
Paths of Institutional Development: A View from Economic History

Karla Hoff

This article surveys an influential new research program on historical paths of institutional development and their consequences for growth. The research program exploits the experience of European colonialism as a kind of “natural experiment” whose results bear on the way institutions affect development. The central hypothesis of this research is that societies that began with more extreme inequality were more likely to develop institutions allowing much of the population only limited access to economic opportunities. The research has uncovered a striking reversal of fortune among the areas colonized by Europe; those that were relatively rich in the 1600s are today far poorer than the areas (such as the United States and Canada) that initially were viewed as relatively undesirable. The timing of the reversal—at the onset of the Industrial Revolution, when there was probably a premium on broad participation in commercial activity—suggests that institutions associated with high inequality may be a causal factor in low aggregate incomes. This research program is still in its early stages. But studies of institutions in India using data rich enough to permit hypothesis-testing provide evidence supporting the hypotheses developed in the analysis of the European colonial experience.

Economists have for some time recognized that institutions—the “rules of the game” that shape incentives and opportunities—are a key determinant of the wealth or poverty of nations.¹ But they are just beginning to understand the process of *institutional change*. In earlier work I surveyed recent developments that led to a consensus that the traditional neoclassical model, because it abstracts from institutional change, leaves out the heart of development economics (Hoff and Stiglitz 2001; Hoff 2001). This article surveys an influential new research program on historical paths of institutional development and their consequences for growth.

Much of the work discussed here is concerned with two puzzling facts about former European colonies (Engerman and Sokoloff 1997, 2002; Sokoloff and Engerman

2000; Acemoglu, Johnson, and Robinson 2002). First, many European colonies viewed by forecasters and migrants of the time as offering the best prospects for wealth are among the poorest countries in the world today. Second, the point in time at which these colonies fell behind was not during the early period of colonialism but at the onset of the Industrial Revolution. In simplest terms, the thesis put forward persuasively in recent work is that the factors that made these European colonies relatively wealthy in 1500, 1600, or 1700 also made possible a colonization strategy that created or perpetuated stark inequalities in wealth and political power—a strategy whose legacy is institutions that made these areas ill-suited for modern economic growth.

In 1700, Mexico and the colonies that were to become the United States had a very similar per capita income (based on approximate estimates), and the sugar-producing islands of Barbados and Cuba were far richer (table 1). In the 16th, 17th, and 18th centuries, the North American mainland was widely considered to offer relatively poor economic prospects compared with the vast opportunities in the Caribbean and Latin America. The colonial powers viewed Canada—famously characterized by Voltaire as “a few acres of snow”—to be of comparable value to the small sugar-producing island of Guadeloupe. But the United States and Canada ultimately proved to be far richer than other economies of the hemisphere. All Latin American countries in Table 1 fell behind the United States in per capita income in the 19th century, and Argentina, Mexico, and Peru fell still further behind in the 20th century.

In the modern period at least, few indicators are as strongly correlated with a country’s level of development as its level of urbanization (see, for example, Glaeser 1999). Moreover, longer time series are available for urbanization than for per capita

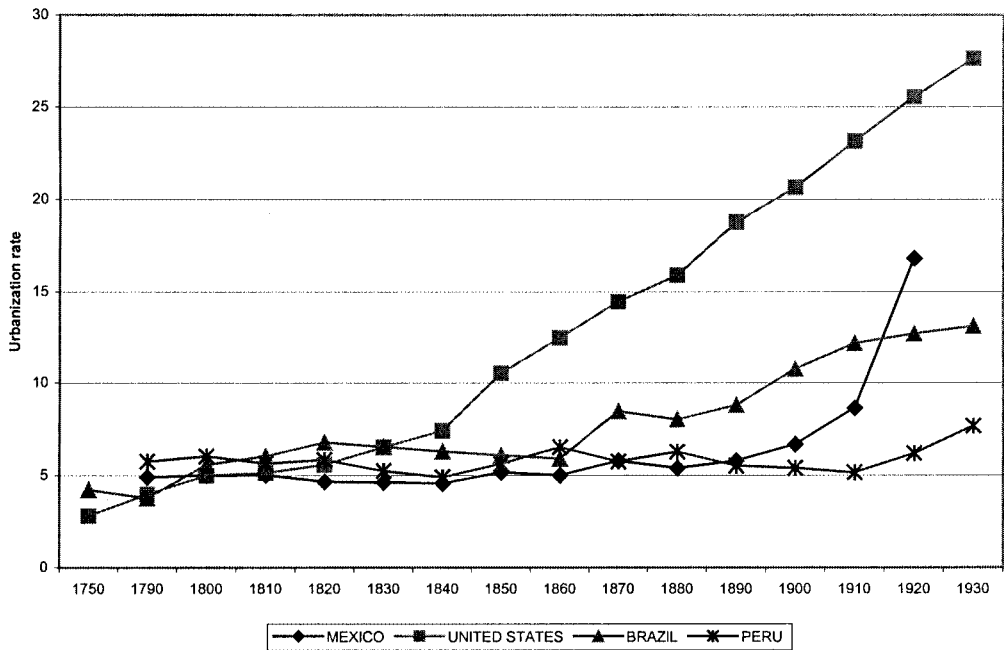
Table 1. Record of Per Capita Gross Domestic Product in Selected Economies of the Americas, Selected Years, 1700–1997

<i>Economy</i>	<i>Per capita GDP as a percentage of U.S. value</i>			
	<i>1700</i>	<i>1800</i>	<i>1900</i>	<i>1997</i>
Argentina	—	102	52	35
Barbados	150	—	—	51
Brazil	—	50	10	22
Chile	—	46	38	42
Cuba	167	112	—	—
Mexico	89	50	35	28
Peru	—	41	20	15
Canada	—	—	67	76
U.S. per capita GDP (1985 U.S. dollars)	550	807	3,859	20,230

Note: — Not available.

Source: Sokoloff and Engerman (2000).

Figure 1. Urbanization Rates in Four New World Economies, 1750–1930



Note: Urbanization rate is the percent of population living in urban areas with a population of at least 20,000.
Source: Acemoglu, Johnson, and Robinson (2001b). Reprinted with permission.

income. If urbanization is taken as a proxy for development, the U.S. economy diverged sharply from the rest of the hemisphere in the mid-19th century, just as the United States began to industrialize (figure 1). Per capita incomes throughout the Americas have increased greatly since 1900, but the gap between the United States and Canada, on the one hand, and Latin America, on the other, has changed little in proportional terms since 1900 (see table 1).

Why did the areas favored by the forecasters of the 1700s fall behind? Clearly development is not just about *having* productive opportunities. If that were the case, it would be hard to explain why all the Latin American colonies, which were generally richer in 1700 than the areas to the north, fell behind beginning in the 18th century. More than *having* productive opportunities, most economists would agree that development depends on being able to *create* a never-ending supply of new opportunities in the future. Economic historians Stanley Engerman and Kenneth Sokoloff (1997, 2002; Sokoloff and Engerman 2000) have argued that the reason that incomes in the United States and Canada diverged from those in the rest of the hemisphere was that a key to successful early industrialization was the *breadth of access to opportunities for social and economic advancement*—the ability to own land,

obtain schooling, borrow, and innovate. Among the colonies of the Americas, only the United States and Canada provided the social infrastructure—the collection of laws, institutions, and government policies (Hall and Jones 1997, 1999)—that made participation in investment and entrepreneurship possible for a broad segment of the population. At the end of the 18th century, the Industrial Revolution generated opportunities whose value depended on broad participation in entrepreneurship, investment, and innovation. Thus a “reversal of fortune” occurred as the United States and Canada surged ahead of societies in which a large fraction of the population was illiterate, disenfranchised, and without assets to borrow against.²

This article first gives an overview of the surprising systematic patterns in the historical paths of institutional development in New World economies and evidence that differences in the choices of the colonial powers relating to the institutions they set up have had a powerful effect on the performance of these economies in the past two centuries. Next it discusses two recent econometric studies that pinpoint the causes and consequences of inefficient institutions of particular types in rural India. Finally, it uses the example of patent institutions to illustrate the possibility that, for institutions, the devil is in the details—the fine details of an institution can make the difference between broad and narrow access to opportunity.

I. Paths of Institutional Development in the New World

The most systematic studies of paths of institutional development have focused on the Americas (see Engerman and Sokoloff 1997, 2002; Sokoloff and Engerman 2000; Acemoglu, Johnson, and Robinson 2002; and Easterly 1999, 2000). The patterns that this work discerns over the past 400 years point to two broad conclusions.

- European colonialism was a historical moment when factor endowments imposed constraints on the choices of a small number of colonial powers with respect to the institutions they set up. Factor endowments, broadly defined, were major influences on the level of economic and political inequality in the early days of colonial rule.
- The institutions established at the outset of colonial rule then influenced policies that shaped factor endowments (levels of physical and human capital) and institutional trajectories in ways that tended to reproduce the initial level of political and economic inequality. In colonies characterized by high initial inequality, the majority of people could not vote, own land, or obtain an education for hundreds of years. Nor, as Acemoglu, Johnson, and Robinson (2002) emphasize, could they enjoy effective protection from expropriation of their property.

Three Kinds of Factor Endowments

The central aspect of factor endowments that shaped initial inequality in the European colonies was labor scarcity.³ The extent of labor scarcity depended not only on the size of the population but also on its density (which influenced how easily the native population could be enslaved) and the local climate and soils (which influenced the profitability of importing slaves). Because there were no prohibitions against slavery in any of the regions colonized by the Europeans, areas where the climate and soils were suitable for plantation agriculture were those where a slave labor force was imported.⁴

Engerman and Sokoloff (1997) distinguish three broad classes of colonies by their factor endowments during the early period of colonization (1500–1650):

- *Dense populations of natives and rich supplies of minerals.* Many parts of Spanish America had large concentrations of populations of Native American descent that survived contact with the colonizers—and rich concentrations of minerals. Spain adopted a policy throughout Spanish America of distributing vast grants of land, including claims to the labor of the native population residing on the land. Spain also limited the immigration of European settlers, which contributed to the persistence of elites and the maintenance of vast landholdings even where production activities (as in Argentina) were not characterized by economies of scale.
- *Climate and soils well-suited to producing sugar and other highly valued crops cultivated with slave labor.* In the Caribbean, factor endowments were suitable for plantations based on slave labor, and scale economies supported the competitive success of large-scale sugar plantations. The reliance on slavery and the disparity in landownership made the distribution of wealth, income, and human capital extremely unequal, not only among the population as a whole but even among free men.
- *Dispersed indigenous population and climate and soils suited for grains and livestock.* In North America the preexisting population of Native American descent was very sparse. Except in the South, the colonies had soils and climates that made the production of crops using slave labor unprofitable. As a result, development depended on labor of European descent. Such labor was scarce because the North American colonies were not viewed as attractive places to settle and because grains could be profitably produced on very small farms (see Engerman and Sokoloff 2002:table 2).

Thus for different reasons the colonies in the first two categories were characterized almost from the outset by an abundance of slave or quasi-enslaved labor with very low human capital. Those in the third category were not. These factor endowments shaped the possible choices of the colonial powers with respect to the institutions

they set up. To concentrate power and wealth in the hands of a few and to extract labor and tribute from the rest proved possible and profitable in the first two classes of environments but not in the third.

The French in Canada and the British in North America began many settlements with an attempt to introduce a seigniorial system for landholdings. But because of the labor scarcity, the system either failed to take hold or collapsed after a short period (Galenson 1996). The economic historian David Galenson (1996:144) emphasizes the pivotal role of labor scarcity in the economic organization of the early colonies:

The extreme labor shortage...allowed many early settlers to gain their economic independence from the manorial lords, and establish separate farms... Although the establishment of large estates to be worked by tenants and landless laborers was the initial model on which these proprietary colonies [Maryland, Jamestown, Carolina, New Jersey, and New York] were usually based, *the greater economic power conferred on settlers by the New World's labor scarcity prevented these English tenures and practices from effectively taking hold*, and proprietors were often forced to adapt by simply selling their land outright to settlers. (emphasis added)

In the Southern colonies of what became the United States, conditions were somewhat different. The Southern climate was suitable for crops—rice, cotton, and tobacco—that exhibited scale economies. But even here the slave population and the degree of inequality were much smaller than those in Latin America, in part because the South was unsuitable for sugar production and in part because institutions in the South were determined at the national level or by competition among states.

All the economies established in the Americas had abundant land relative to labor. But their climates, soils, minerals, and density of native populations differed in ways that meant that most regions were characterized almost from the outset of colonization by extreme inequality, whereas the colonies that were to become the United States and Canada were not.

Paths of Institutional Development

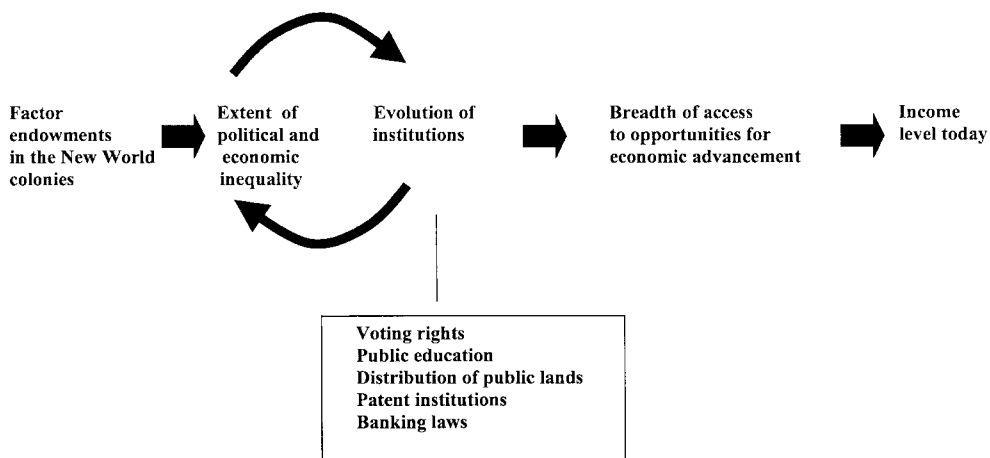
Why should inequality hundreds of years ago matter for per capita income today? What are the mechanisms that link the historical factor endowments to social organization and economic performance today? Recent and ongoing research has established systematic patterns in the institutional development paths of New World economies: In societies with high inequality at the outset of colonization, institutions tended to evolve in ways that restricted to a narrow elite access to political power and opportunities for economic advancement (Engerman and Sokoloff 1997, 2001, 2002; Engerman, Mariscal, and Sokoloff 1999).

Engerman and Sokoloff have discerned such a pattern in a wide range of public policies: the right to vote and to vote in secret, schooling, the distribution of public land and other natural resources, patent institutions, and banking laws. Moreover, they argue that high inequality at the outset of colonization through its effects on institutions may provide an important part of the explanation for the divergence in per capita income among New World economies. Their central thesis is worth quoting at some length (see figure 2 for a schematic representation of this thesis):

The factor endowment and the degree of inequality may influence the directions in which institutions evolve, but these institutions, in turn, can affect the evolution of the factor endowment and of the distributions of human capital, wealth, and political power. The initial conditions had long-lasting effects, however, not only because they were difficult to change, but also because government policies and other institutions tended generally to foster their persistence. (Engerman and Sokoloff 2002:63–64)

Voting Rights. One mechanism through which initial disparities in wealth and political power might be reproduced over time is limitations on suffrage and on the right to vote in secret. Most societies in the Americas were democracies by the middle of the 19th century, but they differed sharply in the breadth of effective access to the vote. Until early in the 19th century all countries, including the United States, limited the right to vote to white men with significant property. But by the mid-19th century, the share of the population voting in the United States and Canada was far greater

Figure 2. Institutional Origins and Persistence in New World Economies



Source: Author's illustration.

than that in Argentina, Brazil, Chile, and Ecuador (table 2). At the beginning of the 20th century the share voting in the United States and Canada was more than twice that in Argentina, one of the most progressive Latin American countries at the time.

In the United States, labor scarcity played a role in broadening voting rights. This is evidenced by the fact that the pioneers in extending the franchise were Western states competing for migrants (Engerman and Sokoloff 2001, 2002). Nearly all new entrants to the union in the 19th century extended voting rights to all white men.

Public education. Most Latin American countries did not provide public primary schooling on a scale sufficient to serve the general population until the 20th century. In contrast, by the early 19th century locally funded public primary schooling was widespread in the United States and Canada (although the southern United States,

Table 2. Voting and Literacy Rates in Selected Economies of the Americas, Various Years, 1850–1925

<i>Period and country^a</i>	<i>Share of the population voting (%)</i>	<i>Literacy rate (%)^b</i>
<i>1850–1900</i>		
Argentina, 1896, 1869	1.8	23.8 ^c
Brazil, 1894, 1872	2.2	15.8 ^d
Chile, 1869, 1865	1.6	18.0 ^d
Ecuador, 1856	0.1	—
Canada, 1867, 1861	7.7	82.5 ^e
United States, 1850, 1870	12.9	80.0
<i>1900–25</i>		
Argentina, 1916, 1925	9.0	73.0
Brazil, 1914, 1920	2.2	30.0
Chile, 1920, 1925	4.4	66.0
Colombia, 1918	—	32.0 ^f
Mexico, 1920, 1925	8.6	36.0
Canada, 1911	18.1	—
United States, 1900, 1910	18.4	92.3

Note: — Not available.

^aThe first year in each row relates to voting and the second to literacy. Where only one year appears, it relates to the indicator for which data are available.

^bExcept where otherwise specified, the literacy rate refers to the age group 10 and above.

^cAge 6 and above.

^dAge 7 and above.

^eAll.

^fAge 15 and above.

Source: Engerman and Sokoloff (2002:tables 7 and 8).

which had greater inequality and greater heterogeneity in the population, lagged behind). As a result, around 1870 the United States and Canada had literacy rates roughly four times those in many Latin American countries.

When literacy rates first diverged, the differences in per capita income between the United States and Canada and the rest of the hemisphere were small. So differences in aggregate resources to invest in schooling do not explain this pattern.

An alternative explanation focuses on political and economic inequality. If only the wealthy have the right to vote, all voters face tradeoffs. Providing mass education would raise the productive potential of the poor majority, which might promote growth and create positive spillovers for the wealthy. But the wealthy would bear a disproportionate share of the cost of that education, and the newly educated poor might agitate for the right to vote, which would threaten the political power of the elite. Among the New World economies, those with greater political inequality (as reflected in the share of the population voting) had a smaller share of the population enrolled in school, controlling for per capita income, time, and region (Engerman, Mariscal, and Sokoloff 1997). Among all economies today, those with greater income inequality tend to have lower spending on schooling, particularly primary schooling, and tend to be less democratic.⁵

Distribution of public lands. The governments of colonies or nations were regarded as the owners of the extensive public lands that existed into the 19th century and beyond. Public policies for transferring these lands to private hands were often the subject of political debates and struggles. The outcomes in the United States and Canada differed greatly from those in Latin America. The two North American countries awarded small landholdings to people who would settle and farm the land for a specified period. In contrast, Argentina and Mexico tended to award large landholdings to developers.

This policy difference led to extreme differences in the degree of inequality in rural landownership in these four countries at the beginning of the 20th century (table 3). In the two Latin American countries a small minority of households owned all the land—19 percent in Argentina and 2 percent in Mexico. In contrast, in Canada 87 percent of heads of household in rural areas owned land, and in the United States, 75 percent did. Throughout much of U.S. and Canadian history the vast majority of occupiers of farmland were landowners (Engerman and Sokoloff 2002).

Thus large differences existed between the United States and Canada, on the one hand, and the rest of the Americas, on the other, in public policies toward suffrage, education, and land. Based on these and similar patterns, Engerman and Sokoloff (1997, 2002) argue that initial differences in the degree of inequality—which can be attributed largely to factor endowments—had long-lasting effects on the paths of development of the economies of the Americas. In economies characterized by stark inequality at the outset of colonization, institutions evolved in ways that tended to

Table 3. Rural Landholding in Argentina, Mexico, Canada, and the United States, around 1900

<i>Country and year</i>	<i>Share of heads of household owning land (%)</i>
Argentina, 1895	19.2 ^a
Mexico, 1910	2.4
Canada, 1901	87.1
United States, 1900	74.5

Note: Landownership is defined as follows: in Argentina, the share of landowners in the male population ages 18–50; in Mexico, heads of household who own land; in Canada, occupiers of farmland who are owners; and in the United States, farms that are owner-operated.

^aThe average across the nine states of Argentina for which data are available.

Source: Engerman and Sokoloff 2002, table 6.

protect the elite and maintain a large class of poor, uneducated, and disenfranchised people. Neither the abolition of slavery nor the creation of democratic governments changed the basic patterns.⁶

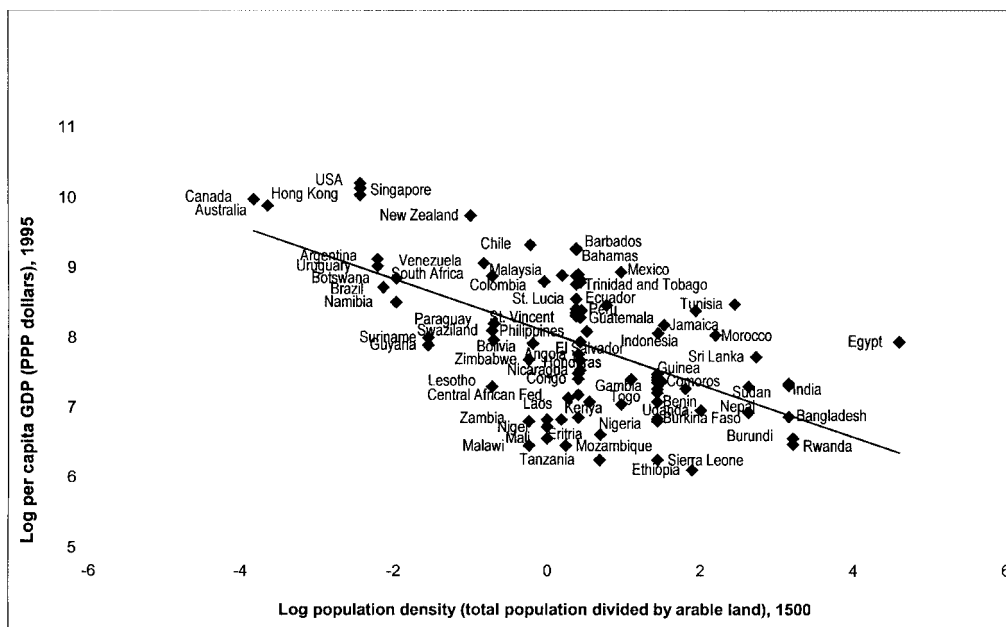
An Empirical Test

The work of Engerman and Sokoloff and Acemoglu, Johnson, and Robinson lays out a provocative theory: that the differences between the institutions that the European colonizers put in place (or took over) at the outset of colonization account for the reversal in relative incomes between 1500 or 1700 and the present day. A testable implication of the theory is that areas characterized by high labor density in 1500 should be relatively poor today. Figure 3 shows that this is indeed the case. There is a strongly negative relationship between population density in 1500 and per capita income today. The approximately linear relationship between these two variables, expressed in logarithms, means that there is a roughly constant relationship between a 1 percent increase in population density in 1500 and a 1 percent decline in per capita income today.

This statistical relationship is robust, remaining essentially the same for subsamples of colonies that are on the same continent, were colonized by the same country, are at the same distance from the equator, and exclude Australia, Canada, New Zealand, and the United States. Acemoglu, Johnson, and Robinson (2001b:2) offer the following explanation:

Relatively poor areas were sparsely settled, and this enabled or induced Europeans to develop settler colonies with institutions encouraging investment and commerce by a broad cross-section of the society. In contrast, a large population and relative prosperity made it profitable for the Europeans to

Figure 3. Population Density in 1500 and Per Capita Income in 1995 for Former European Colonies



Source: Acemoglu, Johnson, and Robinson (2002:figure 2). Reprinted with permission of MIT Press Journals.

set up extractive institutions, with political power concentrated in the hands of a small elite. High population density, for example, meant a large supply of labor that the Europeans could force to work in mines or plantations, or tax heavily by taking over existing tribute systems.

Acemoglu, Johnson, and Robinson (2002) go on to demonstrate that the economic growth of the initially sparsely settled colonies diverged sharply from that of the initially densely settled colonies at the onset of the Industrial Revolution (see figure 1, which illustrates the relationship for four economies of the Americas). This evidence—the “second fact” about the reversal of fortune—is consistent with the view that it was the interaction of national institutions and opportunities for industrialization during the 19th century that caused the divergence of incomes between the initially sparsely settled economies and those that were more densely populated at the outset of colonization.

The historical findings reported here are sharply at variance with two traditional views of economic development. One is the functionalist view that institutions are endogenous, flexible, and efficient. In this view, efficient institutions will “naturally”

emerge over time.⁷ Key institutions for the success of capitalist economies are secure property rights and the rule of law. But Acemoglu, Johnson, and Robinson (2002:table 7) find that colonies that started with dense populations were unlikely to develop effective property rights institutions. Among the former European colonies, areas with a high population density in 1500 not only are relatively poor today but also have relatively poor property rights institutions, as measured by the constraints on the executive and the risk of expropriation. The evidence supports the view that the stark political and economic inequality in some of the European colonies may have had a powerful effect on growth in the modern era through two distinct channels: direct policies that limited the breadth of access to social opportunity (see figure 2), and the insecurity of property rights engendered by the concentration of political and civil rights in a narrow elite.

The second view is that national heritage has a preeminent influence on the institutions of the former European colonies. In this view, the British colonies succeeded because of their heritage of the rule of law and constraints on the executive, and the colonies of Latin America fell behind because of their heritage of absolutist rule.⁸ One problem with this view is that the former British colonies include very poor countries—such as Belize, Guyana, and Jamaica—that resemble neighbors with similar factor endowments and no history of British rule. The theory described in this article suggests quite a different explanation for the comparative success of the former British colonies. The early colonizers of the New World—Spain and Portugal—chose the regions most heavily populated at the time, whereas Great Britain, as a late colonizer, was left with many of the less desirable lands, areas lacking dense populations or rich minerals. The need to solve the problem of labor supply in North America set the British colonies on a path of institutional development that diverged sharply from that followed in Latin America and the British sugar colonies. That path, rather than the national heritage, appears to explain the ultimate success of the North American colonies.

II. Tests of Causal Links among Inequality, Institutions, and Growth

The previous section set out the hypothesis that the level of inequality at the outset of colonization affected the evolution of institutions with strategic importance for growth. Given the limited data available for remote periods, subjecting this hypothesis to a test of causation is difficult. The richer data available for present-day developing economies permit stronger tests of causal links from the level of inequality to institutions and from institutions to growth. This section briefly describes the results of two tests using data from India.

Sugar Cooperatives in Maharashtra

Banerjee and others (2001) document and explain a puzzle in economic performance in the Indian state of Maharashtra over a recent 23-year period (1971–93). The study examines the performance of one present-day institution—sugar cooperatives. Cooperatively organized production of sugarcane has lower yields and lower growth in the fertile eastern region of the state than in the arid, less fertile western region. The surprising resolution of the puzzle illustrates the effects that the distribution of property rights can have on efficiency.

In the eastern region of the state, landholdings are heterogeneous in size. The wealthier members of the sugar cooperatives, who wield disproportionate control rights, set low prices for sugar supplied by members and divert the retained earnings to their own benefit. They prefer a lower price because it increases the rents they can extract. But a side effect of low sugar prices is weak incentives for farmers to improve productivity.

In the western region of the state, landholdings tend to be uniformly small. The cooperatives set the prices for sugar supplied by members near the world price. With high sugar prices, the farmers have strong incentives to improve productivity. Thus growth is higher and capacity more fully exploited than in the east.

Tenancy in West Bengal

Agricultural tenancy reform in the Indian state of West Bengal provides the setting for another test of the causal link between a present-day institution and efficiency. Before 1977, sharecropping contracts in West Bengal, involving around 2 million sharecroppers, generally assigned 50 percent of output to the tenant. In 1977, a new administration came into office with enforcement of a long-dormant tenancy law as one of its highest priorities. The law gave tenants a choice of registering with the government and stipulated that tenants who had registered could not be evicted from the land they sharecropped as long as they paid the landlord a minimum share-rent of 25 percent of output. Thus for most tenants, the reform increased their share of output on sharecropped land from 50 percent to 75 percent. It also gave them permanent, inheritable tenure on the land they sharecropped. In the decade after the reform, West Bengal achieved a breakthrough in agricultural productivity growth (table 4).

Banerjee, Gertler, and Ghatak (2002) use two approaches to measure the effect of the land reform on productivity. The first is a quasi-experimental approach that uses Bangladesh as a control. The second uses the fact that the land reform was implemented more intensively in some areas than in others to identify the effect of the change in tenants' share on agricultural production. The two approaches provide similar estimates, showing that the tenancy reform increased sharecropper yields by between 51 and 62 percent. These estimates imply that the tenancy reform explains

Table 4. Annual Growth in Food Grain Production in West Bengal and All of India, 1968–2001 (%)

<i>Period</i>	<i>West Bengal</i>	<i>All of India</i>
1968–81	0.43	1.94
1981–92	6.97	2.36
1992–2001	0.74	1.52

Source: For 1968–81, Banerjee, Gertler, and Ghatak (2002); for 1981–2001, Centre for Monitoring the Indian Economy (2002).

more than a quarter of the subsequent growth in agricultural productivity in West Bengal from 1979 through 1993. In that period, the rate of agricultural productivity growth in West Bengal rose from one of the lowest among Indian states to one of the highest. In the past decade, however, growth has slowed, for reasons not yet examined.⁹

III. Patent Institutions: Democratic and Undemocratic

As already described, an initial level of inequality tends to reproduce itself through such institutions as suffrage, schooling, land policies, and property rights institutions. But it is not just the broad policies of a government that influence access to economic opportunity. The fine details are also important. Rules that appear to be symmetric, because they apply equally to all, will not provide symmetric access if transaction costs affect different groups in the population differently. A comparison of the 19th-century patent systems in the United States and Great Britain provides an illustration. This comparison is used because of the richness of the data. But the kinds of costs and procedures that limited access to patent rights in 19th-century Britain were also important in Brazil and Mexico in the 19th century (Engerman and Sokoloff 1997; Khan and Sokoloff 2001).

The U.S. patent system of the 19th century was strongly influenced by British law. But the U.S. system made three innovations. First, it dramatically lowered fees. Then it gave inventors unrestricted freedom to assign their patent rights to others, which promoted a market for new technology and made it possible for individuals to become full-time inventors (Lamoreaux and Sokoloff 1999). Finally, it created impersonal administrative procedures for handling applications, which extended the protection of property rights to those with no special influence with government while reducing uncertainty about the value of inventive activity (table 5).

Not surprisingly, the differences in characteristics between the patent systems were associated with very different outcomes. The British system tended to restrict access to intellectual property rights to an elite—those with substantial wealth,

Table 5. Characteristics of Patent Institutions in the United States and Great Britain, 1790–1860

<i>Characteristic</i>	<i>United States</i>	<i>Britain</i>
Patent fee	\$30–35 (20–30 percent of per capita income in 1790, and a declining share thereafter). Patent applications could be mailed to the Patent Office free of postage.	A multiple of per capita annual income, and 35–40 times the U.S. fee in real terms.
Freedom to assign rights to inventions	Unrestricted.	The number of assignees was limited and could be increased only by a private act of Parliament.
Application process	Impersonal, routine administrative procedure.	Processing by seven different offices was required. Many more offices were involved if patent protection was to extend to other British Isles besides England. The signature of the sovereign was required at two distinct stages.
Dispute settlement	Relatively predictable, since many parameters were established by statute.	Judges had jurisdiction in determining the usefulness of an invention and thus the validity of a patent. In doing so, judges exercised broad discretion.
Access of inventors to information on patents previously granted	All records were centralized in one office, to which access was free.	A fee was required to read a patent. Obtaining information was sufficiently difficult that patent agents were generally used as intermediaries. They were few in number and thus able to keep the prices for their services high.

Source: Author's summary based on Sokoloff and Khan (1990) and Khan and Sokoloff (1998).

political connections, or technical knowledge. One measure of the breadth of access to intellectual property rights is the share of patentees who were merchants, professionals, and gentlemen. Among urban patentees in the United States, that share fell from 50 percent in the period 1790–1804 to 18.6 percent in 1836–50, whereas for all English patentees, this share remained roughly constant at 40 percent (table 6).

Another measure of the breadth of access to intellectual property rights is the share of patents granted to people with little previous record in invention. In the early 19th century, the share of patents granted to people who received a single patent over their career was 57.5 percent in the United States and 42.9 percent in Britain (table 6). In the United States, this share rose over the early 19th century, as “men with relatively common skills and knowledge [were pulled] into invention” (Sokoloff and Khan 1990:377). Indirect evidence suggests that such patentees were responsible for more than just low-value patents.

Table 6. Characteristics of Inventive Activity in the United States and the United Kingdom, 1790–1860

<i>Characteristic</i>	<i>United States</i>	<i>Britain</i>
Share of patentees who were doctors, “gentlemen,” or merchants in the periods ^a :		
1790–1804	50.0 percent	41.8 percent
1805–1822	38.7	40.9
1823–1836	24.6	47.7
1836–1850	18.6	39.1
Share of patents granted to individuals who received a single patent over their career, 1812–29	57.5 percent	42.9 percent
Patents per capita, 1810–30	3.5	2
Average education of patentees	Low. Even among those responsible for the most important inventions in 1790–1846, half had little or no formal schooling, and less than a quarter had attended college.	High
Composition of patent inventions	Balanced over sectors	Concentrated in capital-intensive sectors

^aThe data for the United States are based on urban patentees only. Urban patentees received 31.3 percent of all patents in 1805–11 but only 22.1 percent in 1830–36 and 28.0 percent in 1836–42. The data therefore probably understate the difference between Britain and the United States in the share of patentees who are merchants, professionals, and gentlemen.

Source: Author’s summary based on Sokoloff and Khan (1990) and Khan and Sokoloff (1998).

In the United States, even the great inventors, those whom histories of technology credit with at least one important invention, were unexceptional in schooling and technical skills. For example, nearly half of the great inventors had little or no formal schooling, and less than a fourth attended college (table 6). For them as for the small inventors, the expansion of markets during early U.S. industrialization provided incentives for committing resources to inventive activity. The activity of the great inventors was influenced by the same market forces as those driving invention by ordinary patentees, contrary to the popular view of inventors as disinterested geniuses (Khan and Sokoloff 1993).

The broad access to intellectual property rights in the United States spurred an enormous increase in patenting activity in all sectors, in virtually all subregions, and across a broad spectrum of the population. Evidence also points to important interaction effects with public infrastructure. The building of canals, such as the Erie Canal, caused patenting activity in the vicinity to boom, as detailed maps of patent awards illustrate (Sokoloff 1988). This finding suggests a link between patenting and access to extensive markets, because waterways were the only low-cost means of long-haul transport until the 1830s. Sokoloff estimates that the extension of the inland waterways in the northeastern United States accounted for a 10–20 percent increase in U.S. patents per capita between 1805–11 and 1830–36. Rural counties especially realized large and rapid growth in patents per capita as they gained access to waterways. These findings suggest important synergies between extensive markets, effective property rights, and technological change.

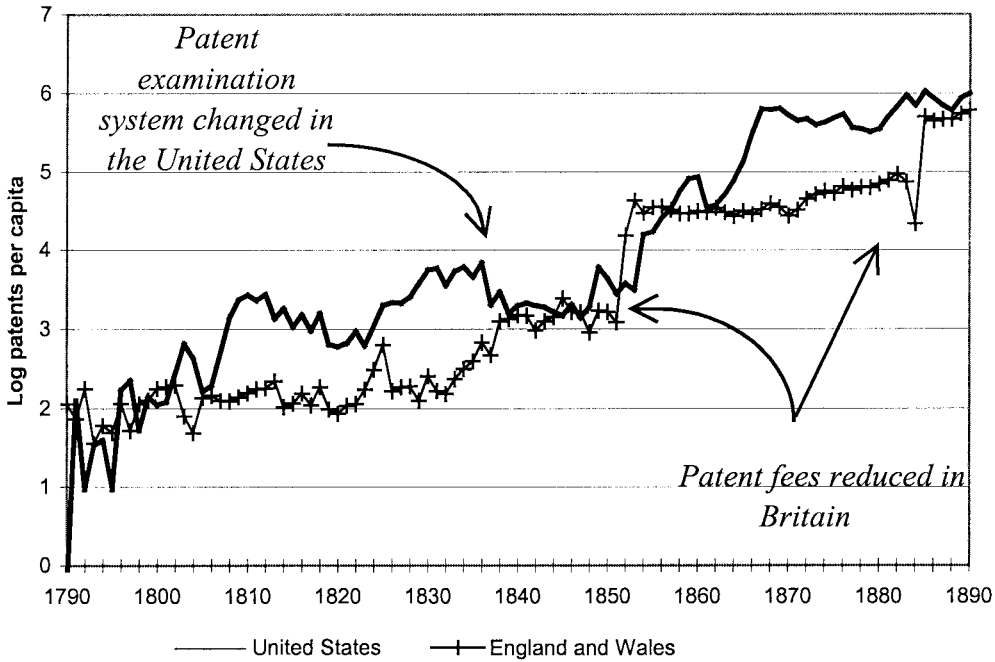
The comparison of U.S. and British patent institutions implies that the details of these institutions had an important influence on the breadth of effective access to intellectual property rights. In each country the same rules applied to everyone. But what looks symmetric in the British legal regime is not, because the system imposed needlessly high transaction costs, whose burden was much greater for the poor and the politically unconnected than for the rich.

The story has an epilogue. In 1852, one year after U.S. technology dazzled the world at the Crystal Palace exhibition in London, Britain made major changes in its patent system. In particular, it radically lowered patent fees. After these reforms the number of patents awarded in Britain jumped sharply, closing the gap with the United States in patents per capita (figure 4).

IV. Conclusion

European colonialism was a historical moment when factor endowments imposed important constraints on the choices of a small number of colonial powers with respect to the institutions they set up. This fact makes it possible to view that moment as an experiment that reveals the effects of differences in institutions. The experiment provides startling evidence of the impact of institutions on growth. It

Figure 4. Patents Per Capita in the United States and England and Wales, 1790–1890



Source: Khan and Sokoloff (1998). Reprinted with permission of Edward Elgar Publishing, Ltd.

shows that even in democracies, institutions can persist that exclude a broad segment of the population from opportunities to vote, to own land, to obtain schooling, and to secure effective protection of property rights. The reversal of fortune—when the economies that had been relatively poor surged ahead of those that had been relatively rich—occurred at the onset of the Industrial Revolution, when there was probably a premium on broad participation in entrepreneurship and innovation.

As Engerman and Sokoloff (2002:46) underscore, a “breathtaking” counterfactual is implicit in treating the moment of European colonialism as a natural experiment. In this counterfactual, but for the accident of high population density in 1500 or soils and climate suitable for sugar cultivation, areas of the Americas that are poor today would tend to be rich. Economists do not yet know what the key mechanisms of institutional change are, and they undoubtedly depend on the context. But viewing the early colonial period as a natural experiment suggests two likely drivers of institutional evolution: the level of political inequality and the scarcity of labor. These factors influence policies and institutions that in both their broad aspects (such as whether communities support public schooling) and their fine details (such as those of the British and U.S. patent systems) affect the breadth of access to opportunities for social advancement.

Although the work described here points to powerful and persistent forces underlying institutions, this work should not be taken to imply that the past is destiny. Institutions are malleable; factor endowments are malleable. Both depend on many influences. When a world exhibition of inventions provided striking evidence that Britain was falling behind in invention, it democratized its patent system. When the balance of political power shifted in favor of disadvantaged groups in West Bengal, India, in 1977, the state enforced a long-dormant tenancy law that expanded the economic rights of tenants and led to a big increase in productivity.

The nature of political competition and the balance between central and local control over government also influence institutions. In much of Latin America, national rather than local initiative finally expanded access to primary education (Engerman and Sokoloff 2002). In the United States, national forces determined many of the institutions of the South. In northeastern Brazil in recent years, a series of small reforms by a state governor created a new kind of competition among municipal governments that led to a radical change in the incentives of municipal health workers and major success in delivering health services to the poor (Tendler 1997:ch. 2).

The central questions in the research program described here are these: What are the forces that sustain inefficient institutions? How large are the effects of changing certain institutions? When can a change in institutions be sustained? Knowing the answers to these questions would take us a long way toward understanding both institutional change and the interventions that promote it. The case studies of rural institutions in India provide an example of the kind of detailed econometric work that will help advance understanding of the scope for interventions to change institutions. The research program suggests that the place to look for the obstacles to economic growth are institutions that protect narrow elites at the expense of broad access to opportunities for economic advancement and effective protection of property rights.

Notes

Karla Hoff is a senior research economist in the Development Research Group at the World Bank. Her e-mail address is: khoff@worldbank.org. The author is indebted to Gunnar Eskeland and Kenneth Sokoloff for helpful discussions, and to Zmarak Shalizi and the editors of this journal for the invitation to write this paper. She thanks Simon Johnson for assistance with Figures 1 and 3.

1. See, for example, North (1981, 1990) and Hall and Jones (1997). See World Bank (2001, 2002) for references to the vast literature.

2. The phrase *reversal of fortune* is from Acemoglu, Johnson, and Robinson (2002). Using urbanization and population density as proxies for income in 1500, the authors show that the reversal of fortune holds generally for New World colonies and for all former European colonies and that it does not hold for Europe and for all noncolonies. A parallel to the view in the 1700s of Canada as a few acres of snow—one whose prospects were quite dim—is the description of Australia as “an unchanging Paleolithic backwater” (McEvedy and Jones 1978:322, as cited in Acemoglu, Johnson, and Robinson 2002:1256).

3. This follows Engerman and Sokoloff (1997, 2002) and Acemoglu, Johnson, and Robinson (2002). Acemoglu, Johnson, and Robinson (2001a) put forward an alternative thesis: the key aspect of factor endowments was the disease environment. They argue that where prospective European settlers faced a high risk of disease, settler communities did not form and so the colonial powers did not wish to create a rule of law (broad access to secure property rights). Many scholars have raised concerns about this provocative thesis. One concern is that it fails to take into account that factor endowments, broadly defined, also imposed constraints on the ability of the colonial powers to set up a highly exploitative form of social organization: If slavery was unprofitable and if almost any settler could make a good living by farming on his own, highly inegalitarian institutions could not take root. In addition, Engerman and Sokoloff (2002) have argued that the pattern of European migration to the colonies does not support the hypothesis that the colonies with high mortality were unattractive to settlers. Finally, that areas with better institutions tend to have lower mortality makes a causal interpretation from 19th-century mortality figures to institutions problematic.

4. Slavery is just one example of an institution that influences labor scarcity. Another is the caste system in India, which barred all members of the lowest castes (shudras and “Untouchables”) from owning property and thus becoming self-employed cultivators. Many other examples are provided by Binswanger, Deininger, and Feder (1995).

5. Benabou (2000) provides a theoretical model of the tradeoffs that voters face. Easterly (1999, 2001) tests the implications empirically. Acemoglu and Robinson (2000) present a model in which political constraints block the adoption of policies that would increase wealth. For an accessible overview of the ways in which social polarization can hurt growth, see Easterly (2001:ch. 13).

6. North (1990) has suggested that the past matters because it shapes not only present opportunities but also ideologies and expectations. A characteristic expression of stark political inequality is an ideology that stigmatizes the subordinate group. In a controlled experiment, Hoff and Pandey (2003) find that the performance of low castes when they undertake a task for money is debilitated by the public announcement of caste and that the influence of caste may be mediated by distrust that effort will be rewarded.

7. Douglass North adopted this position in his early work (North and Thomas 1970) but abandoned it in his later work (North 1990). Broadly speaking, this view underlies the traditional neoclassical model, which assumes that fundamental forces of endowments, preferences, and technology drive allocations, history does not matter, and economic outcomes are the same as those that would emerge as the equilibrium allocation under a competitive market system.

8. A parallel explanation focuses on religious heritage. In one version, the distinctiveness of the North American colonies was due to their Protestant heritage. A remarkable example that again points to the importance of factor endowments and the much lesser importance of religion is explored in Kupperman’s (1993) comparative study of two Puritan colonies, the Massachusetts Bay Colony and Providence Island off the coast of Nicaragua.

9. These two tests provide evidence of a causal link between inequality and institutions and between institutions and growth but do not bear on the third aspect of the work by Engerman and Sokoloff and Acemoglu, Johnson, and Robinson—institutional overhang, the persistent effect of institutions that have been formally changed. Banerjee and Iyer (2002) test for institutional overhang in their work on agricultural productivity differences across districts in India. In some areas, the British colonial regime conferred de facto governmental powers on a landlord by giving him the right to collect taxes and retain a substantial share of the revenue. In other districts, landlords were bypassed in favor of village councils or direct taxation of cultivators, and the cultivators were given implicit property rights to the land. The land revenue system introduced by the British ended in 1947—independent India—does not impose direct taxes on agricultural incomes. Yet Banerjee and Iyer find that districts that were controlled by landlords have lower current agricultural productivity today, stemming from lower rates of investment and lower use of modern inputs, than do areas where property rights to the land were given to the cultivators.

References

- Acemoglu, Daron and James A. Robinson. 2000. "Political Losers as a Barrier to Economic Development." *American Economic Review Papers and Proceedings* 90(2):126–30.
- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001a. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5):1369–401.
- . 2001b. *Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution*. NBER Working Paper 8460. National Bureau of Economic Research, Cambridge, Mass.
- . 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117(4):1231–94.
- Banerjee, Abhijit and Lakshmi Iyer. 2002. "History, Institutions and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." Massachusetts Institute of Technology, Department of Economics, Cambridge, Mass.
- Banerjee, Abhijit, Paul Gertler, and Maitreesh Ghatak. 2002. "Empowerment and Efficiency: Tenancy Reform in West Bengal." *Journal of Political Economy* 110(2):239–80.
- Banerjee, Abhijit, Dilip Mookherjee, Kaivan Munshi, and Debraj Ray. 2001. "Inequality, Control Rights, and Rent Seeking: Sugar Cooperatives in Maharashtra." *Journal of Political Economy* 109(1):138–90.
- Benabou, Roland. 2000. "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review* 90(1):96–129.
- Binswanger, Hans, Klaus Deininger, and Gershon Feder. 1995. "Power, Distortions, Revolt and Reform in Agricultural Land Relations." In T. N. Srinivasan and Jere Behrman, eds., *Handbook of Development Economics*. Vol. 3B. Amsterdam: Elsevier Science.
- Centre for Monitoring the Indian Economy. 2002. "Agriculture." New Delhi, India.
- Easterly, William. 1999. "The Middle Class Consensus and Economic Development." Policy Research Working Paper 2346. World Bank, Development Research Group, Washington, D.C.
- . 2000. "It's Not Factor Accumulation: Stylized Facts and Growth Models." World Bank, Development Research Group, Washington, D.C. Online document available at www.worldbank.org/html/prdmg/grthweb/stylized.htm.
- . 2001. *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics*. Cambridge, Mass.: MIT Press.
- Engerman, Stanley L. and Kenneth L. Sokoloff. 1997. "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies: A View from Economic Historians of the United States." In Stephen Haber, ed., *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800–1914*. Stanford, Calif.: Stanford University Press.
- . 2001. *The Evolution of Suffrage Institutions in the New World*. NBER Working Paper 8512. National Bureau of Economic Research, Cambridge, Mass.
- . 2002. "Factor Endowments, Inequality, and Paths of Development among New World Economies." *Economia* 3(1):41–109.
- Engerman, Stanley L., Elisa V. Mariscal, and Kenneth L. Sokoloff. 1997. "The Persistence of Inequality in the Americas: Schooling and Suffrage, 1800–1945." University of California at Los Angeles, Department of Economics.
- Galenson, David W. 1996. "The Settlement and Growth of the Colonies: Population, Labor and Economic Development." In Stanley L. Engerman and Robert Gallman, eds., *The Cambridge Economic History of the United States*. Vol. 1. New York: Cambridge University Press.
- Glaeser, Edward L. 1999. "The Future of Urban Research: Nonmarket Interactions." Harvard University, Department of Economics, Cambridge, Mass.

- Hall, Robert E., and Charles I. Jones. 1997. "Levels of Economic Activity across Countries." *American Economic Review* 87(May):173–77.
- . 1999. "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114:83–116.
- Hoff, Karla. 2001. "Beyond Rosenstein-Rodan: The Modern Theory of Coordination Problems in Development." In Boris Pleskovic and Nicholas Stern, eds., *Annual World Bank Conference on Development Economics 2000*. Washington, D.C.: World Bank.
- Hoff, Karla and Priyanka Pandey. 2003. "Why Are Social Inequalities So Durable? An Experimental Test of the Effect of Indian Caste on Performance." World Bank, Development Research Group, Washington, D.C.; and Pennsylvania State University, Department of Economics, College Park, Pa.
- Hoff, Karla and Joseph E. Stiglitz. 2001. "Modern Economic Theory and Development." In Gerald Meier and Joseph E. Stiglitz, eds., *Frontiers of Development Economics*. Oxford: Oxford University Press.
- Khan, B. Zorina and Kenneth L. Sokoloff. 1993. "'Schemes of Practical Utility': Entrepreneurship and Innovation among 'Great Inventors' in the United States, 1790–1865." *Journal of Economic History* 53(2):289–307.
- . 1998. "Patent Institutions, Industrial Organization and Early Technological Change: Britain and the United States, 1790–1850." In Maxine Berg and Kristine Bruland, eds., *Technological Revolutions in Europe*. Cheltenham, U.K.: Elgar.
- . 2001. "The Innovation of Patent Systems in the Nineteenth Century: A Historical Perspective." University of California at Los Angeles, Department of Economics.
- Kupperman, Karen Ordahl. 1993. *Providence Island, 1630–1641: The Other Puritan Colony*. New York: Cambridge University Press.
- Lamoreaux, Naomi R. and Kenneth L. Sokoloff. 1999. *Inventive Activity and the Market for Technology in the United States, 1840–1920*. NBER Working Paper 7101. National Bureau of Economic Research, Cambridge, Mass.
- McEvedy, Colin and Richard Jones. 1978. *Atlas of World Population History*. New York: Facts on File.
- North, Douglass C. 1981. *Structure and Change in Economic History*. New York: Norton.
- . 1990. *Institutions, Institutional Change and Economic Performance*. New York: Cambridge University Press.
- North, Douglass C. and Robert Paul Thomas. 1970. "An Economic Theory of the Growth of the Western World." *Economic History Review* 2d ser., 23:1–17.
- Sokoloff, Kenneth L. 1988. "Inventive Activity in Early Industrial America: Evidence from Patent Records, 1790–1846." *Journal of Economic History* 48(4):813–50.
- Sokoloff, Kenneth L. and Stanley L. Engerman. 2000. "History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World." *Journal of Economic Perspectives* 14(3):217–32.
- Sokoloff, Kenneth L. and B. Zorina Khan. 1990. "The Democratization of Invention during Early Industrialization: Evidence from the United States, 1790–1846." *Journal of Economic History* 50(2):363–78.
- Tendler, Judith. 1997. *Good Government in the Tropics*. Baltimore, Md.: Johns Hopkins University Press.
- World Bank. 2001. *World Development Report 2000/2001: Attacking Poverty*. New York: Oxford University Press.
- . 2002. *World Development Report 2002: Building Institutions for Markets*. New York: Oxford University Press.