The Importance of History for Economic Development

Nathan Nunn

Department of Economics, Harvard University and NBER, Cambridge, Massachusetts 02138; email: nnunn@fas.harvard.edu

Key Words
path dependence, colonialism, institutions, norms, culture

Abstract
This article provides a survey of a growing body of empirical evidence that points toward the important long-term effects that historic events can have on economic development. The most recent studies, using microlevel data and more sophisticated identification techniques, have moved beyond testing whether history matters and attempt to identify exactly why history matters. The most commonly examined channels include institutions, culture, knowledge and technology, and movements between multiple equilibria. The article concludes with a discussion of the questions that remain and the direction of current research in the literature.
1. INTRODUCTION

In recent years, an exciting new literature has emerged empirically examining whether historic events are important determinants of economic development today. The origins of this literature can be traced to three lines of research that began roughly one decade ago. Engerman & Sokoloff (1997, 2002) examined the importance of factor endowments and colonial rule for the subsequent economic development of colonies within the Americas. Acemoglu et al. (2001, 2002) developed a research agenda that sought to better understand the historical origins of current institutions and their importance for long-term economic development. The line of inquiry undertaken by La Porta et al. (1997, 1998) also examined the importance of colonial rule, but they focused on the legal institutions that were transplanted by the different colonial powers and the long-term consequences this had for investor protection and financial development.

What united these three lines of research, and what made them particularly novel at the time, was their analysis of the potential importance of an historic event, colonial rule, for long-term economic development. These three studies spawned a large literature of empirical studies seeking to identify the importance of historic events for economic development. The earliest subsequent studies typically examined correlations between variables quantifying the impact of historic events, which almost exclusively was colonial rule, with a country as the unit of observation. These initial studies were successful at highlighting correlations in the data consistent with the notion that history can matter, even in the long-run. However, because of their inability to establish causality, the evidence presented was suggestive at best. For examples of these early studies, see Grier (1999), Englebert (2000a,b), Bertocchi & Canova (2002), and Price (2003).

Since these early contributions, the literature has developed in a number of significant ways. Much more effort has been put into collecting and compiling new variables based on detailed historic data. Recent studies, exploiting these richer data sources, are also able to employ much more satisfying identification strategies that typically rely on instrumental variables, falsification tests, regression discontinuities, differences-in-differences estimation, or propensity score matching techniques: See Acemoglu & Johnson (2004), Banerjee & Iyer (2005), Iyer (2007), Berger (2008), Dell (2008), Huillery (2008a), Nunn (2008a), Nunn & Qian (2008), Nunn & Wantchekon (2009), and Feyrer & Sacerdote (2009).

The literature has also moved beyond simply estimating reduced-form causal relationships between historic events and economic development. For many studies, the goal is also to explain exactly how and why specific historic events can continue to matter today. That is, the literature has moved from asking whether history matters to asking why history matters: See Acemoglu & Johnson (2004), Acemoglu et al. (2005a), Iyer (2007), Dell (2008), Munshi & Wilson (2008), Nunn (2008b), Nunn & Qian (2008), Nunn & Wantchekon (2009), and Becker & Woessmann (2009).

This paper provides a survey of this body of empirical research. I begin by reviewing the seminal articles by Acemoglu et al. (2001), Engerman & Sokoloff (1997, 2002), and La Porta et al. (1997, 1998) as well as the body of literature that each contribution has generated. Section 3 reviews the additional evidence from second-generation studies that provide identification-based evidence that history matters. Section 4 then surveys the precise channels of causality that have been examined in the literature. The evidence for the importance of (a) multiple equilibria and path dependence, (b) domestic institutions, (c) cultural norms of behavior, and (d) knowledge and technology is examined.
The penultimate section of the paper, Section 5, discusses the interesting relationship between geography and history that has developed in the literature. Whereas some studies have pitted these two factors against each other as alternative determinants of economic development, other studies have shown that the two factors interact in interesting and important ways. As is discussed, the existing body of evidence indicates that the greatest effect that geography has on economic development is through its influence on history. Section 6 concludes by discussing the current direction of future research.

2. THE SEMINAL CONTRIBUTIONS

The literature linking history to economic development has its origins in three distinct but related strands of research: Acemoglu et al. (2001), La Porta et al. (1997, 1998), and Engerman & Sokoloff (1997, 2002). All three examine one of the largest and most important events in the world's history: European expansion and colonization of the globe, which began in the sixteenth century.

The studies document the lasting impact that Europe's colonization had on the development paths of former colonies. They also share a common view that an important part of the causal mechanism was the impact that colonial rule had on the domestic institutions that persisted after independence. Viewed in this light, all three lines of research are conceptually consistent with one another. All three argue that the institutions of a society are an important determinant of long-term economic development and that historical events can be an important determinant of the evolution and long-term persistence of domestic institutions. Where the studies differ, however, is in their views of which aspects of colonial rule were crucial for shaping institutions and in the specifics of the proposed causal mechanisms.

For La Porta et al. (1997, 1998), the identity of the colonizer determined whether a civil law or common law legal system was established, which was important for long-term development. Unlike La Porta et al., Engerman & Sokoloff (1997, 2002) and Acemoglu et al. (2001) share the common view that the characteristics of the region being colonized were crucial factors that determined the effect of colonial rule on long-term economic development. For Acemoglu et al., the initial disease environment shaped the extent to which secure property rights were established in the colony, and through their persistence, these initial institutions had a large effect on long-term economic development. Engerman and Sokoloff focused on the importance of a region's endowment of geography suitable for growing lucrative globally traded cash crops that were best cultivated using large-scale plantations and slave labor. These large plantations resulted in economic and political inequality, which in turn impeded the development of institutions that promoted commercial interests and long-term economic growth. I now examine the three seminal contributions as well as the resulting literature that has been generated by each.

2.1. La Porta, Lopez-de-Silanes, Shleifer, and Vishny

The core of the analysis by La Porta et al. (1997, 1998) is their emphasis on the differences between legal systems based on British common law versus Roman civil law. They argue that countries with legal systems based on British common law offer greater investor
protection relative to countries with legal systems based on civil law. They then recognize that in British colonies common law–based legal systems were transplanted, whereas the European countries with a legal system based on Roman civil law—namely Spain, France, and Portugal—transplanted civil law legal systems. La Porta et al. (1997, 1998) used this historic fact to examine the causal effect of the strength of legal rules protecting investor rights on financial development. The authors argue that for former colonies legal origin is largely exogenous to country characteristics and is therefore a potential instrument that can be used to estimate the effect of the protection of investor rights on financial development. The first stage of their instrumental variables (IV) estimates shows that civil law countries, relative to common law countries, do have better investor protection, and their second-stage estimates indicate that countries with weaker investor protection have smaller debt and equity markets.

Since these initial studies, a large literature has emerged exploring the potential effect that legal origin may have on other factors [La Porta et al. (2008) provide a survey of these early studies as well as the subsequent literature that they generated]. These studies show that legal origin is also correlated with a host of other country characteristics, such as military conscription (Mulligan & Shleifer 2005a,b), labor market regulation (Botero et al. 2004), contract enforcement (Djankov et al. 2003, Acemoglu & Johnson 2004), comparative advantage (Nunn 2007b), and economic growth (Mahoney 2001). These results are both good and bad for the initial studies by La Porta et al. (1997, 1998). They suggest that legal origin may have effects that are even more wide ranging than originally assumed in La Porta et al. (1997, 1998). However, if this is the case, then the validity of their use of legal origin as an instrument for investor protection is called into question. Given that legal origin appears to be correlated with a host of other country characteristics that may also affect financial development, it is unlikely that the exclusion restrictions from original papers by La Porta et al. are satisfied. As discussed in La Porta et al. (2008), the authors are clearly aware of this fact.

Subsequent studies also look for similar relationships involving legal origin within the United States. Ten U.S. states that were first settled by either France, Spain, or Mexico initially developed civil law legal systems. Berkowitz & Clay (2005, 2006) found that today these civil law states have less independent judiciaries, lower quality courts, and less stable constitutions. Although both studies rely on ordinary least squares (OLS) estimates, they show that the correlations remain robust to controlling for a number of additional factors, such as slavery, date of entry into the Union, state size, and climatic characteristics.

Other studies also highlight correlations in the data and show that a relationship exists between the identity of the colonizer and various measures of long-term economic development. For example, Grier (1999) found that, at independence, former British colonies had on average a larger share of their populations in school. Bertocchi & Canova (2002) found that, within Africa, former British and French colonies have higher levels of investment and education after independence. Although these correlations do not provide proof of the causal importance of the identity of colonizer, they are consistent with the emphasis by La Porta et al. on the impact that the identity of the colonizer (specifically, its legal system) has on the long-term economic development of its colony.

---

2The ten states are Alabama, Arizona, Arkansas, California, Florida, Louisiana, Mississippi, Missouri, New Mexico, and Texas. Of these states, only Louisiana continues to have a civil law legal system.
2.2. Acemoglu, Johnson, and Robinson

Like La Porta et al. (1997), Acemoglu et al. (2001) also examined the effect of colonial rule on the institutions that were implemented and their long-term impact on economic development. However, Acemoglu et al. (2001) focused on an alternative determinant of the differences in institutions that evolved across former colonies. They hypothesize that, because colonies with a less deadly disease environment had greater European settlement, growth-promoting institutions were established in these colonies to protect property rights during colonial rule. In colonies in which European mortality was high and settlement low, the colonizers did not have an incentive to establish strong property rights and instead established extractive rent-seeking institutions. Using this logic, the authors estimate the causal effect of current domestic institutions on per capita income, using early European mortality rates as an instrument for institutions. One of the assumptions underlying the IV strategy is that initial settler mortality is not correlated with current income other than through domestic institutions. In the first stage of their IV procedure, the authors found a strong negative relationship between initial settler mortality and current institutional quality. The second-stage estimates indicate that domestic institutions exert a strong positive effect on per capita income.

The elegance of the paper lies in its ability to develop a clear and convincing historical narrative with supporting empirical evidence and to show how an historic event can affect past institutions, which through their persistence have an influence on income levels today. The study provides an empirical foundation to support the seminal works on the importance of institutions written by North & Thomas (1973) and North (1981, 1990); for a more recent analysis, see Greif (2006). The study emerged at a time when the literature was in the process of trying to estimate convincingly the causal impact of domestic institutions on economic development: Early papers in this literature include De Long & Shleifer (1993), Knack & Keefer (1995), Mauro (1995), Hall & Jones (1999), and Englebert (2000a,b). An important contribution of Acemoglu et al. (2001) was to develop a much more satisfying identification strategy than that provided by previous empirical studies.

A number of studies have attempted to extend Acemoglu et al.’s line of research, providing evidence for the importance of historic institutions for current economic development. Two recent studies by Banerjee & Iyer (2005) and Dell (2008), rather than taking a broader, more macro perspective, focus on a specific region. By doing this, the authors are able to collect and analyze richer data at a more micro level. The use of these richer data also allows the authors to employ additional estimation strategies that help identify the causal effects of history on economic outcomes today.

Dell (2008) examines the mita forced mining labor system, which was instituted by the Spanish in Peru and Bolivia between 1573 and 1812. The study combines contemporary household survey data, and geographic data, as well as data from historic record, and uses a regression discontinuity estimation strategy to identify the long-term impacts of the mita system. Her identification exploits the fact that there was a discrete change in the boundaries of the mita conscription area and that other relevant factors likely vary smoothly around the mita boundary. As a result, comparing the outcomes of mita and non-mita districts very close to the border provides an unbiased estimate of the long-term effects of the mita. The study found that the mita system had an adverse effect on long-term economic development. All else being equal, former mita districts now have an average
level of household consumption that is 32% lower than households in former non-mita
districts. The study also found that a significant proportion of the difference can be
explained by lower levels of education and less developed road networks. Dell’s study
provides valuable evidence showing that the institutions established during colonial rule
can have long-term impacts that continue to be felt today.

Like Dell (2008), Banerjee & Iyer (2005) also analyze the long-term effects of colonial
institutions, but they examine differences in revenue collection institutions across districts
within colonial India. The authors compared districts where revenue was historically
collected directly by British officials against districts where revenue was collected by
native landlords. They found that, after independence, districts with nonlandlord systems
have higher levels of health, education, and agricultural technology investments relative to
those levels in landlord systems. To determine the extent to which the correlation is causal,
the authors exploit the fact that, in the parts of India conquered between 1820 and 1856,
nonlandlord revenue collection was implemented. They argue that the historic reasons for
this pattern are orthogonal to district characteristics, and therefore, the date of British
conquer can be used as an instrument for the revenue collection system. Their IV estimates
are consistent with their OLS estimates. They also show that their OLS results are robust
when the sample is restricted to 35 districts, in which all landlord districts are bordered by
nonlandlord districts.

Although the analysis of Banerjee & Iyer (2005) and Dell (2008) provides evidence of
the long-term impacts of initial colonial institutions, the studies do differ from that by
Acemoglu et al. (2001) because the transmission mechanism is not the persistence of these
initially implemented institutions. In Dell (2008), the hypothesized mechanism is the
concentration of wealth and power and the resulting provision of public goods. Similarly,
in the analysis by Banerjee & Iyer the transmission mechanism is not through the persis-
tence of these initially implemented institutions, because the differences in colonial land
revenue collection systems no longer exist.

One study that does empirically link early colonial institutions to institutional out-
comes today is sociologist Matthew Lange’s (2004) analysis of the differential effects of
indirect rule relative to direct rule on the quality of institutions and governance today.
Using colonial documents housed in Britain’s Public Records Office, Lange compiled
information on court cases held in 33 former British colonies in 1955. He then used the
fraction of the court cases that were presided over by local chiefs, rather than colonial
officials, as a measure of the extent of indirect rule in each country. The measure was
intended to provide a proxy for the overall extent to which colonial rule in the country
relied on traditional legal, political, and institutional structures. The study found a posi-
tive relationship between the extent of indirect rule and a variety of measures of institu-
tional quality and good governance. The primary shortcoming of the study, however, lies
in its lack of a convincing identification strategy. Because the paper relies on OLS esti-
mates, it is unknown whether the correlations between past and current institutions
capture the causal effect of historic institutions on institutions today or simply reflect a
spurious correlation driven by omitted country characteristics.

2.3. Engerman and Sokoloff

The studies by Engerman & Sokoloff (1997, 2002) focus on the differential paths of
development among the New World countries of the Americas. Engerman and Sokoloff
argue that the different development experiences of the countries in the Americas can be explained by initial differences in endowments of land and geography suitable for growing globally traded crops like sugar, which were best grown on large-scale plantations using slave labor. These areas were characterized by severe economic and political inequality, which resulted in the subsequent evolution of domestic institutions that protected the privileges of the elites and restricted the participation of the rest of the population in the commercial economy. In former Spanish colonies, endowments of rich mineral resources further strengthened the tendency toward political and economic inequality.

Because the analysis of Engerman & Sokoloff (1997, 2002) is primarily qualitative in nature, the literature has taken the natural next step of beginning to examine empirically the core predictions that arose from their analysis. A number of studies have uncovered correlations in the data that support the authors’ hypothesis that slavery was detrimental for subsequent economic development. These studies found a negative relationship between various measures of economic development and past slave use by looking across U.S. states (Mitchener & McLean 2003) or counties (Lagerlöf 2005) or across New World countries (Nunn 2008b).

Studies have also examined additional hypotheses that arise from the study by Engerman & Sokoloff (1997, 2002). Nunn (2008b) used information from the 1860 U.S. Census to examine whether the data support the inequality channel proposed by Engerman and Sokoloff. The study, based on OLS estimates, found a strong positive relationship between slavery and inequality in the size of land holdings, when looking across either states or counties. This finding is consistent with the notion that slavery resulted in economic inequality. However, Nunn failed to find any evidence of an empirical relationship between initial economic inequality and current income levels. Although the counties and states with higher levels of land inequality in 1860 also have higher levels of income inequality today, these states and counties do not have lower levels of income today.3 In other words, there is no evidence of a relationship between historic inequality and current economic underdevelopment as hypothesized by Engerman and Sokoloff. Instead, the data are most consistent with slavery having had two unrelated consequences: It increased economic inequality, which persists today, and it resulted in lower levels of economic development.

Nunn (2008b) also looked outside of the United States to examine the inequality channel proposed by Engerman and Sokoloff. Specifically, he examined their argument that slavery was detrimental because it took the form of large-scale plantation slavery, which promoted extreme inequality. Using data from British Slave Registrars, Nunn disaggregated slave populations into those that were used on plantations and those that worked in other occupations such as industry, livestock, salt, timber, fishing, and shipping. In a second exercise, Nunn disaggregated slave populations by the size of the holdings on which they were held. By doing this, he is able to examine whether the negative correlation between slavery and economic development is being driven by large-scale plantation slavery. He found that this is not the case. All forms of slavery are negatively correlated with subsequent economic development, and when statistically significant differences in the estimates exist, large-scale plantation slavery is less correlated with underdevelopment than is smaller-scale nonplantation slavery.

3Related to this is Frankema’s (2006) finding of a positive relationship between colonial land inequality and postcolonial income inequality across former colonies.
Additional evidence on the validity of the inequality channel hypothesized by Engerman and Sokoloff comes from the study by Acemoglu et al. (2008) that examines variation across municipalities within the state of Cundinamarca, Colombia. The analysis shows that within Cundinamarca there is a positive relationship between nineteenth-century land inequality and current economic development proxied by measures of primary- and secondary-school enrollment rates. This relationship is opposite to the negative relationship that one would expect to observe if the hypothesis by Engerman and Sokoloff was correct. This finding can also be contrasted to evidence from the United States showing that until 1940 a negative relationship existed between inequality and education (Galor et al. 2009). The study also showed that in Cundinamarca economic inequality and political inequality are negatively correlated, a fact that also runs counter to the hypothesis by Engerman and Sokoloff that initial economic inequality resulted in political inequality.

Additional evidence can be gleaned from Melissa Dell’s (2008) study of colonial Peru and Bolivia. In her analysis, Dell explains her finding of a negative effect of the mita system on education and the development of road networks by the fact that in mita districts the colonial government restricted the formation of large Spanish land holdings, called haciendas. Because the large-land owners typically lobbied for the greater provision of public goods, the non-mita districts, without these large-land owners, had lower levels of public goods. Dell provides empirical evidence to support her argument. Comparing districts on both sides of the mita boundary, one observes a significantly lower presence of haciendas in former non-mita districts both during and after the mita as well less dense road networks and less market integration today. These estimates support her explanation that because mita districts had fewer haciendas, they did not have large-land owners that were able to lobby for local public goods. Dell’s explanation and her supporting evidence run contrary to the inequality channel put forth by Engerman and Sokoloff. In Dell’s setting, not enough (rather than too much) concentration and inequality in land ownership was the cause of long-term economic underdevelopment.

3. ADDITIONAL EVIDENCE THAT HISTORY MATTERS

Much of the early literature examining the importance of history for economic development examined correlations between different measures of colonial rule and current economic performance across countries (Grier 1999; Englebert 2000a,b; Bertocchi & Canova 2002; Price 2003). Although these studies are useful in highlighting correlations that exist in the data, they stop short of providing causal evidence of the effect of history on long-term development. The more recent empirical studies at the forefront of this literature combine much richer data, typically at a level of analysis much finer than the country, with increasingly sophisticated identification techniques to provide causal evidence of the importance of history for long-term economic development.

One example of this line of research is Huillery’s (2008a) analysis of the differential effects of colonial rule across districts within French West Africa. Her study combines data from historic documents from archives in Paris and Dakar with household surveys from the 1990s. She showed that looking across districts there is a positive relationship between early colonial investments in education, health, and infrastructure and current levels of schooling; health outcomes; and access to electricity, water, and fuel. The study then tested for the causal effect of colonial policy on these subsequent outcome measures. Exploiting
the richness of her district-level data set, Huillery used a regression discontinuity estimation strategy, comparing the differences in her outcomes of interest between neighboring districts only. Her estimates provide strong support that colonial investments on education, health, and infrastructure have positive and persistent effects on current education, health, and infrastructure.

Like Huillery (2008a), the recent study by Berger (2008) also uses microlevel data to identify long-term impacts of colonial policies. The study also used a regression discontinuity estimation strategy and examined the difference in outcomes between Northern Nigeria and Southern Nigeria, two regions of colonial Nigeria. During colonial rule, the two districts were split by a straight line running east-west located at 7°10′ N. Among the many differences between the two regions, one of the primary ones was the form of taxation. In the coastal Southern region, revenue was raised primarily through tariffs on internationally traded goods. The landlocked North relied on a poll tax to raise revenue. As a result, in the North, it was necessary to develop an institutional structure and network to collect taxes and to develop mechanisms to control corruption and incompetence of government officials and native chiefs.

Berger (2008) used vaccination rates as his measure of local government quality today. Comparing households within one degree on either side of the historic border between Northern and Southern Nigeria, he found that vaccination rates are significantly lower on the Northern side of the border. A falsification test showed something distinct about the historic border located at 7°10′ N. As one moves the border north or south, the estimates show no difference in vaccination rates of households on either side of the false borders. It is only for this historic border, which does not correspond to any modern boundary, that a significant difference exists.

Other studies, using more sophisticated identification techniques, have revisited old questions that were never satisfactorily answered in the early literature. One of these questions concerns the effect that the length of colonial rule had on subsequent development. Grier (1999) and Price (2003) examined this question previously, but their analyses relied on OLS estimates. Feyrer & Sacerdote (2009) used a novel instrumental variables estimation strategy to provide the first causal estimates of the impact of the length of colonial rule. The authors, as part of their identification strategy, considered only former island colonies. They argued that how early an island was discovered was determined in part by its location relative to prevailing wind patterns and that these wind patterns most likely do not affect long-term development through channels other than through the island's date of discovery. If this is the case, then the wind vectors surrounding an island can be used as instruments to estimate the causal effect of the length of colonial rule on subsequent development. Their baseline set of instruments, which are constructed from satellite imagery data reported monthly on a one degree by one degree global grid, includes the annual mean and variance of monthly east-west wind vectors.

Their first-stage estimates show that stronger westerly winds are associated with earlier discovery and more years under colonial rule. According to their second-stage estimates, the length of colonial rule has a positive effect on per capita income in 2000. This result may seem surprising because it appears to provide evidence that colonial rule was good for economic development. However, this conclusion does not follow from their study. Because the estimated effects are for the length of colonial rule conditional on being colonized, it does not provide an estimate of the average effect of being colonized relative to not being colonized.
A second long-standing question that has recently been revisited with improved identification regards the long-term effect of direct colonial rule relative to indirect rule. The distinction between direct and indirect rule, although not always clear-cut, lies in the amount of bureaucratic responsibility given to native agents rather than European administrators. Direct rule occurs when only the lowest levels of responsibility are given to natives and all other positions are occupied by Europeans. Under indirect rule, much more power is given to natives, and typically local governance structures and local elites are integrated into the colonial governance structure.

The recent study by Iyer (2007) employed an empirically persuasive identification strategy to examine the relative effects of direct versus indirect rule across regions within India. She addressed the problem of the endogeneity of the form of British rule by exploiting the “Doctrine of Lapse,” which was in place between 1848 and 1856. The doctrine, which stated that a native ruler’s adopted heirs were not to be recognized by the British government, allowed the British to annex several states where the native ruler died without a natural heir. Exploiting the Doctrine of Lapse, Iyer constructed a district-specific indicator variable that equals one if the ruler died without a natural heir between 1848 and 1856 when the Doctrine of Lapse was in place. In her analysis, Iyer further considered only the subset of states that were not annexed before 1848.

Looking across 415 districts, Iyer estimated the effect of colonial indirect and direct rule on investment in agriculture and agricultural productivity today. She first estimated the relationship using OLS and found a positive and statistically significant relationship between direct rule on agricultural investments and productivity. By contrast, her IV estimates showed no statistically significant difference between the two types of districts. The difference between the IV and OLS estimates is consistent with the British selectively annexing the most productive states. Iyer (2007) also examined the effect of direct rule on the availability of public goods such as health, education, and roads. Again, she found that the IV estimates are consistently lower than the OLS estimates, which is consistent with the British selecting the “better” states. According to the IV estimates, direct rule exerted a negative effect on the long-term availability of public goods.4

As is always the case with IV estimates, the validity of the estimates rests crucially on whether the instruments satisfy the necessary exclusion restrictions. As a test of the validity of her instruments, Iyer performed a number of falsification tests. She first reported the reduced-form relationship between her instrument and her measures of the availability of public goods. Consistent with the IV estimates, her reduced form shows a statistically significant negative relationship between her instrument and the public goods measures. Iyer then constructed a “fake instrument,” which is an indicator variable that equals one if a state’s ruler died without a natural heir between 1858 and 1884, a period when the Doctrine of Lapse was no longer in place. According to the channel proposed by Iyer, the results with the fake instrument should not provide the same reduced-form estimates as with the actual instruments. This is exactly what she found. For all outcome measures except one, she found no statistically significant relationship between the fake instrument and the outcome measures.

One of the most dramatic examples of how history can shape long-term economic development comes from Nunn’s (2008a) study of the impact that the external trade in

---

4Banerjee et al. (2005) examined the robustness of this result across 27 different public good measures. Overall, the results confirm the findings of Iyer (2007) looking across a more restricted range of public goods.
slaves had on long-term political and economic development within Africa. Combining data from historic shipping records, which report the total number of slaves shipped on each voyage from Africa, with historic records that report the ethnic identity of the slaves being shipped, Nunn constructed estimates of the number of slaves shipped during the trans-Atlantic Slave Trade. He also constructed similar estimates for Africa’s three other slave trades: the Indian Ocean, Red Sea, and trans-Saharan slave trades. The study found that the parts of Africa from which the largest number of slaves were taken (normalized by land area or historic population) are today the poorest parts of Africa. The core issue in interpreting this correlation is selection into the slave trades. If, for example, the societies with the most poorly functioning institutions and the poorest future growth prospects selected into the slave trades, then this would explain the negative relationship even if the external trade in slave trades had no direct impact on the societies within Africa.

Nunn tested whether selection is driving the results by looking at the evidence on the nature of selection during the slave trades. He found that the descriptive and quantitative evidence suggest that the least developed societies were not the ones selected into the slave trade. Instead, the more developed and more densely populated societies supplied the largest numbers of slaves. Nunn also constructed instruments based on the distance of each country from the external locations of demand for the slaves. The key identification assumption is that the location of demand influenced the location of supply and not vice versa. If, as Sokoloff & Engerman (2000) argue, the demand for slaves was determined primarily by geographic characteristics, such as soil quality suitable for plantation agriculture, then this assumption is reasonable. The IV estimates provide estimates that are consistent with the OLS estimates. Overall, Nunn concluded that the empirical evidence suggests that Africa’s external trade in slaves did have a significant negative impact on the subsequent economic development of the regions within Africa.

4. WHY HISTORY MATTERS

4.1. Multiple Equilibria and Path Dependence

If one thinks about classical models in economics, why historic events should have any impact in the long run is far from obvious. For example, in the classic Solow model, for a given set of parameter values, there is a unique steady-state level of capital and income per worker. Any event that shocks either capital or output has a temporary impact only. In the long run, the economy eventually converges back to a unique steady-state equilibrium. Viewed in this light, why history should matter is unclear. However, once models with multiple equilibria are considered, then historic events have permanent effects if they cause a movement from one equilibrium to another. For early examples of models of this type, see Murphy et al. (1993), Acemoglu (1995), and Mehlum et al. (2003).

A recent paper by Nunn (2007a) illustrates this point by developing a model that features multiple equilibria in the security of property rights and output per worker. It is then shown that a period of severe extraction, by temporarily causing the optimal equilibrium to disappear, can cause a permanent movement to an equilibrium with significantly lower income levels. Because this equilibrium is stable, the society remains “stuck” in this equilibrium even after the high-income equilibrium returns. Using historic evidence, Nunn argued that the movement from one equilibrium to another explains why in Africa the slave trade and colonial rule appear to have had permanent impacts.
On many levels, it is clear that the world we live in is filled with multiple equilibria. Suggestive evidence of this fact can be found in a variety of real-world phenomena, such as the fact that this article was typed on a QWERTY keyboard rather than the significantly more efficient DVORAK keyboard\(^5\) or the fact that cars are driven on the right-hand side of the road in the United States but on the left side in the United Kingdom. Historic events, such as the battle for market dominance between VHS and Betamax in the 1980s, provide further evidence that multiple equilibria exist.

Very recently, a number of studies have undertaken the task of testing empirically for the existence of multiple equilibria. One strategy that some studies have employed is to examine cases with an extremely large shock to an equilibrium. The studies then test whether the shock causes a movement from the previous equilibrium to a new equilibrium. Davis & Weinstein [2008; also see their precursor, Davis & Weinstein (2002)] examined the effect of bombings on 114 Japanese cities during World War II. The authors found that the cities returned to their prebombing populations, regained their shares in total manufacturing output, and, most surprisingly, also regained their pre-existing industrial composition. Overall, the results point toward the existence of a unique stable equilibrium of production, rather than the existence of multiple equilibria. A similar result is found in the analysis by Miguel and Roland (2006) of the long-term effects of the U.S. bombings in Vietnam. The authors found that the bombings had no long-term effects on populations, poverty, or consumption 25 years later.

These tests, although very suggestive, do not provide airtight proof against the existence of multiple equilibria in these environments. This is true for a number of reasons: Other equilibria may exist, but the shocks were not large enough to cause a movement to a new equilibrium. Alternatively, the shocks may have been large enough to cause a movement, but because the prebombing equilibrium is focal, this equilibrium may have been selected after the bombings among all possible equilibria (see Schelling 1960).

An innovative study by Redding et al. (2007) tests for the existence of multiple equilibria in an environment distinct from those studied by Davis & Weinstein (2008) and Miguel & Roland (2006). The study examines the location of airport hubs in Germany before and after the division of Germany following World War II. They showed that, after division, the location of West Germany’s primary airport hub switched from Berlin to Frankfurt. After reunification in 1990, the location of the hub did not switch back to Berlin. Redding et al. (2007) showed that this shift cannot be explained by changes in fundamentals over the time period. Thus, the evidence suggests that the temporary division of Germany resulted in a permanent movement of the location of Germany’s largest airport hub.

One of the most dramatic examples of how an historic event can drastically alter the subsequent path of economic development is the history of Tasmania. According to archaeological evidence, humans first arrived in what today is Tasmania as early as 22,000 years ago across a land bridge that connected present-day Australia to Tasmania. Approximately 12,000 years ago, at the end of the Pleistocene, rising sea levels drowned the land bridge, turning it into what is now the shallow Bass Straight (Diamond 1978). The archeological evidence indicates that the technologies used by the Tasmanians deteriorated over time. The Tasmanians lost their ability to construct bone tools, make cold-weather clothing, and catch fish.

\(^5\)For a discussion of the historic reasons for the adoption of the QWERTY keyboard see David (1985).
An explanation for Tasmania’s dramatic deterioration has been put forth by anthropologist Joseph Henrich (2004). He developed a model where skills are imperfectly transmitted from the most skillful or successful individual in the society to all other members of the society. Henrich showed that in this environment the size of the population matters, because in a larger population the most skilled individual is, on average, more skilled. He also showed that, because of imperfect transmission, there exists a minimum population size below which the average skill of the society decreases over time. Therefore, his model explains the Tasmanian technology losses as a result of the drop in effective population size after rising sea levels isolated Tasmania from Australia.

The Tasmanian experience provides a dramatic example of path dependence. One historic event—the rising sea levels and flooding of the Bass Straight—resulted in a permanent change in the evolutionary process and moved the society to a new trajectory, where technologies and skills deteriorated over time. One historic event permanently changed, in a very dramatic way, the long-term social and economic development of the Tasmanians.

### 4.2. Domestic Institutions

Even without the existence of multiple equilibria, historic events can still affect economic development in the long run if they alter deep determinants of long-term economic growth. The deep determinant that has received the greatest attention in the literature is domestic institutions. The emphasis on institutions can be seen clearly from the fact that in each of the seminal papers by Acemoglu et al. (2001, 2002), Engerman & Sokoloff (1997, 2002), and La Porta et al. (1997, 1998), the mechanism through which colonial rule affects current development is domestic institutions.

The focus on institutions as a causal mechanism has also continued in subsequent research, for example the study by Acemoglu et al. (2005a) of the effect that early Atlantic trade had in Europe. The authors argue that, in countries with access to the lucrative Atlantic three corner trade, economic and political power shifted toward commercial interests. As the merchant class became more powerful, they were able to alter domestic institutions to protect their interests against the interests of the royalty, and these institutional changes, in turn, had a positive effect on long-term prosperity. Using data on historic urbanization rates and per capita incomes, the authors first showed that the rise of Europe was actually a rise of the nations with access to the lucrative Atlantic trade, namely Britain, France, the Netherlands, Portugal, and Spain. Their explanation is robust even when controlling for alternative explanations of the rise of Western Europe, such as the rise of a Protestant work ethic (Weber 1930, Landes 1998), war and interstate competition (Tilly 1990), and the legacy of the Roman Empire.

Acemoglu et al (2005a) argue that profits alone cannot explain the divergent growth of Atlantic traders and that the evolution of domestic institutions played an important role in the process. To test this hypothesis, the authors extended the Polity IV data back to 1350

---

6More recently, Becker & Woessmann (2009) provided empirical support for the importance of the Protestant religion, but they focused on the religion’s emphasis on the importance of literacy so that each individual was able to read the Bible. The analysis identifies a positive effect of the Protestant religion on literacy and industrialization in nineteenth-century Prussia. [Also related is the evidence put forth by Iyigun (2008) showing that the rise of the Ottoman Empire and its movement into Europe are partly responsible for the Protestant Reformation.]

7See Inikori (2002) for the alternative view that the profits that accrued to Western Europe during the three corner Atlantic trade explain much of its growth during the time.
and showed that Atlantic trade increased the quality of domestic institutions as measured by an index of the constraints on the executive. They further hypothesized that the process of institutional change could occur only in countries that initially had nonabsolutist political institutions. They showed that the data are also consistent with this hypothesis. They showed that the increase in economic growth generated by the Atlantic trade was higher for countries with better initial domestic institutions, again measured by the constraint on the executive.

The recent study by Gennaioli & Rainer (2007) also provides evidence of the persistence of domestic institutions, but within the African context. The authors used ethnographic data to construct a measure of the level of state development in precolonial African societies. Their OLS estimates show that there is a positive correlation between precolonial political development and the provision of public goods today. This result can be combined with evidence from Nunn (2008a) showing that the parts of Africa from which more slaves were taken had less developed political systems after the slave trade (and before official colonial rule). Because the evidence for both relationships is based on correlations in the data, one must be cautious when drawing conclusions. However, the combined evidence from Gennaioli & Rainer (2007) and Nunn (2008a) is consistent with a chain of causality where the slave trade resulted in a deterioration of domestic political institutions, which in turn had a long-term adverse impact on the provision of public goods. Therefore, the evidence from the two studies provides support for the notion that history can matter through the evolution and persistence of early institutions.

The persistence of historically determined institutions is also the channel of causality in the recent study by Jha (2008) that examines the effect of early medieval trade on the formation and persistence of institutions promoting religious tolerance. Jha’s study shows that, when looking across cities within India, one finds a positive relationship between participation in overseas trade during the medieval period and religious conflict during the late nineteenth and early twentieth centuries. Jha addressed the endogeneity of the selection of medieval ports by using the existence of natural harbors as an instrument for whether a coastal city was a trading port and by using propensity score-matching techniques. His estimates show that being a town that was a medieval trading port made it less likely that the town later experienced Hindu-Muslim riots. Using historical evidence, Jha argued that, because Muslims provided access to the markets of the Middle East, in the towns connected to this overseas trade the returns to Hindu-Muslim cooperation were much higher. As a result, institutions that supported exchange and a peaceful coexistence between Hindus and Muslims were developed. Although Jha’s analysis provides evidence of a causal effect of medieval trade on long-term religious tolerance, it stops short of verifying empirically that the intervening channel is the persistence of historic institutions.

A number of studies also provide suggestive evidence of the importance of historic institutions. Rather than empirically documenting the evolution and persistence of institutions over time, they show that historic institutions have a strong effect on economic outcomes today. Examples include Dell’s (2008) analysis of the impact of the early forced-labor institutions in colonial Peru and Bolivia as well as studies by Banerjee & Iyer (2005, 2008) of the effects of early land-tenure institutions in colonial India.

Overall, the literature since Acemoglu et al. (2001) has succeeded at providing additional evidence showing that institutions are an important channel through which history matters. However, much work remains to be done before we have a clear understanding of the effect that historic events have on the formation of early institutions and
their persistence and importance for long-term development. For example, in past studies (typically at the macro level) institutions have been conceptualized and measured as a broad cluster of institutions. As a result, institutions have largely remained a black box whose details we do not clearly understand. As empirical research continues to examine specific examples of institutional change and persistence at the micro level, our understanding of the causes and consequences of specific institutions will naturally improve.

The recent study by Acemoglu & Johnson (2004) directly tackles the task of unpacking the black box we call institutions. The authors make a distinction between two broad clusters of institutions that they refer to as “property rights institutions” and “contracting institutions.” According to their definitions, property rights institutions protect individuals from theft or expropriation by the government or elites, and contracting institutions enforce private contracts written between individuals. The study explores the historical determinants of both types of institutions. They show that, although a strong relationship exists between a country’s legal origin and the quality of its contracting environment today [consistent with the previous findings by Djankov et al. (2002), Djankov et al. (2003) and Lerner & Schoar (2005)], legal origin is not correlated with measures of property rights institutions. Extending previous research by Acemoglu et al. (2001), the study shows that there is a strong correlation between initial European settler mortality rates and property rights institutions, but no correlation between settler mortality and contracting institutions.

To unbundle the effects of these different types of institutions on economic outcomes, the authors used legal origin as an instrument for contracting institutions and initial settler mortality as an instrument for property rights institutions. Using this IV procedure, the authors estimated the effect of both types of institutions on four outcome variables: current income, investment, financial development, and the form of financial intermediation (i.e., equity versus debt contracts). On the one hand, they found that property rights institutions have a positive and significant effect on income, investment, and financial development. On the other hand, contracting institutions appear to have a much more limited impact, affecting only the form of financial intermediation.

4.3. Cultural Norms of Behavior

Another way in which historic events can have long-term impacts is if these past events permanently affect culture or norms of behavior. Although economics often bats about the notion of culture with little meaning, other disciplines often give it much more content. For example, evolutionary anthropologists have long recognized that there are clear microfoundations that explain the existence of a phenomenon like culture (e.g., Cavalli-Sforza & Feldman 1981, Boyd & Richerson 1985). If information acquisition is either imperfect or costly, then selection favors shortcuts to learning. Rather than use scarce resources to acquire all the information needed to make every decision, individuals will develop “rules of thumb.” These shortcuts or norms of behavior allow individuals to free

---

8Motivated by Acemoglu et al. (2002), the authors also used initial population density in 1500 as an alternative instrument for property rights institutions.

9This is not to say that no serious studies in economics seek to understand and model culture and its transmission in a concrete way. Examples include Verdier (2000, 2001) and Tabellini (2008).
ride on the learning of others. These shortcuts then become internalized as individuals come to believe that certain behaviors are the “right” behaviors in particular situations.

The idea that norms of behavior may be a channel through which history can affect long-term economic development is not new. One of the most famous links among history, culture, and development is Max Weber’s (1930) hypothesis that the Protestant Reformation was instrumental in facilitating the rise of industrial capitalism in Western Europe. He argued that Protestantism, in contrast to Catholicism, approves the virtues of hard work and the accumulation of wealth and that these values, referred to as the “Protestant work ethic,” provided the moral foundation that spurred the transition to a modern market-based industrial economy. A more recent example is Mokyr’s (2008) argument that an important determinant of the Industrial Revolution was the development of a social norm he calls “gentlemanly culture” that emphasized honesty, commitment, and cooperation.

Empirical studies exploring the link between history and culture typically attempt to isolate culture by comparing some measure or proxy of a cultural belief across individuals in the same external environment. In this context, systematic differences in individuals’ stated beliefs or actions when faced with the same situation provide suggestive evidence of the importance of culture.

This logic builds on earlier studies that attempt to identify the existence of cultural differences between groups. In this literature, two strategies are employed in an attempt to identify systematic differences in behavior among individuals in the same environment. The first strategy is to bring the same environment to people of different backgrounds. Henrich et al. (2001, 2005) undertook this strategy in their studies where the “ultimatum game” was conducted in 15 extremely remote small-scale societies across the world. The second strategy that has been employed is to examine situations where individuals from different backgrounds have been brought into the same environment. Miguel & Fisman (2007) looked for a culture of corruption using the accumulation of unpaid parking violations among foreign diplomats stationed in Manhattan. Following a similar logic, Miguel et al. (2008) sought to identify the existence of a culture of violence by examining the number of yellow- and red-card fouls given to soccer players from different countries playing in six professional soccer leagues. Fernandez & Fogli (2007) examined the behavior of individuals whose parents were born in different countries, but today live in the United States. They found that the labor force participation and fertility of second-generation women are positively correlated with the historic labor force participation and fertility of the individual’s country of ancestry.

The earliest studies examining the possibility that cultural norms may be historically determined use an experimental setting to control for the external environment of individuals. Psychologists Cohen et al. (1996) tested whether in the U.S. South there is a “culture of honor,” where a special importance is placed in defending one’s reputation and honor, even if this requires aggression and violence. Their explanation for why this culture exists in the U.S. South and not the North lies in the different histories of settlement in the two areas. The North was settled by groups with a farming background, whereas the South was settled primarily by the Celts who had been herders since prehistoric times and had never engaged in large-scale agriculture. They argued that, historically, in herding cultures, with their low population densities and weak states, protection of one’s property was left to the individual. The result of this is that some aspects of the norms of aggressive behavior that were developed as a means to protect one’s herds continue to persist today.

80  Nunn
To test the culture of honor hypothesis, Cohen et al. (1996) conducted a series of experiments involving white men from the U.S. North and South. In the experiments, each individual was bumped by an accomplice and called an “asshole.” (The participants did not know this was part of the experiment.) Using a number of methods including direct observation, psychological tests, and saliva tests, the authors compared the effects of this incident on Southerners relative to Northerners. They found that the Southerners became more upset, were more likely to feel their masculinity was threatened, became more physiologically and cognitively primed for aggression as measured by a rise in testosterone and cortisol levels, and were more likely to engage in subsequent aggressive behavior.

Salaman (1980) provides a study well-known in the sociology literature that compares two similar towns located in the same county in Eastern Illinois. Both towns were settled approximately 100 years ago, but one was settled by German immigrants from East Frisia and the other by Irish immigrants. Using survey data, Salaman documented the persistent of norms relating to land inheritance, which reflect the ethnic heritage of the two communities.

Greif (1994), relying on game theory and historic evidence, examines the historical origins of the divergent evolution of collectivist and individualist cultures among the Genoese and Maghribi, respectively. Relying on qualitative historical evidence, Greif argued that the differences between the two groups have their origins in the varying enforcement strategies undertaken by merchants toward overseas agents during long-distance medieval trade. Among the Maghribi, merchants relied on a collective enforcement strategy, where all merchants collectively punish any agent who had cheated. Among the Genoese, enforcement was achieved through an individual punishment strategy (Greif 1993).

Although these pioneering studies provide scarce evidence supporting the role of norms as a channel through which history matters, the shortcoming of each study is that only two groups are being compared. Although the studies by Cohen et al. (1996), Salaman (1980), and Greif (1994) carefully document differences between the two groups being examined, they are unable to prove that the source of these differences is the historical channel being examined. In all three examples, the two groups being examined differ along many dimensions other than the historic dimension under consideration. Because of this, one cannot conclude that the cultural differences observed today do not arise as a result of other factors.

More recent studies attempt to use the same comparative logic of these early studies, but they extend the sample to include more observations. By doing this, more sophisticated statistical techniques can be employed to help identify with greater certainty the importance of history in determining current cultural norms of behavior. Guiso et al. (2008) empirically examined the well-known hypothesis put forth by Putnam et al. (1993) that within Italy city states that became independent during the 1000–1300 period developed higher levels of social capital and that these higher levels of social capital continue to persist. The authors bring the hypothesis by Putnam et al. to the data by collecting various city-level measures of social capital. Looking across 400 Italian cities, they showed a positive relationship between their measures of social capital and whether the city was free in 1176.

The core issue when interpreting this correlation is that these independent cities may also be different along other dimensions that may also affect social capital today. In an attempt to overcome this identification issue, the authors rely on historically motivated instrumental variables. Motivated by the historic fact that, after the collapse of the Holy Roman Empire, local autonomy was often formed around existing religious authorities, the authors used whether a town had a local bishop in the fifth century as one instrument.
Their second instrument is motivated by the authors’ argument that a city that was a part of the Etruscan civilization during the eighth century was more likely to have become an independent city during the Middle Ages. Using these instruments, the authors provided IV estimates that are consistent with their OLS estimates.

The authors also used the historic presence of the Norman Kingdom in southern Italy, which prevented these cities from gaining independence, to develop a falsification test to assess the validity of their instruments. Using the first-stage estimates from their IV procedure, the authors derived, for each southern city, the predicted probability of its being free if it was not located in the south. They showed that, among these southern cities, this predicted probability is uncorrelated with their measures of social capital. The authors argued that this result provides suggestive evidence that the exclusion restrictions for their IV estimates are satisfied. Having a local bishop during the Holy Roman Empire and being an Etruscan city do not appear to have an effect on social capital in southern Italy, where their potential effect through independence was shut down historically. If the relationship between their instruments and social capital was because of a correlation with factors other than independence, then one would also expect to find a similar relationship in southern Italy.

The recent study by Tabellini (2007) also considers the historical origins of norms of behavior, but it is interested in using this relationship to identify the causal effect of norms on economic development. His analysis examined differences in trust in others, respect for others, and confidence in the benefit of individual effort across regions within Europe. Using an IV estimation strategy to isolate exogenous variation in cultural norms, Tabellini used two historically based instruments: (a) the literacy rate at the end of the nineteenth century and (b) the political institutions in place over the past several centuries. According to the first-stage estimates, European regions with historically lower literacy and less well-developed institutions today have less trust in others, less respect for others, and less confidence in the benefit of individual effort. The second stage of the IV estimates reports a positive effect of the measured norms on current per capita income levels and average annual growth between 1977 and 2001.

Although the instruments are relevant, it is less clear whether the necessary exclusion restrictions are satisfied. The historic variables may have affected the evolution of factors other than cultural norms. Tabellini clearly understands this concern and controlled for a variety of factors including current literacy, current political institutions, and the historic urbanization rate. He also controlled for the capital stock and the sectoral composition of employment in the 1970s, arguing that these are also possible channels through which the historic variables may affect current economic development. The extent to which one takes the estimates as causal rather than correlations observed in the data is a matter of individual judgement based on the plausibility of the identification strategy.

Like Tabellini (2007), Nunn & Wantchekon (2009) also considered the historical determinants of trust, but in a very different environment. The authors examined whether the Trans-Atlantic and Indian Ocean slave trades were responsible for a culture of mistrust within Africa. This is done by combining survey data from the 2005 Afrobarometer with estimates of the number of slaves taken from each ethnic group in Africa. The study found a very strong negative relationship between an individual’s reported trust in others and the number of slaves taken from the individual’s ethnic group during the slave trades. As always, the core concern with this correlation is the nonrandom nature of the selection of the historic event. To better understand whether the relationship is causal, the authors...
instrument for slave exports using the historic distance from the coast of each individual’s ethnic group, while controlling for each individual’s current distance from the coast. The IV results confirm the OLS estimates. They also performed a number of falsification tests, the results of which suggest that the instrument’s exclusion restrictions are likely satisfied.

The study then attempted to distinguish between the two most plausible channels through which the slave trade could have adversely affected trust. In one channel, the slave trade altered the cultural norms of the ethnic groups exposed to the trade, making them inherently less trusting. In a second channel, the slave trade resulted in a long-term deterioration of the legal and political institutions, and these weak institutions cause individuals to be less trusting of others. To distinguish between these two channels, the authors constructed a second measure of slave exports: the average number of slaves that were taken from the geographic location in which each individual is currently. This is different from the first measure, which is the average number of slaves taken from an individual's ethnic group.10

Identification between the two channels is based on the fact that, in relocating, individuals take their internal norms with them, but leave their external institutional environment behind. In other words, institutions, which are external to the individual, are much more geographically fixed, relative to cultural beliefs, which are internal to the individual. Therefore, if one accepts that the slave trade had a causal effect on trust, then the two variables can be used to distinguish the extent to which the slave trade affects trust through both the culture and institutions channel. If the slave trade affects trust primarily through internalized norms and cultural beliefs, which are ethnically based and internal to the individuals, then, when looking across individuals, what should matter is whether their ancestors were heavily enslaved. If the slave trade affects trust primarily through its deterioration of domestic institutions, which are external to the individual and geographically immobile, then what should matter is whether the external environment in which the individual is living was heavily affected by the slave trades.

The empirical results provide evidence consistent with the slave trades adversely affecting trust through both cultural norms and institutions, but the magnitude of the culture channel is approximately twice the magnitude of the institutions channel. This result is significant because it provides one of the first pieces of evidence attempting to quantify the impact of an historic event through its effect on both institutions and culture.

Although there has been a recent surge in papers seeking to test whether norms of behavior are one channel through which history matters, the literature is still very young and far from developing a clear understanding about norms. For example, we still do not understand a number of fundamental and important issues related to history, norms, and economic development: In what environments do norms change and when do they tend to persist? Which norms are more persistent and which are more prone to change? What is the relationship between norms and institutions? How do the two interact? Are they complements or substitutes?

4.4. Knowledge and Technology

A number of studies have hypothesized that knowledge, education, and technology may also be channels through which historic events have long-term impacts. Glaeser et al.

---

10Both measures are averages over a given land area. Because the two variables are measured in the same units, if the individual lives in the same location as did his ancestors, then the two variables will take on the same value.
(2004) conjectured that part of the positive relationship between European settlement and economic growth documented by Acemoglu et al. (2001) may reflect the knowledge and “know how” brought by the settlers to the colony. Although their study is unable to provide causal evidence proving their hypothesis, it does highlight correlations in the data that are consistent with their view. Using OLS estimates, the authors showed that, across countries, school enrollment in 1900 is positively correlated with income in 2000 and that a country’s initial settler mortality rate is strongly correlated with schooling in 1900. The authors also examined panel data over five-year intervals between 1960 and 2000. They found a strong positive relationship between initial education and the subsequent five-year change in democracy [for additional evidence on the possible link between education and democratic institutions, see Glaeser et al. (2006)]. However, in a subsequent study, Acemoglu et al. (2005b) showed that this result is driven by the worldwide increase in average education and democracy that has occurred between 1960 and 2000. Once one includes time-period fixed effects, then the correlation disappears.

Other studies also provide additional evidence consistent with education being a channel through which historic events affect long-term development. A recent study by sociologist Robert Woodberry (2004) documents correlations between measures of the historic presence of missionaries and current per capita income and democracy across former nonsettler colonies. Woodberry argued that the historical presence of missionaries, particularly Protestant missionaries, resulted in a number of benefits for the native populations, such as increased education and a decrease in injustice and abuse by the colonizers. Because these abuses angered the indigenous populations, making missionary work more difficult, missionaries had an incentive to fight against these injustices. According to Woodberry, these consequences in turn served to promote democracy after independence. Additional examples include Huillery’s (2008a) finding that, in French West Africa, differences in colonial investments in schooling can still be observed today, which provides some evidence consistent with a persistent impact of education. Similarly, Bolt & Bezemer (2008) showed that there is a positive correlation between colonial education levels and current income across African countries.

As Glaeser et al. (2004) acknowledge, their analysis (as well as the other studies cited here) is unable to prove the primacy of education as a fundamental factor affecting long-term development. The primary difficulty is establishing a causal link. Both the demand for and the supply of education are potentially endogenous to a large number of factors, including the quality of domestic institutions. Countries with better institutions also tend to have better governments that provide a higher level of public goods, including education. In addition, countries with better institutions tend to have more secure property rights, which raises the returns to investments, including human capital investments. Therefore, correlations between colonial education and current economic development may be driven by other omitted factors.

Although these studies focus on the transfer of education and knowledge from Europe to the colonies, historically, knowledge is also transferred in the opposite direction. An example is the Columbian Exchange. A study by Nunn & Qian (2008) examines one part of this exchange: the introduction of the potato to the Old World from the New World. Their study estimates the impact of the new food technology on population growth and development in the Old World. The authors showed that the potato was calorically and nutritionally superior to all Old World crops, including wheat, barley, rye, and rice. Because of this, for the parts of the Old World that were able to adopt the potato, its
diffusion from the New World resulted in a large positive shock to agricultural productivity. The authors used a difference-in-differences estimation strategy and compared the difference in population growth before introduction relative to the population growth after introduction between countries that were able to adopt potatoes and those that were not. A country’s ability to adopt potatoes is measured using climate and soil data based on geographic information systems taken from the Food and Agricultural Organization. According to their estimates, the introduction of the potato had a very large positive impact on population growth, accounting for approximately one quarter of the observed increase in population after 1700. The authors also found that the introduction of the potato spurred economic development, which they measured by the urbanization rate, but that this effect is felt approximately 100 years after the increase in population growth.

The recent study by Comin et al. (2007) provides a unique analysis of the possible historical persistence and importance of knowledge and technology. The authors documented a positive relationship between historic technology levels (as far back as 1000 BC) and current income per capita across different parts of the world. The authors interpreted their findings using a model where the stock of existing technology decreases the cost of adopting new technologies. They showed that if a society has a more advanced technology in the past, then, because of lower costs to technology acquisition, this will increase its stock of technology and income today. Therefore, according to the model, past technology levels can have a direct causal impact on current economic development. However, other interpretations of their correlations are also possible. The most obvious is that the relationship between historic technology and current technology is being driven by some omitted factor, which tends to persist over time. Viewed in this light, their correlation is not surprising, especially given that previous studies have also documented a similar persistence in economic and political performance over time. For example, Bockstette et al. (2002) empirically documented a positive relationship between state antiquity and current economic performance. The persistence of technology may simply be another manifestation of the persistence of economic performance over long periods of time.

5. THE RELATIONSHIP BETWEEN HISTORY AND GEOGRAPHY

An alternative determinant of economic development that is often pitted against history is geography. Proponents of the geography view argue that unchanging fixed geographic factors, such as ecology, climate, natural endowments, and the disease environment, are the primary determinants of long-term economic development. This view has been emphasized by a number of studies that highlight strong statistical correlations between economic development and a variety of geographic characteristics, such as the climate (Kamarck 1976), the disease environment (Sachs et al. 2001, Sachs & Malaney 2002), natural openness (Rappaport & Sachs 2003), and resource endowments (Sachs & Warner 2001).

An example of the tension between the “geography matters” and the “history matters” views can be seen if we return to the examination by Acemoglu et al. (2001) of the effects of institutions on long-term growth.11 In their study, because initial settler mortality is used as an instrument for domestic institutions to explain the causal effect of institutions on income, the exclusion restriction that must be satisfied is that initial settler mortality

11This debate often takes the form of “history matters” versus “institutions matter.” See, for example, Rodrik et al. (2004).
affects only current income through its historic effect on institutions. Sachs (2003), a proponent of the geography view, argued that the IV estimates are biased because there is a direct and persistent effect of geographic characteristics on income today. The argument is that areas with high settler mortality during colonial rule continue to have deadly disease environments, resulting in low income levels today. Therefore, initial settler mortality is correlated with income through a channel other than institutions and the exclusion restriction is not satisfied.

Acemoglu et al. (2001) understood this concern and addressed it directly in their study. They argued that the primary diseases that killed Europeans were malaria and yellow fever and that these diseases have a limited effect on indigenous populations that have developed immunities to the diseases. Because of native immunities, it is unlikely that European settler mortality rates are highly correlated with the disease burdens faced by native populations either historically or today. The authors also showed that their results remain robust to the inclusion of a host of controls that measure geographic characteristics and the local disease environment.

Another piece of evidence of the relative importance of history versus geography has been put forth by Acemoglu et al. (2002), who documented a “reversal of fortunes” among countries that were colonized. The countries that were the most prosperous in 1500 are the poorest countries today.12 They argued that if geography is a fundamental determinant of long-term economic development, then we would not expect this drastic change in the relative prosperity of countries over time.

Part of the reason for the debate lies in the fact that geography affects human actions in the past as well as today. In other words, in addition to affecting income directly, geography also influences history, which in turn affects current income. Many examples show that small and seemingly innocuous geographic differences become magnified through historic events and as a result end up having large impacts on long-term economic development. One example is the differences in soil and climate that made plantation agriculture and its reliance on slavery more or less profitable in different parts of the Americas (Engerman & Sokoloff 1997, 2002). Even more dramatic examples of the effect of geography through history come from Diamond’s (1997) book Guns, Germs and Steel. The book is devoted to exploring the answer to the question of why Europeans colonized the rest of the world and not the other way around. The proximate answer to this question is clear: Europeans had superior technology, such as guns, swords, and ships. They also brought with them germs that killed large numbers of native populations, particularly in the Americas and Australia. But why did Europeans have superior technology? And why did their germs kill native populations and not the other way around?

Diamond’s answer to the first question is that Europe’s technological advantage, which stemmed from more complex specialized societies, arose because crops and animals were domesticated earlier and in more varieties than in other parts of the world. Among the nine locations where food production initially arose, the Fertile Crescent was the first and contained the most domesticated species. Despite its name, the Fertile Crescent was not

---

12Huillery (2008b) tested for a reversal of fortunes within districts of former French West Africa. She found no evidence of a reversal at this increasingly micro level. Instead, she found that Europeans tended to settle in the initially more prosperous parts of French West Africa and European settlement exerted a strong positive effect on long-term development.
the most fertile region on Earth. Instead, its advantage lay in its having the largest number of species of wild plants suitable for domestication. Similarly, the Fertile Crescent, and Eurasia more generally, also had the largest endowment of animals that could be domesticated.

The domestication of plants and animals quickly spread east and west throughout Eurasia, but it diffused much less quickly south to the African continent. This is because when moving east or west, the length of the day does not change, and the climate does not drastically differ. However, this is not true when moving north or south, where the length of the day changes and the climate may significantly differ. More generally, for continents with a north-south orientation, such as the Americas or Africa, domestication or technological advance tended not to spread as quickly as in Eurasia with its east-west orientation. Because of the early domestication of animals in Eurasia, humans lived in close proximity to animals. As a result, new animal-based diseases, such as measles, tuberculosis, influenza, and smallpox, developed. Over time, genetic resistances to the diseases were developed. The parts of the world without domesticated animals did not develop these diseases or a genetic resistance to the diseases. This explains why European diseases decimated native populations and not the other way around.

Overall, Diamond’s explanation for Europe’s global dominance illustrates clearly the large effects that geography can have through history. The historical origins of European colonization of the globe lie in two deep determinants: (a) being endowed with wild plants and animals suitable for domestication and (b) being located on a continent with an east-west orientation.

Although Diamond’s analysis provides descriptive evidence showing that geography can have large impacts through history, it does not provide a quantitative assessment of the relative importance of the historic effect of geography relative to its direct contemporaneous effect on development [but see Olsson & Hibbs Jr. (2005), whose estimates provide strong support for Diamond’s hypotheses]. A recent study by Nunn & Puga (2007) attempts to estimate the magnitude of these two channels for one geographic characteristic, terrain ruggedness, which is measured as the average absolute slope of a country’s surface area. The study identifies two channels through which terrain ruggedness should matter for economic development. One is a direct contemporaneous effect of ruggedness on income. All else being equal, constructing buildings, roads, bridges, and other infrastructure on rugged terrain is more difficult. Agriculture and irrigation are also more difficult, and trade is more costly. For these reasons, terrain ruggedness is expected to have a direct negative effect on income.

The study also identifies an historic effect of terrain ruggedness. During Africa’s slave trade, societies were able to use rugged terrain to protect and hide from slave raiders and kidnappers. This allowed individuals, villages, and societies to partially defend against the negative effects of the slave trades documented in Nunn (2008a). Thus, for the African continent, which was exposed to the slave trade, ruggedness also had an historic indirect positive effect on income. Ruggedness allowed people in certain areas to escape from the slave trade, thereby increasing long-term economic growth. Exploiting the fact that the slave trade affected only Africa, Nunn & Puga (2007) estimated the magnitude of both effects of ruggedness. They found that the indirect historic effect of ruggedness is consistently twice the magnitude of the direct contemporaneous effect of ruggedness, suggesting that, in terms of this geographic characteristic, the importance of geography through history swamps its importance today.
Overall, the body of evidence reviewed here suggests that the largest effects of geography on current economic development may work through its influence on past events rather than through its direct effect on current economic outcomes.

6. CONCLUSIONS

The empirical literature examining the relationship between history and current economic development has developed considerably in the past years. The main fact established by this literature is that history matters. A wide variety of papers have documented the important effects that certain historic events have had on long-term economic development. Examples of these include the study by Acemoglu et al. (2001) showing that colonial rule affected the subsequent development of domestic institutions and economic development, the study by Banerjee & Iyer (2005) showing the importance of colonial land revenue systems within India, and Nunn’s (2008a) analysis showing that Africa’s slave trade adversely affected subsequent development.

Although the literature has made considerable progress in showing that history matters, less well understood are the exact channels of causality through which history matters. For most of the studies undertaken to date, their greatest shortcomings lie in their inability to identify the exact mechanism or channel of causality. For example, Nunn’s (2008a) study of Africa’s slave trade documents the adverse long-term effects of the slave trade, but it is unable to identify the precise channels and mechanisms at play. Similarly, the analysis by Banerjee & Iyer (2005) is unable to specify the precise channels through which the historical land revenue system in India, established by the British 150 years ago, continues to have an effect long after it was abolished. Even in the analysis by Acemoglu et al. (2001), which identifies a broad cluster of institutions as the intervening mechanism, this broad cluster remains, by and large, a black box left to be unpacked. Although the follow-up paper by Acemoglu & Johnson (2004) takes initial steps in this direction, much remains to be done before this historic channel of causality is clearly understood.

The most recent studies that rely on microlevel data and analysis (e.g., Iyer 2007, Acemoglu et al. 2008, Dell 2008, Huillery 2008a, Nunn & Wantchekon 2009) are beginning to identify finer causal factors and more precise mechanisms. However, because these studies rely on data at a much more micro level, the scope of their analysis is more limited. For example, it is difficult to know if the effects of the mita forced-labor system convincingly identified by Dell (2008) are similar to the effects of other coercive labor systems in other parts of the world. For this reason, as studies become increasingly focused, evidence from complementary studies from other locations and times is also needed before one can assess whether specific results are part of broad systematic patterns that exist in the data or whether the results are specific to a given environment. For this reason, although this literature has grown greatly in recent years, many historic events and mechanisms need to be examined before we have a clear picture of the overall importance of historic events and the specific channels through which they continue to affect economic development today.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.
ACKNOWLEDGMENTS

The JEL classification codes for this article, prepared for the Annual Review of Economics, are as follow: H11, N00, O10, 050, P51, R58, Z13. I have benefited from useful discussions with Daron Acemoglu, Robert Bates, Joseph Henrich, Joel Mokyr, and James Robinson.

LITERATURE CITED


Nunn N. 2007a. Historical legacies: a model linking Africa’s past to its current underdevelopment. *J. Dev. Econ.* 83:157–75


Nunn N, Qian N. 2008. *Columbus’s contribution to world population and urbanization: a natural experiment examining the introduction of potatoes.* Mimeogr., Harvard Univ.


Contents

Some Developments in Economic Theory Since 1940: An Eyewitness Account
   Kenneth J. Arrow ............................................. 1

School Vouchers and Student Achievement: Recent Evidence and Remaining Questions
   Cecilia Elena Rouse and Lisa Barrow.............................. 17

Organizations and Trade
   Pol Antrás and Esteban Rossi-Hansberg ......................... 43

The Importance of History for Economic Development
   Nathan Nunn ..................................................... 65

Technological Change and the Wealth of Nations
   Gino Gancia and Fabrizio Zilibotti ............................ 93

CEOs
   Marianne Bertrand ............................................. 121

The Experimental Approach to Development Economics
   Abhijit V. Banerjee and Esther Duflo ......................... 151

The Economic Consequences of the International Migration of Labor
   Gordon H. Hanson ............................................. 179

The State of Macro
   Olivier Blanchard ............................................. 209

Racial Profiling? Detecting Bias Using Statistical Evidence
   Nicola Persico ................................................. 229

Power Laws in Economics and Finance
   Xavier Gabaix ................................................. 255

Housing Supply
   Joseph Gyourko ................................................. 295
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantitative Macroeconomics with Heterogeneous Households</td>
<td>319</td>
</tr>
<tr>
<td><em>Jonathan Heathcote, Kjetil Storesletten, and Giovanni L. Violante</em></td>
<td></td>
</tr>
<tr>
<td>A Behavioral Account of the Labor Market: The Role of Fairness Concerns</td>
<td>355</td>
</tr>
<tr>
<td><em>Ernst Fehr, Lorenz Goette, and Christian Zehnder</em></td>
<td></td>
</tr>
<tr>
<td>Learning and Equilibrium</td>
<td>385</td>
</tr>
<tr>
<td><em>Drew Fudenberg and David K. Levine</em></td>
<td></td>
</tr>
<tr>
<td>Learning and Macroeconomics</td>
<td>421</td>
</tr>
<tr>
<td><em>George W. Evans and Seppo Honkapohja</em></td>
<td></td>
</tr>
<tr>
<td>Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods</td>
<td>451</td>
</tr>
<tr>
<td><em>Raj Chetty</em></td>
<td></td>
</tr>
<tr>
<td>Networks and Economic Behavior</td>
<td>489</td>
</tr>
<tr>
<td><em>Matthew O. Jackson</em></td>
<td></td>
</tr>
<tr>
<td>Improving Education in the Developing World: What Have We Learned from Randomized Evaluations?</td>
<td>513</td>
</tr>
<tr>
<td><em>Michael Kremer and Alaka Holla</em></td>
<td></td>
</tr>
<tr>
<td>Subjective Probabilities in Household Surveys</td>
<td>543</td>
</tr>
<tr>
<td><em>Michael D. Hurd</em></td>
<td></td>
</tr>
<tr>
<td>Social Preferences: Some Thoughts from the Field</td>
<td>563</td>
</tr>
<tr>
<td><em>John A. List</em></td>
<td></td>
</tr>
<tr>
<td><strong>Errata</strong></td>
<td></td>
</tr>
<tr>
<td>An online log of corrections to <em>Annual Review of Economics</em> articles may be found at <a href="http://econ.annualreviews.org">http://econ.annualreviews.org</a></td>
<td></td>
</tr>
</tbody>
</table>