions in changing the institutional equilibrium. Though social science research is of intrinsic interest, one would hope that a convincing fundamental theory of comparative growth based on institutions would lead to policy conclusions that would help us improve the institutions and thus the lives and welfare of people in poor countries. It should be obvious that, at the moment, we are a long way from being in a position to draw such conclusions. In a world where political choices are made rationally and are endogenous to the structure of institutions, which are themselves ultimately endogenous, giving policy advice is a conceptually complex issue [see Acemoglu et al. (2003) for reflections on this]. Recognizing our current ignorance on this topic in no way diminishes its importance, and its role as the Holy Grail of political economy research, however. And we believe that better and empirically more realistic theoretical frameworks in the future will take us closer to this Holy Grail.

Acknowledgements

We thank the editors for their outstanding patience and Philippe Aghion, Robert Barro, Avinash Dixit, Leopoldo Fergusson, Herschel Grossman, Pablo Querubín and Barry Weingast for their helpful suggestions.

References

Acemoglu, D. (1995). "Reward structures and the allocation of talent". European Economic Review 39, 17–33.

Acemoglu, D. (1997). "Training and innovation in an imperfect labor market". Review of Economic Studies 64, 445–464.

Acemoglu, D. (2003a). "Why not a Political Coase Theorem?". Journal of Comparative Economics 31, 620–652.

Acemoglu, D. (2003b). "The form of property rights: Oligarchic versus democratic societies". NBER Working Paper #10037.

Acemoglu, D., Aghion, P., Zilibotti, F. (2002). "Distance to frontier, selection, and economic growth". NBER Working Paper #9066.

Acemoglu, D., Johnson, S., Robinson, J.A. (2001). "The colonial origins of comparative development: An empirical investigation". American Economic Review 91 (5), 1369–1401.

Acemoglu, D., Johnson, S., Robinson, J.A. (2002). "Reversal of fortune: Geography and institutions in the making of the modern world income distribution". Quarterly Journal of Economics 118, 1231–1294.

Acemoglu, D., Johnson, S., Robinson, J.A. (2005). "The rise of Europe: Atlantic trade, institutional change and economic growth." American Economic Review, in press.
Acemoglu, D., Robinson, J.A. (2000a). "Why did the West extend the franchise? Democracy, inequality and growth in historical perspective". Quarterly Journal of Economics 115, 1167–1199.
Acemoglu, D., Robinson, J.A. (2000b). "Political losers as a barrier to economic development". American Economic Review 90, 126–130.
Acemoglu, D., Robinson, J.A. (2001). "A theory of political transitions". American Economic Review 91, 938–963.
Acemoglu, D., Robinson, J.A. (2002). "Economic backwardness in political perspective". NBER Working Paper #8831.
Acemoglu, D., Robinson, J.A. (2003). Economic Origins of Dictatorship and Democracy, forthcoming book to be published by Cambridge University Press.
Acemoglu, D., Zilibotti, F. (1997). "Was Prometheus unbound by chance? Risk, diversification and growth". Journal of Political Economy 105, 709–751.
Acemoglu, D., Johnson, S., Robinson, J.A., Thaicharoen, Y. (2003). "Institutional causes, macroeconomic symptoms: Volatility, crises and growth". Journal of Monetary Economics 50, 49–123.
Ades, A., Verdier, T. (1996). "The rise and fall of elites: A theory of economic development and social polarization in rent-seeking societies". CEPR Discussion Paper #1495.
Aghion, P., Howitt, P.W. (1992). "A model of growth through creative destruction". Econometrica 60, 323–351.
Aghion, P., Howitt, P.W. (1994). "Growth and unemployment". Review of Economic Studies 61, 477–494.
Aghion, P., Howitt, P.W. (2005). "Growth with quality-improving innovations: An integrated framework". In: Aghion, P., Durlauf, S.N. (Eds.), Handbook of Economic Growth, vol. 1A. Elsevier, Amsterdam.
Aghion, P., Harris, C., Howitt, P.W., Vickers, J. (2001). "Competition, imitation, and growth with step-by-step innovation". Review of Economic Studies 68, 467–492.
Alesina, A., Perotti, R. (1996). "Income distribution, political instability and investment". European Economic Review 40, 1203–1225.
Alesina, A., Rodrik, D. (1994). "Distributive politics and economic growth". Quarterly Journal of Economics 109, 465–490.
Allen, R.C. (1982). "The efficiency and distributional consequences of eighteenth century enclosures". Economic Journal 92, 937–953.
Allen, R.C. (1992). Enclosure and the Yeoman. Oxford University Press, New York.
Aston, T.H., Philpin, C.H.E. (Eds.) (1985). The Brenner Debate: Agrarian Class Structure and Economic Development in Pre-industrial Europe. Cambridge University Press, New York.
Azariadis, C., Stachurski, J. (2005). "Poverty traps". In: Aghion, P., Durlauf, S.N. (Eds.), Handbook of Economic Growth, vol. 1A. Elsevier, Amsterdam.
Bairoch, P. (1988). Cities and Economic Development: From the Dawn of History to the Present. University of Chicago Press, Chicago.
Bairoch, P., Batou, J., Chèvre, P. (1988). La Population des villes Europeenees de 800 a 1850: Banque de Données et Analyse Sommaire des Résultats. Centre D'histoire Économique Internationale de l'Uni. de Genève, Libraire Droz, Geneva.
Banerjee, A., Duflo, E. (2005). "Growth theory through the lens of development economics". In: Aghion, P., Durlauf, S.N. (Eds.), Handbook of Economic Growth, vol. 1A. Elsevier, Amsterdam.
Banerjee, A., Newman, A.F. (1993). "Occupational choice and the process of development". Journal of Political Economy 101, 274–298.
Banfield, E.C. (1958). The Moral Basis of a Backward Society. University of Chicago Press, Chicago.
Bardhan, P., Udry, C. (1999). Development Microeconomics. Oxford University Press, New York.
Barro, R.J. (1997). The Determinants of Economic Growth: A Cross-Country Empirical Study. MIT Press, Cambridge.
Barro, R.J., McCleary, R. (2003). "Religion and economic growth". NBER Working Paper #9682.
Barzel, Y. (2001). A Theory of the State: Economic Rights, Legal Rights, and the Scope of the State. Cambridge University Press, New York.

Bates, R.H. (1981). Markets and States in Tropical Africa. University of California Press, Berkeley, CA.
Bates, R.H. (1989). Beyond the Miracle of the Market. Cambridge University Press, New York.
Bates, R.H. (1997). Open Economy Politics. Princeton University Press, Princeton.
Becker, G.S. (1958). "Competition and democracy". Journal of Law and Economics 1, 105–109.
Benhabib, J., Rustichini, A. (1996). "Social conflict and growth". Journal of Economic Growth 1, 125–142.
Berman, B.J., Lonsdale, J. (1992). Unhappy Valley. James Currey, London.
Besley, T.F., Coate, S.T. (1998). "Sources of inefficiency in a representative democracy: A dynamic analysis". American Economic Review 88, 139–156.
Bloom, D.E., Sachs, J.D. (1998). "Geography, demography, and economic growth in Africa". Brookings Papers on Economic Activity 2, 207–295.
Blum, J. (1943). "Transportation and industry in Austria, 1815–1848". Journal of Modern History 15, 24–38.
Bourguignon, F., Verdier, T. (2000). "Oligarchy, democracy, inequality and growth". Journal of Development Economics 62, 285–313.
Brenner, R. (1973). "The Civil War politics of Londons merchant community". Past and Present 58, 53–107.
Brenner, R. (1976). "Agrarian class structure and economic development in preindustrial Europe". Past and Present 70, 30–75.
Brenner, R. (1982). "Agrarian roots of European capitalism". Past and Present 97, 16–113.
Brenner, R. (1993). Merchants and Revolution: Commercial Change, Political Conflict, and London's Overseas Traders, 1550–1653. Princeton University Press, Princeton.
Brewer, J. (1988). The Sinews of Power: War, Money, and the English State, 1688–1783. Harvard University Press, Cambridge.
Brunton, D., Pennington, D.H. (1954). Members of the Long Parliament. Allen and Unwin, London.
Cameron, R. (1993). A Concise Economic History of the World. Oxford University Press, New York.
Cantor, N.F. (2001). In the Wake of the Plague: The Black Death and the World it Made. The Free Press, New York.
Cardosso, F.H., Faletto, E. (1979). Dependency and Development in Latin America. University of California Press, Berkeley, CA.
Cass, D. (1965). "Optimum growth in an aggregate model of capital accumulation". Review of Economic Studies 32, 233–240.
Chandler, T. (1987). Four Thousand Years of Urban Growth: An Historical Census. St. David's University Press, Lewiston, NY.
Charlesworth, A. (1983). An Atlas of Rural Protest in Britain 1548–1900. Croon Helm, London.
Coase, R.H. (1937). "The nature of the firm". Economica 3, 386–405.
Coase, R.H. (1960). "The problem of social cost". Journal of Law and Economics 3, 1–44.
Coatsworth, J.H. (1993). "Notes on the comparative economic history of Latin America and the United States". In: Bernecker, W.L., Werner Tobler, H. (Eds.), Development and Underdevelopment in America: Contrasts in Economic Growth in North and Latin America in Historical Perspective. Walter de Gruyter, New York.
Cumings, B. (2004). North Korea: Another Country. The New Press, New York.
Curtin, P.D. (1989). Death by Migration: Europe's Encounter with the Tropical World in the Nineteenth Century. Cambridge University Press, New York.
Curtin, P.D. (1998). Disease and Empire: The Health of European Troops in the Conquest of Africa. Cambridge University Press, New York.
Davis, R. (1973a). The Rise of the Atlantic Economies. Cornell University Press, Ithaca.
Davis, R. (1973b). English Overseas Trade 1500–1700. Macmillan, London.
Demsetz, H. (1967). "Toward a theory of property right". American Economic Review 57, 61–70.
de Vries, J., van der Woude, A. (1997). The First Modern Economy: Success, Failure, and Perseverance of the Dutch Economy, 1500–1815. Cambridge University Press, New York.
Diamond, J.M. (1997). Guns Germs and Steel: The Fate of Human Societies. W.W. Norton & Co., New York.
Dixit, A.K. (1996). The Making of Economic Policy: A Transaction Cost Politics Perspective. MIT Press, Cambridge.

Dixit, A.K. (2004). Lawlessness and Economics: Alternative Modes of Governance. Princeton University Press, Princeton.

Dixit, A.K., Londregan, J.B. (1995). "Redistributive politics and economic efficiency". American Political Science Review 89, 856–866.

Djankov, S., La Porta, R., Lopez-de-Silanes, F., Shleifer, A. (2002). "The regulation of entry". Quarterly Journal of Economics 117, 1–37.

Djankov, S., La Porta, R., Lopez-de-Silanes, F., Shleifer, A. (2003). "Courts". Quarterly Journal of Economics 118, 453–517.

Dobb, M.H. (1948). Studies in the Development of Capitalism. Cambridge University Press, Cambridge, UK.

Domar, E. (1970). "The causes of slavery or serfdom: A hypothesis". Journal of Economic History 30, 18–32.

Durlauf, S.N. (1993). "Nonergodic economic growth". Review of Economic Studies 60, 349–366.

Durlauf, S.N., Fafchamps, M. (2004). "Empirical studies of social capital: A critical survey". In: Aghion, P., Durlauf, S.N. (Eds.), Handbook of Economic Growth, vol. 1A. Elsevier, Amsterdam.

Edgerton, R.B. (1992). Sick Societies: Challenging the Myth of Primitive Harmony. Free Press, New York.

Eggimann, G. (1999). La Population des Villes des Tiers-Mondes, 1500–1950. Centre d'histoire économique Internationale de l'Uni. de Genève, Libraire Droz, Geneva.

Engerman, S.L., Sokoloff, K.L. (1997). "Factor endowments, institutions, and differential growth paths among new world economies". In: Haber, S. (Ed.), How Latin America Fell Behind. Stanford University Press, Stanford, CA.

Ertman, T. (1997). Birth of the Leviathan: Building States and Regimes in Medieval and Early Modern Europe. Cambridge University Press, New York.

Evans, E.J. (1983). The Forging of the Modern State: Early Industrial Britain, 1783–1870. Longman, New York.

Farrell, J. (1987). "Information and the Coase Theorem". Journal of Economic Perspectives 1, 113–129.

Frank, A.G. (1978). Dependent Accumulation and Underdevelopment. Macmillan, London.

Freudenberger, H. (1967). "State intervention as an obstacle to economic growth in the Hapsburg Monarchy". Journal of Economic History 27, 493–509.

Galor, O., Zeira, J. (1993). "Income distribution and macroeconomics". Review of Economic Studies 40, 35–52.

Garrard, J. (2002). Democratization in Britain: Elites, Civil Society and Reform since 1800. Basingstoke, Palgrave.

Gerschenkron, A. (1970). Europe in the Russian Mirror: Four Lectures in Economic History. Cambridge University Press, Cambridge, UK.

Glaeser, E.L., Shleifer, A. (2002). "Legal origins". Quarterly Journal of Economics 117, 1193–1230.

Graves, M.A.R. (2001). The Parliaments of Early Modern Europe. Longman, New York.

Gregory, P.R. (1991). "The role of the state in promoting economic development: The Russian case and its general implications". In: Sylla, R., Toniolo, G. (Eds.), Patterns of European Industrialization: The Nineteenth Century. Routledge, New York.

Greif, A. (1994). "Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies". Journal of Political Economy 102, 912–950.

Gross, N. (1973). "The industrial revolution in the Hapsburg monarchy, 1750–1914". In: Cipolla, C.M. (Ed.), The Fontana Economic History of Europe, vol. 4. Fontana Books, London.

Grossman, G.M., Helpman, E. (1991). Innovation and Growth in the Global Economy. MIT Press, Cambridge.

Grossman, H., Kim, M. (1995). "Swords or ploughshares? A theory of the security of claims to property rights". Journal of Political Economy 103, 1275–1288.

Grossman, H., Kim, M. (1996). "Predation and accumulation". Journal of Economic Growth 1, 333–350.

Grossman, H., Noh, S.J. (1994). "Proprietary public finance and economic welfare". Journal of Public Economics 53, 187–204.

Grossman, S.J., Hart, O.D. (1986). "The costs and benefits of ownership: A theory of vertical and lateral integration". Journal of Political Economy 94, 691–719.

Gutierrez, H. (1986). "La mortalite des eveques Latino-Americains aux XVIIe et XVIII siecles". Annales de Demographie Historique, 29–39.

Haber, S.H. (2001). "Political institutions and banking systems: Lessons from the economic histories of Mexico and the United States, 1790–1914". Unpublished. Department of Political Science, Stanford University.

Haber, S.H., Maurer, N. (2004). "Related lending and economic performance: Evidence from Mexico". Unpublished. Department of Political Science, Stanford University.

Hall, R.E., Jones, C.I. (1999). "Why do some countries produce so much more output per worker than others?". Quarterly Journal of Economics 114, 83–116.

Hardin, R. (1995). All For One. Princeton University Press, Princeton.

Hart, O.D. (1995). Firms, Contracts, and Financial Structure. Oxford University Press, New York.

Herbst, J.I. (1990). State Politics in Zimbabwe. University of California Press, Berkeley.

Hicks, J.R. (1969). A Theory of Economic History. Oxford University Press, New York.

Hill, C. (1961a). The Century of Revolution, 1603–1714. W.W. Norton & Co., New York.

Hill, C. (1961b). "Protestantism and the rise of capitalism". In: Fisher, F.J. (Ed.), Essays in the Economic and Social History of Tudor and Stuart England. Cambridge University Press, Cambridge.

Hill, C. (1969). From Reformation to Industrial Revolution 1530–1780. Penguin Books, Baltimore.

Hilton, R. (1981). Bond Men Made Free. Routledge, Oxford.

Hirshleifer, J. (2001). The Dark Side of the Force: Economic Foundations of Conflict Theory. Cambridge University Press, New York.

Horowitz, D.L. (1991). A Democratic South Africa? Constitutional Engineering in Divided Societies. University of California Press, Berkeley.

Israel, J.I. (1989). Dutch Primacy in World Trade, 1585–1740. The Clarendon Press, Oxford.

Israel, J.I. (1995). The Dutch Republic: Its Rise, Greatness and Fall 1477–1806. Oxford University Press, New York.

Jones, E.L. (1981). The European Miracle: Environments, Economies, and Geopolitics in the History of Europe and Asia. Cambridge University Press, New York.

Keeler, M. (1954). The Long Parliament, 1640–1641; A Biographical Study of its Members. American Philosophical Society, Philadelphia.

Keyssar, A. (2000). The Right to Vote: The Contested History of Democracy in the United States. Basic Books, New York.

Killick, T. (1978). Development Economics in Action: A Study of Economic Policies in Ghana. Heinemann, London.

Knack, S., Keefer, P. (1995). "Institutions and economic performance: Cross-country tests using alternative measures". Economics and Politics 7, 207–227.

Knack, S., Keefer, P. (1997). "Does social capital have an economic impact? A cross-country investigation". Quarterly Journal of Economics 112, 1252–1288.

Koopmans, T.C. (1965). "On the concept of optimal economic growth". In: The Economic Approach to Development Planning. North-Holland, Amsterdam.

Krugman, P.R., Venables, A.J. (1995). "Globalization and the inequality of nations". Quarterly Journal of Economics 110, 857–880.

Krusell, P., Ríos-Rull, J.-V. (1996). "Vested interests in a theory of stagnation and growth". Review of Economic Studies 63, 301–330.

Krusell, P., Ríos-Rull, J.-V. (1999). "On the size of government: Political economy in the neoclassical growth model". American Economic Review 89, 1156–1181.

Kupperman, K.O. (1993). Providence Island, 1630–1641: The other Puritan Colony. Cambridge University Press, New York.

Kuznets, S. (1968). Towards a Theory of Economic Growth. Yale University Press, New Haven, CT.

Landes, D.S. (1998). The Wealth and Poverty of Nations: Why Some Are so Rich and Some so Poor. Norton & Co, New York.

Lang, S. (1999). Parliamentary Reform, 1785–1928. Routledge, New York.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R. (1998). "Law and finance". Journal of Political Economy 106, 1113–1155.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R. (1999). "The quality of government". Journal of Law, Economics and Organization 15, 222–279.

Lee, S.J. (1994). Aspects of British Political History, 1815–1914. Routledge, New York.

Levine, R. (2005). "Finance and growth: Theory and evidence". In: Aghion, P., Durlauf, S.N. (Eds.), Handbook of Economic Growth, vol. 1A. Elsevier, Amsterdam.

Lewis, W.A. (1978). The Emergence of the International Economic Order. Princeton University Press, Princeton.

Li, H., Squire, L., Zou, H.-f. (1998). "Explaining international and intertemporal variations in income inequality". Economic Journal 108, 26–43.

Lichbach, M.I. (1995). The Rebel's Dilemma. University of Michigan Press, Ann Arbor.

Lindert, P.H. (2000). "Three centuries of inequality in Britain and America". In: Atkinson, A.B., Bourguignon, F. (Eds.), Handbook of Income Distribution. North-Holland, Amsterdam.

Lindert, P.H. (2004). Growing Public: Social Spending and Economics Growth since the Eighteenth Century, vols.1, 2. Cambridge University Press, Cambridge, MA.

Lucas, R.E. (1988). "On the mechanics of economic development". Journal of Monetary Economics 22, 3–42.

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

Marshall, A. (1890). Principles of Economics. Macmillan, London. 1949.

Maurer, N. (2002). The Power and the Profits: The Mexican Financial System, 1876–1932. Stanford University Press, Stanford.

McDaniel, T. (1991). Autocracy, Modernization and Revolution in Russia and Iran. Princeton University Press, Princeton.

McEvedy, C., Jones, R. (1978). Atlas of World Population History. Facts on File, New York.

Mitch, D. (1993). "The role of human capital in the first industrial revolution". In: Mokyr, J. (Ed.), The British Industrial Revolution: An Economic Perspective. Westview Press, San Francisco.

Mokyr, J. (1990). The Lever of Riches: Technological Creativity and Economic Progress. Oxford University Press, New York.

Montesquieu, C.S. (1748). The Spirit of the Laws. Cambridge University Press, New York. 1989.

Moore, B. Jr. (1966). Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World. Beacon Press, Boston.

Morgan, E.S. (1975). American Slavery, American Freedom: The Ordeal of Colonial Virginia. W.W. Norton & Co., New York.

Mosse, W.E. (1958). Alexsandr II and the Modernization of Russia. University of London Press, London, UK.

Mosse, W.E. (1992). An Economic History of Russia, 1856–1914. I.B. Taurus Press, London, UK.

Murphy, K.M., Shleifer, A., Vishny, R.W. (1989a). "Industrialization and the big push". Journal of Political Economy 97, 1003–1026.

Murphy, K.M., Shleifer, A., Vishny, R.W. (1989b). "Income distribution, market size and industrialization". Quarterly Journal of Economics 104, 537–564.

Murphy, K.M., Shleifer, A., Vishny, R.W. (1991). "The allocation of talent: Implications for growth". Quarterly Journal of Economics 106, 503–530.

Myrdal, G. (1968). Asian Drama; An Inquiry into the Poverty of Nations, vols. 1–3. Twentieth Century Fund, New York.

Neal, L. (1990). The Rise of Financial Capitalism: International Capital Markets in the Age of Reason. Cambridge University Press, New York.

Newton, A.P. (1914). The Colonizing Activities of the English Puritans. Yale University Press, New Haven.

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

North, D.C. (1990). Institutions, Institutional Change, and Economic Performance. Cambridge University Press, New York.

North, D.C., Thomas, R.P. (1973). The Rise of the Western World: A New Economic History. Cambridge University Press, Cambridge, UK.

North, D.C., Summerhill, W., Weingast, B.R. (2000). "Order, disorder, and economic change: Latin America versus North America". In: de Mesquita, B.B., Root, H.L. (Eds.), Governing for Prosperity. Yale University Press, New Haven.

North, D.C., Weingast, B.R. (1989). "Constitutions and commitment: Evolution of institutions governing public choice in seventeenth century England". Journal of Economic History 49, 803–832.

Olson, M.C. (1982). The Rise and Decline of Nations: Economic Growth, Stagflation, and Economic Rigidities. Yale University Press, New Haven and London.

Olson, M.C. (1993). "Dictatorship, democracy and development". American Political Science Review 87, 567–575.

Olson, M.C. (2000). Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships. Basic Books, New York.

Overton, M. (1996). Agricultural Revolution in England: The Transformation of the Agrarian Economy 1500–1850. Cambridge University Press, New York.

Parente, S., Prescott, E.C. (1999). "Monopoly rights as barriers to riches". American Economic Review 89, 1216–1233.

Parente, S., Prescott, E.C. (2005). "A unified theory of the evolution of international income levels". In: Aghion, P., Durlauf, S.N. (Eds.), Handbook of Economic Growth, vol. 1B. Elsevier, Amsterdam.

Parry, J.H. (1966). The Spanish Seaborne Empire. University of California Press, Berkeley.

Perotti, R. (1993). "Political equilibrium, income distribution and growth". Review of Economic Studies 61, 755–776.

Persson, T., Tabellini, G. (1994). "Is inequality harmful for growth?". American Economic Review 84, 600–621.

Piketty, T. (1995). "Social mobility and redistributive politics". Quarterly Journal of Economics 100, 551–584.

Pincus, S. (1998). "Neither Machiavellian moment nor possessive individualism: Commercial society an the defenders of the English commonwealth". American Historical Review 103, 705–736.

Pincus, S. (2001). "From holy cause to economic interest: The study of population and the invention of the state". In: Houston, A., Pincus, S. (Eds.), A Nation Transformed: England after the Restoration. Cambridge University Press, New York.

Pincus, S. (2002). "Civic republicanism and political economy in an age of revolution: Law, politics, and economics in the revolution of 1688–89". Unpublished. Department of History, University of Chicago.

Pirenne, H. (1937). Economic and Social History of Medieval Europe. Harcourt, Brace and Company, New York.

Pissarides, C. (2000). Equilibrium Unemployment Theory, second edn. MIT Press, Cambridge.

Postan, M.M. (1937). "The chronology of labour services". Transactions of the Royal Historical Society 20.

Postan, M.M. (1966). "Medieval agrarian society in its prime: England". In: Postan, M.M. (Ed.), The Cambridge Economic History of Europe. Cambridge University Press, London.

Powell, R. (2004). "The inefficient use of power: Costly conflict with complete information". American Political Science Review 98, 231–241.

Putnam, R.D., Leonardi, R., Nanetti, R.Y. (1993). Making Democracy Work: Civic Traditions in Modern Italy. Princeton University Press, Princeton.

Randall, A. (1991). Before the Luddites: Custom, Community, and Machinery in the English Woollen Industry, 1776–1809. Cambridge University Press, New York.

Ray, D. (1998). Development Economics. Princeton University Press, Princeton.

Reynolds, A. (1999). Electoral Systems and Democratization in Southern Africa. Oxford University Press, New York.

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

Robinson, J.A. (1998). "Theories of bad policy". Journal of Policy Reform 3, 1–46.

Rodney, W. (1972). How Europe Underdeveloped Africa. Howard University Press, Washington, DC.

Rodrik, D. (1999). "Democracies pay higher wages". Quarterly Journal of Economics CXIV, 707–738.

Romer, D. (2003). "Misconceptions and political outcomes". Economic Journal 113, 1–20.

Romer, P.M. (1986). "Increasing returns and long-run growth". Journal of Political Economy 94, 1002–1037.
Romer, P.M. (1990). "Endogenous technical change". Journal of Political Economy 98, 71–102.
Rosenstein-Rodan, P. (1943). "Problems of industrialization in Eastern and South-Eastern Europe". Economic Journal 53, 202–211.
Ross, J.I., Gurr, T.R. (1989). "Why terrorism subsides: A comparative study of Canada and the United States". Comparative Politics 21, 405–426.
Sachs, J.D. (2000). "Notes on a new sociology of economic development". In: Harrison, L.E., Huntington, S.P. (Eds.), Culture Matters: How Values Shape Human Progress. Basic Books, New York.
Sachs, J.D. (2001). "Tropical underdevelopment". NBER Working Paper #8119.
Saint-Paul, G., Verdier, T. (1993). "Education, democracy, and growth". Journal of Development Economics 42, 399–407.
Schattschneider, E.E. (1935). Politics, Pressures and the Tariff; a Study of Free Private Enterprise in Pressure Politics, as Shown in the 1929–1930 Revision of the Tariff. Prentice-Hall, New York.
Scott, J.C. (2000). "The moral economy as an argument and as a fight". In: Randall, A., Charlesworth, A. (Eds.), Moral Economy and Popular Protest: Crowds, Conflict and Authority. MacMillan, London.
Skaperdas, S. (1992). "Cooperation, conflict, and power in the absence of property rights". American Economic Review 82, 720–739.
Smith, A. (1776). The Wealth of Nations, vols.1, 2. Penguin Classics, London. 1999.
Solow, R.M. (1956). "A contribution to the theory of economic growth". Quarterly Journal of Economics 70, 65–94.
Stasavage, D. (2003). Public Debt and the Birth of the Democratic State: France and Great Britain, 1688–1789. Cambridge University Press, New York.
Stevenson, J. (1979). Popular Disturbances in England, 1700–1870. Longman, New York.
Tarrow, S. (1991). "Aiming at a moving target: Social science and the recent rebellions in Eastern Europe". PS: Political Science and Politics 24, 12–20.
Tarrow, S. (1998). Power in Movement: Social Movements and Contentious Politics, second edn. Cambridge University Press, New York.
Tawney, R.H. (1926). Religion and the Rise of Capitalism: A Historical Study. J. Murray, London.
Tawney, R.H. (1941). "The rise of the gentry, 1558–1640". Economic History Review 11, 1–38.
Thomis, M.I. (1970). The Luddites; Machine-Breaking in Regency England. David & Charles, Newton Abbot.
Tilly, C. (1995). Popular Contention in Britain, 1758–1834. Harvard University Press, Cambridge.
Tornell, A. (1997). "Economic growth and decline with endogenous property rights". Journal of Economic Growth 2, 219–250.
Tornell, A., Velasco, A. (1992). "Why does capital flow from poor to rich countries? The tragedy of the commons and economic growth". Journal of Political Economy 100, 1208–1231.
Townsend, R.M. (1993). The Medieval Village Economy: A Study of the Pareto Mapping in General Equilibrium Models. Princeton University Press, Princeton.
Vargas Llosa, M. (1989). The Storyteller. Farrar, Straus, Giroux, New York.
Veitch, J.M. (1986). "Repudiations and confiscations by the medieval state". Journal of Economic History 46, 31–36.
Véliz, C. (1994). The New World of the Gothic Fox: Culture and Economy in English and Spanish America. University of California Press, Berkeley.
Wallerstein, I.M. (1974–1980). The Modern World-System, vols. 1–3. Academic Press, New York.
Weber, M. (1930). The Protestant Ethic and the Spirit of Capitalism. Allen and Unwin, London.
Weber, M. (1958). The Religion of India. Free Press, Glencoe.
Weingast, B.R. (1997). "The political foundations of democracy and the rule of law". American Political Science Review 91, 245–263.
Weingast, B.R. (1998). "Political stability and Civil War: Institutions, commitment, and American democracy". In: Bates, R., Greif, A., Levi, M., Rosenthal, J.-L., Weingast, B.R. (Eds.), Analytic Narratives. Princeton University Press, Princeton.
Wiarda, H.J. (2001). The Soul of Latin America: The Cultural and Political Tradition. Yale University Press, New Haven.

Williams, E.E. (1944). Capitalism and Slavery. University of North Carolina Press, Chapel Hill.
Williamson, O. (1985). The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting. Free Press, New York.
Wittman, D. (1989). "Why democracies produce efficient results". Journal of Political Economy 97, 1395–1424.