The causal history of Africa: A response to Hopkins

Fenske, James

Yale University

August 2010

Online at https://mpra.ub.uni-muenchen.de/24458/
MPRA Paper No. 24458, posted 25 Aug 2010 06:26 UTC
THE CAUSAL HISTORY OF AFRICA: A RESPONSE TO HOPKINS

JAMES FENSKE†

ABSTRACT. In a recent paper for the Journal of African History, A.G. Hopkins writes that economists have spent the last decade writing a “new” economic history of Africa that has escaped the notice of historians. He labels the “ethnolinguistic fractionalization” and “reversal of fortune” theses as this literature’s key insights. I argue that the most valuable contributions to the new economic history of Africa are not distinguished by their broad theories, but by their careful focus on causal inference. I survey recent contributions to this literature, contrast them with the “old” economic history of Africa, and revise Hopkins’ advice to historians accordingly.

1. THE “OLD” AND “NEW” ECONOMIC HISTORIES OF AFRICA

In the latest edition of the CIA World Factbook, 25 of the 30 poorest countries on Earth are African. Using its 2009 PPP-adjusted GDP per capita of $1900 as a yardstick, the North Korean basket case is twice as well off as Ethiopia, Mozambique, Togo, Sierra Leone, or Malawi. It is three times as wealthy as Guinea-Bissau, Somalia, or Liberia, more than six times as rich as Burundi or the Democratic Republic of the Congo, and nearly ten times as productive as Zimbabwe. The Factbook also estimates that 63% of North Korea’s population is urban, versus 48% in Nigeria, 38% in Sierra Leone, 25% in Lesotho, and 10% in Burundi. According to the UN, life expectancy at birth in North Korea is 67.3, just above the world average. In Nigeria it is 46.9. In Swaziland it is 39.6. The UN estimates child mortality under five at 62.4 per 1000 live births in North Korea. More than 40 African countries are doing worse; in Botswana, this rate is 67.5; in Ghana, 89.6; in Liberia, 205.2; in Sierra Leone, 278.1. African economic failure needs explanation.

The economic history of Africa attempts to account for the long-term origins of African poverty. McPhee’s (1926) Economic Revolution in British West Africa is often pointed to as the first academic study of Africa’s economic history. While many of his conclusions remain valid today, his book suffered from the colonial prejudice that African economies were spurred to life by intervention from outside, and so he overemphasized the discontinuity in economic

†DEPARTMENT OF ECONOMICS, UNIVERSITY OF OXFORD, MANOR ROAD BUILDING, OXFORD OX1 3UQ
E-mail address: james.fenske@economics.ox.ac.uk.

Date: August 16, 2010.

I would like to thank my advisors Timothy Guinnane, Benjamin Polak, and Christopher Udry for their guidance. I am also grateful for the feedback of two anonymous referees. This paper was originally circulated under the title “Institutions in African History and Development: A Review Essay.”
change brought about by the first twenty years of British rule in West Africa. Fage (1971) surveys writing on the history of West Africa that occurred between the start of the colonial period and the establishment of history departments in West African universities from 1948 onwards. During these years, professional imperial historians, social anthropologists, and other white authors had practical reasons for studying the histories of their subject peoples, since knowledge was necessary for indirect rule to function and materials were needed for classrooms in the colonies.

For most Africanists, though, it is Dike’s 1956 *Trade and politics in the Niger Delta* that marks the beginning of the study of historical change in African economies. Political and economic history were favored subjects of the early pioneers of the academic study of African history – Fage (1955), Biobaku (1957), and Davidson (1961) (among others) gave substantial attention to economic questions. Throughout the 1960s, 1970s, and 1980s, historians such as Ralph Austen, Anthony Hopkins, Martin Klein, Robin Law, Paul Lovejoy, Suzanne Miers, Richard Roberts, John Thornton, and Paul Zeleza wrote Africa’s economic history. While some, notably David Eltis, Jan Hogendorn, and Patrick Manning, were economists or employed tools of modeling, regression and simulation, this was largely an “old” economic history, relying on qualitative accounts and pulling together sparse sources of descriptive statistics to uncover the workings of African economies in the past.

In a recent paper for the *Journal of African History*, Hopkins (2009) writes that the economic history of Africa by the close of the 1980s “was in failing health; in the 1990s its public appearances were limited.” He explains this by invoking the general decline of economic history and the rise of postmodernism. Historians have turned to focus more on social and cultural history than on economic subjects. It is not that there has been no work since 1990; the collected studies in Law (1995) and the synthesis by Lynn (1997) have added to our knowledge about economic transformation between the abolition of the Atlantic slave trade and the advent of colonial rule. Austin (2005) has provided an economic history of Asante. In many cases, however, the “old” economic history has been subsumed within other works that focus on the economic, social and political histories of particular places (e.g. Mann (2007) on Lagos) or subjects (e.g. Klein (1998) on slavery in French West Africa). Notably, the relatively new literature on African environmental history (Beinart, 2000; Fairhead and Leach, 1996; Harms, 1987; Kreike, 1996; McCann, 1999) is inseparable from the economies of the peoples who have transformed the continent’s landscape.

Hopkins (2009) notes that, while historians have turned from the subject, economists have spent the past decade writing “a new type of economic history of Africa.” While this is true, Hopkins (2009) errs in identifying the “reversal of fortune” thesis of Acemoglu et al. (2002) and (apparently) Nunn (2008) and the “ethnolinguistic fractionalization” thesis of Easterly and Levine (1997) as the twin forefronts of this literature. While his focus is on
“big ideas,” what originally set the “new” economic history apart was its particular focus on quantitative methods. In the past decade, a collection of careful econometric studies have made incremental and credible contributions to our understanding of African economic history.

One branch of this “new” economic history has focused on uncovering causal relationships, using the same research designs (especially natural experiments, panel data methods, matching, instrumental variables, and regression discontinuity) that Angrist and Pischke (2010) have called the “credibility revolution” in empirical economics. Huillery (2009) has used matched pairs of districts within French West Africa to find robust relationships between colonial-era investments in health and education and modern outcomes. Nunn (2008) employs distance from New World slave markets as an instrument for the number of slaves exported by each African country, to show that the slave trade hinders modern growth. Bubb (2009) has used regression discontinuity to search for differences in policy outcomes on opposite sides of the border between Ghana and Côte d’Ivoire. These papers and other similar contributions are those to which Hopkins (2009) should have looked in seeking the “new” economic history of the African continent.

These contributions have gone unnoticed by historians of Africa because their methods and topics do not speak to recent trends in history, because the culture of circulating working papers common to economics is rare in history, and because many of these studies do not see themselves as part of the tradition that leads from McPhee (1926) to the present. Instead, they situate themselves within the economic literature that uncovers the long term causal effects of historical events and institutions, recently surveyed by Nunn (2009b). This includes work by Banerjee and Iyer (2005) and Iyer (2008) on the long reach of colonial rule in India, or Dell (2009) on the mita system in Peru and Bolivia.

In this review essay, I draw these “old,” “new,” and causal economic histories of Africa together. I argue that studies with careful research designs and a focus on causation have been the most credible recent contributions to the “new” economic history of Africa. Rather than counseling historians to undertake context-specific case studies that engage with the big ideas of the “new” economic history of Africa, Hopkins’s (2009) advice should have been for historians to draw on their knowledge of specific contexts to find cases where it is possible to uncover plausible causal mechanisms. These can tell us how history matters for modern African development or how institutions and economies functioned in the past. These are the sort of results can be extrapolated to other contexts in order to inform development policy.

The rest of this paper proceeds as follows. In Section 2, I describe some of the econometric methods that set causal history apart from other contributions and give examples of these from outside of African history. In Section 3, I discuss the first of the two big
ideas Hopkins (2009) reviews – ethnolinguistic fractionalization. While he is right to suggest that economists and economic historians have not done enough to understand the origins of ethnicity, I argue that a collection of empirical studies of the origins and consequences of fractionalization does exist. In Section 4, I look at institutions and the “reversal of fortune” thesis, the second major theme of his review. I argue that Hopkins (2009) overstates the degree to which this literature pushes a single, cohesive narrative, but understates the robustness of this literature’s conclusions. In particular, the critiques he makes can best be understood as suggestions for future work. I also outline empirical studies of the operation and implications of historical institutions in Africa. In Section 5, I show that causal historical studies of Africa have made contributions outside of these two theses, in particular to our understanding of how geography, the slave trades, colonial rule, and trade have shaped African development. In Section 6, I conclude.

2. Methods of causal history

In 1983, Edward Leamer published an article titled “Let’s take the con out of econometrics.” He charged that empirical economic studies lacked robustness, since published results were sensitive to the choice of specification. Since then, Angrist and Pischke (2010) argue that empirical economists have had better data, have been less distracted by unimportant questions of functional form, and have increasingly focused on better research designs. This has created a literature based on credible causal inference. Economic history too has been transformed by this change in methods. Thirteen of thirty-two papers published in the Journal of Economic History in 2009 employed either panel data methods or instrumental variables to lend credibility to their causal claims. Two more (Vizcarra (2009) and Pereira (2009)) study the impacts of quasi-experimental shocks – the Peruvian government’s credit commitments made possible by guano exports and the 1755 Lisbon earthquake – even if their approach is not econometric. Nunn (2009b) refers to these “more satisfying identification strategies” as one of the major developments in recent studies of the importance of history for economic development.

Practitioners of causal history focus on identifying causal historical relationships, which sets them apart both from the qualitative “old” economic historians and from much of the “new” economic history. In this section, I briefly outline the most common methods used in this literature to identify the causal relationships of interest – regression analysis, panel methods, instrumental variables, and regression discontinuities – and give examples from the causal history literature outside of Africa.¹

¹This is intended only as a very brief introduction. Angrist and Pischke (2009), Imbens and Wooldridge (2009), Wooldridge (2002) and Angrist and Krueger (1999) are indispensable guides to these techniques. I have also relied heavily on Christopher Blattman’s syllabus on “Research Design and Causal Inference” in preparing this section.
2.1. Regression. Causal research designs generally begin with a linear hypothesis:

\[ y_{it} = \alpha + \beta x_{it} + \gamma_1 c_{it}^1 + \ldots + \gamma_K c_{it}^K + \epsilon_{it}, \]

where \( y_{it} \) is the effect of interest, observed for unit \( i \) in period \( t \), \( x_{it} \) is the possible cause of interest, \( c_{it}^1, \ldots, c_{it}^K \) are other controls that may potentially also affect the outcome \( y_{it} \), and \( \epsilon_{it} \) is “error,” which includes other causes of \( y_{it} \) that the researcher does not observe. The goal of causal research design is to uncover an unbiased estimate of \( \beta \), the effect of \( x_{it} \) on \( y_{it} \). It is only under the very restrictive assumption that \( x_{it} \) and \( \epsilon_{it} \) are uncorrelated that simply estimating (1) using ordinary least squares will yield an unbiased estimate of \( \beta \). There must be no reverse causation, and no unobserved causes of \( y_{it} \) in \( \epsilon_{it} \) that are correlated with \( x_{it} \).

Chaney (2010) is one such example; he shows that deviations in the year-to-year flood of the Nile away from its ideal value increased the probability of a change in the Sultinate and lowered the probability that the judgeship changed hands in pre-modern Egypt. This approach is also valid with natural experiments. Chaney (2008) argues that the 1609 expulsion of 300,000 Muslims from Iberia created exogenous variation in institutions; lords established extractive institutions in those areas formerly populated by Muslims, but not those inhabited by Christians. He regresses the share of the active population engaged in agriculture in 1787 on a dummy for whether a municipality was Muslim in 1609 to show that these extractive institutions retarded the growth of the non-agricultural sector.

2.2. Fixed effects. One strategy to eliminate the bias that results from correlations in \( \epsilon_{it} \) and \( x_{it} \) is to use panel data in which each unit \( i \) is observed during multiple periods. Fixed effects for \( i \) and \( t \) will remove bias due to correlation between \( x_{it} \) and unobserved time-invariant heterogeneity or time-specific errors that affect all units equally. The estimating equation becomes:

\[ y_{it} = \alpha + \beta x_{it} + \gamma_1 c_{it}^1 + \ldots + \gamma_K c_{it}^K + \mu_i + \delta_t + \epsilon_{it}, \]

where \( \mu_i \) and \( \delta_t \) are vectors of dummy variables for each unit \( i \) and time period \( t \). Only variation within individual units across time is used for identification. For example, Naidu (2010) includes both state and year fixed effects in estimating the impact of fines on making offers to those already employed on sharecropper mobility, tenancy choice, and agricultural wages in the post-bellum US south. Naidu and Yuchtman (2009), similarly, include county and year fixed effects when demonstrating that positive shocks to labor demand drove Master and Servant prosecutions in England before 1875, relying on an interaction between industry-level output prices and county-level measures of industry presence for identification. Acemoglu et al. (2009a), study the effect of the French revolution on later European
urbanization by regressing urbanization rates on interactions of year dummies with the duration of French occupation as well as country fixed effects. Cantoni (2009), similarly, in assessing the impact of the Protestant reformation on economic growth, regresses city populations on Protestant-times-year interactions and both city and year fixed effects. Hornbeck (2009) regresses changes in county-level land values between 1930 and various later years on measures of Dust Bowl erosion and state-year fixed effects in order to identify the long term effects of the Dust Bowl solely using variation across counties within a single state in a single year.

Related to the use of fixed effects are matching estimators. A variety of parametric and non-parametric techniques exist for matching as a means of purging unobserved heterogeneity (e.g. Imbens (2004)). The basic approach is to match observations that are otherwise similar into pairs and add a vector of dummy variables $\theta_{pt}$ for each pair to (1), so that only variation in $x_{it}$ and $y_{it}$ within each pair is used to identify $\beta$. For example, Naidu (2009) uses only variation in state-level poll taxes and literacy tests within pairs of border-adjacent counties to identify the impact of voting restrictions on black political participation, schooling outcomes, and land values in the post-bellum US south. Moser (2005), similarly, constructs synthetic matches in her study of the impact of national patent laws on innovations recorded in nineteenth century worlds fairs.

2.3. **Instrumental variables.** An conceptually clean approach to overcoming correlation between $\epsilon_{it}$ and $x_{it}$ is the use of an instrumental variable, that predicts $x_{it}$ but is not correlated with $\epsilon_{it}$. A system of equations is now estimated:

\[
y_{it} = \alpha + \beta \hat{x}_{it} + \gamma_1 c_{1it} + \ldots + \gamma_K c_{Kit} + \epsilon_{it},
\]

\[
x_{it} = \lambda + \pi z_{it} + \tau_1 c_{1it} + \ldots + \tau_K c_{Kit} + \eta_{it}.
\]

In (3), $z_{it}$ is the instrument, a source of “exogenous” variation in $x_{it}$. In (2), $x_{it}$ has been replaced with $\hat{x}_{it}$, the values of $x_{it}$ predicted by estimating (3). The critical assumption is that $z_{it}$ affects $y_{it}$ only through its impact on $x_{it}$, and is not correlated directly with $\epsilon_{it}$. Boustan et al. (2010) use the generosity of New Deal programs and extreme weather events as instruments for migrant flows into and out of US cities in assessing the labor market impacts of migration during the Great Depression. Boustan (2010), similarly, uses economic conditions in the South and past settlement patterns to instrument for black movements into northern US cities, showing that white flight into the suburbs was indeed a response to black migration. Meng and Qian (2009) use suitability for grain cultivation as an instrument for childhood exposure to China’s Great Famine in measuring the famine’s effect on adult health, education, and labor market outcomes; they justify this strategy by arguing that China’s
institutional procurement system meant that traditional grain-growing regions were hardest hit by the famine.

2.4. Regression discontinuity. Another strategy for identifying $\beta$ is to exploit discontinuous breaks in $x_{it}$ that occur once some continuous variable has passed a specified value, on the assumption that $\epsilon_{it}$ is similar on both sides of this boundary. The classic application of this method is Thistlethwaite and Campbell (1960), who use test score cutoffs to uncover the impact of scholarships on career aspirations. There are many types of regression discontinuity design (e.g. Imbens and Lemieux (2008), Lee and Lemieux (2009)), but a parsimonious representation can be given by using $D_{it}$ to represent distance from the cutoff in the case where $x_{it}$ is a dummy variable for a discrete treatment. In this case, $\beta$ can be recovered by estimating:

$$y_{it} = \alpha + \beta x_{it} + f(D_{it}) + x_{it} \ast g(D_{it}) + \gamma_1 c_{it}^1 + \ldots + \gamma_K c_{it}^K + \epsilon_{it},$$

where $f$ and $g$ are polynomial functions of $D_{it}$. This method has been under-employed in causal history, but examples do exist. Oreopoulos (2006) uses changes in the school leaving age in Great Britain and Northern Ireland in 1947 and 1957, after which the proportion of 14-year olds leaving school dropped discontinuously, to identify the returns to education despite a long-term secular decline in the proportion of teenagers leaving school. Dell (2009) uses a similar procedure to compare districts on either side of the historical boundary within which it was compulsory to provide forced mining labor before 1812 in Peru and Bolivia. Districts just within the catchment area have worse outcomes in terms of consumption, health and public goods today.

3. Ethnicity and development

African ethnic diversity is unusually high; using the standard index computed from the Soviet Atlas Narodov Mira, the fifteen most fractionalized countries in the world are African. Easterly and Levine (1997) demonstrate that ethnolinguistic fractionalization has a modest negative effect on growth. More importantly, it predicts the under-provision of public goods and the adoption of bad policies. Posner (2004a) shows that this effect becomes even stronger when the measure is restricted to politically relevant groups. This holds even within countries; schools in the more ethnically diverse parts of western Kenya receive less funding and have worse facilities (Miguel and Gugerty, 2005). Arcand et al. (2000) find that the fractionalization effect is stronger in Africa than in the rest of the world. In this section, I outline the critiques Hopkins (2009) makes of this literature, while showing that more empirical efforts have been made to study the causes and consequences of ethnic diversity than he indicates.
Hopkins (2009) charges that Easterly and Levine (1997) treat their data on fractionalization as “unproblematic,” though he does note how others (e.g. Alesina et al. (2003), Fearon (2003), and Posner (2004a)) have improved on their measure. He suggests that an understanding of the origins of fractionalization is lacking. He points to Englebert’s (2000a) argument that it is the congruence of post-colonial states with pre-colonial institutions, not fractionalization, that matters today as an important contribution in this direction. He suggests that Ahlerup and Olsson’s (2007) theory that ethnic groups emerge to provide public goods, and have been doing so the longest in Africa, is another important step.

Hopkins (2009) advises historians to make contributions to this literature on two directions. First, he writes that they should further existing work that has looked at how ethnicities were constructed during the colonial period. Originally the “invention of tradition” was seen as a top-down, mechanical process, but it is now recognized that ethnicity built on existing institutions and in some cases remained ambiguous. Second, he argues for a resurgence of the “land abundance” view of African history (e.g. Austin (2008), Hopkins (1973), Iliffe (1995)), in which it is argued that sparse population hindered political consolidation, leading to dispersal and fragmentation. If low population densities were geographically determined, geography needs to be part of the story of ethnicity. He also warns historians not to employ functionalist defenses of ethnicity if these are no longer relevant today, and to recognize that ethnicity is only one component of identify. I comment on this advice below.

Hopkins (2009) is right to note that data on fractionalization has been taken as given, when it instead should be explained. Ahlerup and Olsson (2007) make one such contribution. Hopkins (2009) outlines their argument, but neglects to mention that their theory has empirical implications, and that they test these with data. First, their theory predicts that fractionalization should be increasing with duration of first human settlement. Second, greater population density should reduce the fractioning of new ethnic groups. Third, state antiquity should have similarly homogenizing effects. All of these predictions hold in a sample of 191 countries. While these are not strictly causal tests (population density and state antiquity, in particular, may be caused by omitted variables that also affect fractionalization, or may indeed be outcomes of fractionalization), Ahlerup and Olsson (2007) are careful to show that their results are robust to removing influential observations and to alternative measures of ethnic diversity and the duration of human settlement.

Other econometric studies exist of the causes of fractionalization. Michalopoulos (2008) proposes that heterogenous environments create localized human capital, limiting mobility and giving rise to ethnicities and languages. He supports this argument by demonstrating that variability in both land quality and elevation are correlated with human diversity. To overcome the endogeneity of country borders, he uses “virtual” countries as his unit of observation, constructed from a global grid. He can include “real country” and continental
fixed effects to control for omitted heterogeneity. His results for heterogeneous land quality hold even if the sample is restricted to the tropics, or if pairs of adjacent regions are matched together. The cleanest evidence of the causes of ethnic difference comes from randomized trials, and Dunning and Harrison (2010) use experimental evidence in Mali to show that “cousinage” ties that cross-cut ethnicity help explain why ethnic identity is unimportant in that country. Participants in their study rated political speeches more highly if the politician’s surname indicated cousinage with the respondent, even if he was from another ethnic group.

Other empirical investigations of the origins of ethnic diversity have not dealt as carefully with causality, but present suggestive results that await robust testing. Green (2010) conducts a set of exploratory correlations, regressing country-level measures of fractionalization in a sample of African countries on a large set of possible predictors. He finds that the strongest predictors of ethnic diversity in Africa today are temperature, state size, and (negatively) urbanization. Slave exports also predict diversity, though their significance is marginal. Fletcher and Iyigun (2010) show that histories of Christian-Muslim and Sunni-Shi’a conflict from 1400 to 1900 predict greater religious homogeneity today, though there are no similar effects for linguistic and ethnic heterogeneity. Because their data comes as a cross section, their research design does not admit for a clean causal test. Fletcher and Iyigun (2010) do, however, demonstrate that their result is robust to alternative measures and sample periods, and interpret their findings as causal.

Statistical studies have uncovered a significant strategic element to ethnic identification. Habyarimana et al. (2007) conduct an experiment in Kampala. Within their sample, preferences for public goods are uncorrelated with ethnicity. Participants’ offers in an anonymous dictator game are no larger to their co-ethnics than to others. Similarly, co-ethnics performed no better on a cooperative puzzle-solving game than other pairs. In a non-anonymous dictator game, however, their offers are larger to co-ethnics. Eifert et al. (2010) find that “ethnic” identification in the Afrobarometer survey increases at the expense of class and occupational identification as a competitive presidential election nears – ethnicity matters because it is politically salient. Critically, they limit their sample to countries in which more than one round of data are available, so that they can control for country fixed effects, trend, and survey round controls. While they can not control for individual-respondent fixed effects, they show their results are robust to aggregation to the country level.

There have also been several empirical studies of the consequences of fractionalization. Alesina and La Ferrara (2005) suggest that diversity has benefits and costs that operate through individuals’ preferences, their strategies, and through the production function. For them, the benefit of diversity is in facilitating variety in production, an effect most pronounced at already high levels of output. They demonstrate this effect in a cross-country
growth regression, though their causal inference assumes both fractionalization and initial income are exogenous. The cost is that it is difficult to agree on policies and provide public goods. For Knack and Keefer (1997), fractionalization inhibits trust between individuals, which restricts growth. They show that ethnic diversity reduces trust in a cross-country growth regression, though their results again take their ethnic measure as exogenous. Mauro (1995), similarly, finds that fractionalization raises corruption, lowering investment.

Ashraf and Galor (2008) show that genetic homogeneity, predicted by migratory distance from Addis Ababa, has a robust and non-monotonic relationship with development (i.e., population density) in 1500 AD in a cross-country regression.\(^2\) They posit that at low levels of diversity, increasing diversity makes it easier to accumulate complementary human capital and to adopt new technologies. At higher levels, however, greater diversity leads to mis-coordination and distrust. To control for unobservable heterogeneity, they show that their results are robust to including the timing of the transition to agriculture, several other controls, and repeating their analysis with population density estimated in other centuries.

Not all of these studies have relied on cross-country differences. Kimenyi (2006) adds that heterogeneity facilitates corruption in the form of nepotism and patronage through preferential treatment of co-ethnics. Looking at 11 African countries, he shows that politically dominant ethnic groups preferentially receive public goods such as schooling, immunization, roads, and health facilities, though he does take political dominance as a given. Related studies using microeconomic data have looked at the importance of race in individuals’ economic choices and outcomes within African countries. Bigsten et al. (2000) show that African-owned small manufacturers in Kenya are more likely than Asian-owned firms to choose informal over formal status. Fafchamps (2000) finds that black entrepreneurs in Kenya and Zimbabwe face no disadvantages in obtaining bank credit, but do have difficulties acquiring supplier credit. He interprets this to mean that they are penalized in for their lack of network connections with the businesses community. Political decisions also have an ethnic dimension – tribal identification correctly predicts whether respondents voted for Mwai Kibaki in the 2007 Kenyan elections for 51\% of the respondents in Bratton and Kimenyi’s (2008) sample; adding assessment of Kibaki’s performance and respondents’ views on federalism raises this to 64\%.

Other empirical studies have shown that the effects of fractionalization are mediated by policies, context, and institutions. Miguel (2004) suggests that nation-building efforts in Tanzania have been successful at overcoming the negative effects of diversity. Ethnic diversity predicts worse outcomes in Kenyan schools than Tanzanian schools. His evidence is not experimental and the results lack statistical robustness, but his research design makes careful

\(^2\)Ramachandran et al. (2005) show that heterozygosity has a very tight and negative correlation with migratory distance out of East Africa, but their sample size is too small to use heterozygosity directly as a regressor.
use of survey data. These were deliberately collected from similar districts in each country in order create a comparable sample. The mediation of fractionalization effects also holds in cross-country growth regressions; Easterly (2001) demonstrates that including an interaction term between fractionalization and institutions does not eliminate the direct effect of fractionalization on schooling, assassinations, financial depth and several other variables, but does substantially weaken it. He controls for the possible endogeneity of the institutional interaction using the fraction of years independent since 1776 as an instrument. Collier (2003), by contrast, finds that interacting fractionalization with political rights completely eliminates its main effect, though he treats both as exogenous. The policy implications of these results are stark; rather than advocating separation of ethnic groups in order to quell conflict, Habyarimana et al. (2008) advocate strategies such as Nyerere’s imposition of Swahili as the national language in Tanzania that break down barriers to cooperation.

Hopkins (2009) is right to note that both economists and economic historians could make significant contributions to our understanding of the origins of fractionalization. Mamdani (2001), for example, has documented how the identities of Hutu and Tutsi were made rigid as a consequence of Belgian rule. Tignor (1993) demonstrates how British rulers stoked animosity between Northern and Southern Nigerians in the lead-up to independence. Hopkins’s advice to historians to contribute studies of the formation of ethnicity under colonial rule is a positive step, but without a focus on forming testable hypotheses that can be brought to data in multiple contexts, there is a risk that this will result in case studies with no external validity or policy relevance. His second recommendation, that attention be given to the hypothesis that African geographic features limited population densities, restricting state formation and giving rise to fractionalization, is a promising testable hypothesis if a case can be made that some geographic features are valid instruments – that they shape the rise of states only through their effect on population densities.

Some of the most interesting historical work on the formation of ethnicity comes from Daniel Posner. Posner (2003) demonstrates that missionary activity, colonial education, and labor migration consolidated the “Babel of more than fifty languages” in Zambia so that, by independence, nearly 80% of the population spoke one of the four main tongues. He supports his argument for missionaries by showing in a regression that the intensity of colonial missionary education reduces the relative prevalence of linguistic to tribal heterogeneity across districts in 1990. In Malawi, Posner (2004b) argues that Chewas and Tumbukas are adversaries because they are both large enough to form bases for political coalitions. In Zambia, both groups are small, and they are political allies. While he does not present econometric results, he interprets this as the outcome of a natural experiment, in which the border between Malawi and Zambia arbitrarily cut these groups across two countries. Posner (2005) argues that the operation of ethnic cleavages is mediated by institutions. He
constructs a simple model of coalition formation in Zambia, in which voters attempt to form the smallest winning coalition in order to extract resources using the state. Under multiparty democracy, the focus of politics is national, and voters emphasize their linguistic identities. Under single party rule, politics is local, and becomes about “tribal” identities. Posner (2005) supports his conclusions by examining the language used by political candidates and by looking descriptive statistics about where candidates choose to run, rather than econometric evidence. Still, his results support historically-grounded and context-specific hypotheses that could be brought to data in Zambia or other contexts.

4. INSTITUTIONS AND REVERSING FORTUNES

Institutions are, in Douglass North’s famous definition, the humanly devised constraints that structure human interaction. They are made up of formal constraints (rules, laws, constitutions), informal constraints (norms of behavior, conventions, and self-imposed codes of conduct), and their enforcement characteristics. Together they define the incentive structure of societies and specifically economies.

Institutions are believed to matter, because “good” institutions create incentives to invest, while “bad” institutions (including extractive institutions) lower these. Insecure property rights, for example, are held to limit investment because they lower its expected return (e.g. Besley (1995)), because they divert effort towards protecting the investment (e.g. Field (2007)), and can be collateralized to free up funds for investment (e.g. de Soto (2003)), among other reasons. Good institutions also create norms that reduce uncertainty, making it easier to conduct business on a day-to-day level.

The cross-country growth literature on institutions has settled on some broader measures of institutional quality as important drivers of growth – protection against expropriation (Acemoglu et al., 2001), property rights (Acemoglu and Johnson, 2005), rule of law (Rodrik et al., 2004), investor protections and legal origin (Porta et al., 1997), institutions such as guilds and banks that overcome coordination failures (Bardhan, 2005), although surprisingly the net effect of democracy itself on growth may be weakly negative (Barro, 1996; Tavares and Wacziarg, 2001), and important mostly for reinforcing the rule of law (Rigobon and Rodrik, 2005). In Africa, the weakness of the state as an institution has been identified as a major source of policies detrimental to growth (Bates, 1984; Collier and Gunning, 1999). In this section, I outline Hopkins’s (2009) concerns with the literature on institutions, argue that he is conflating the “reversal of fortune” and “institutions matter” views, and review recent studies of the consequences and operation of historical institutions in Africa.

Hopkins (2009), in his discussions of both the “new institutionalist history” and the “reversal of fortune” theory, recognizes that institutions have been a major focus of recent
Figure 1. Institutional quality and economic performance

Institutions are the measure of expropriation risk from Acemoglu et al. (2001). African countries are indicated by a shaded marker. The top panel depicts the results of an OLS regression of economic performance on institutions, while the lower panel shows the results of an IV regression with their measure of settler mortality as an instrument for institutions.
empirical work on African poverty. The reason for this is clear – the quality of modern institutions has a very tight statistical relationship with GDP per capita. Figure 1 replicates Column (1) of Tables (2) and (4) in Acemoglu et al.’s (2001) influential paper. The top panel plots the results of an OLS regression of log per capita GDP in 1995 against their “expropriation risk” measure of institutional quality. The correlation is strong and positive, and the $R^2$ of this regression is 0.62. African countries are indicated with a shaded marker; it is clear from the figure that African countries have both low levels of GDP and poor protection against expropriation. Acemoglu et al. (2001) claim to have uncovered causality, since they instrument for institutions using the mortality of early European settlers. The bottom panel of Figure 1 shows that, if institutions are replaced with the institutions predicted by settler mortality, the relationship remains strong.

Hopkins (2009) treats the “Reversal of Fortune” hypothesis of Acemoglu et al. (2002) – that former colonies that were relatively rich in 1500 have since become poor, and vice versa – as identical to the broader claim that “Institutions Matter.” This is not necessary. Acemoglu et al. (2002) discard Africa from their results that concern urbanization, and show that Africa can be dropped from their results concerning population density without undermining these; the “Reversal of Fortune” need not apply at all to Africa. Further no “reversal” is necessarily inherent in the claim of Porta et al. (1997) that legal origin matters, or the argument of Nunn (2008) that the slave trade hindered African development. Even the argument in Acemoglu et al. (2001) that settler mortality determined the nature of colonization does not rely on the finding in Acemoglu et al. (2002) that it was the least urbanized, least populous regions in 1500 that received the most European settlers. Hopkins (2009) is artificially conflating two literatures.

Hopkins (2009) charges that, in the institutional literature, different types of institutions and non-institutional influences cannot be disentangled by best practice, but instead need to be placed in the context of each case. In cross-country growth regressions, he suggests that comparisons between very different countries may discount differences that are not part of the investigation, and should be supplemented with country studies. He makes three broad critiques of this institutional view. First, the measures of population data used as a proxy for African well-being in 1500 are not very good. These come from McEvedy and Jones (1978), who effectively made guesses based on back projection. Second, the conceptual distinction in Acemoglu et al. (2001) between “productive” and “extractive” colonies has abandoned a historiography that at the very least distinguishes colonies of settlement, concession, and trade. Third, the literature compresses both time and history, with causes and effects that may be five hundred years apart, ignoring temporal and spatial differences.

\(^3\)Regressing log population density in 2002 on their estimates of log density in 1400 for the 47 non-island African nations gives a coefficient of 0.92 with a t-statistic of 11.79, and an $R^2$ of 0.76.
These critiques are misplaced. First, the reliability of the McEvedy and Jones (1978) population estimates is an unimportant distraction. If these are measured with unsystematic error, they would bias the results in Acemoglu et al. (2002) towards zero. With classical measurement error their results understate their own case. More fundamentally, Hopkins (2009) argues that research in the last thirty years has revised upwards the best estimates of the pre-colonial populations of Australia and the Americas. This is not a damning criticism, but rather the suggestion for an additional robustness check that might take an afternoon including data entry – re-running their regressions with the most up-to-date population estimates. Further, the Acemoglu et al. (2002) result is robust (Column (5) of Table (5)) to the inclusion of dummy variables for Africa, Asia and the Americas, and so this critique only matters if there is systematic measurement error within continents, and not solely across them.

The second criticism is more an invitation for future work than a critique. Englebert (2000a) has already looked to some extent at the heterogeneous effects of colonialism across different pre-colonial states. Similarly, if Hopkins (2009) believes that it is the distinction between colonies of settlement, concession, and trade that matters, these could easily be turned into regressors. Olsson (2004) has made progress in this regard, showing that the settler mortality instrument is a poor predictor of modern institutions in Latin America and Africa; identification in Acemoglu et al. (2001) comes from within Europe and Asia.

While Acemoglu et al. (2001) make a very simple distinction between “productive” and “extractive” colonies, their measures of interest – settler mortality and expropriation risk – are continuous variables that admit for more conceptual gradations than are possible in any discrete categorization. There is nothing about the fact that production was peasant-based in much of colonial Africa that implies colonies were not extractive; head taxes and export duties could easily funnel the surplus available to producers into salaries, pensions, metropolitan revenue, and other avenues that had little benefit to producers.

Third, it is not clear what to make of the charge that this literature compresses history. If \(X\) causes \(Y\), this is no less the case if \(X\) and \(Y\) are centuries apart. If his suggestion is that future research should uncover the specific mechanisms through which \(X\) caused \(Y\), this could be illuminating and policy-relevant, but unless these studies test measurable hypotheses, they will have no external validity and will teach us nothing that can inform policy.

Empirical studies have added to the lessons from Acemoglu et al. (2001) and Acemoglu et al. (2002) by demonstrating the importance in Africa of economic institutions that have deep historical roots, while others have examined the working of these institutions. A large fraction of this literature has focused on the consequences of weak states, though other
institutions such as local land tenure systems, slavery, polygyny have received econometric study.

11 of the 20 most “failed states” according to the Fund for Peace’s 2009 index are in Africa. This is in part a legacy of the process of colonization. Many of Africa’s borders were set arbitrarily during the Berlin Conference of 1884-5. Alesina et al. (2006) demonstrate that these sorts of “artificial” states – those with straight-line borders and particularly those for which major ethnic groups are split along multiple countries – have lower GDP per capita today. For robustness, they show that these relationships remain after several controls are included. Englebert et al. (2002), similarly, find that borders that divide African ethnic groups, regardless of their pre-colonial state centralization, are more likely to lead to international disputes. The same is true of straight-line borders, though inclusion of a dummy for North Africa pushes this latter result into marginal insignificance. Because of their sample size, Englebert et al. (2002) are not able to make strong causal claims, since they can only add a small number of additional controls.

Other studies interested in causation have attempted to show that pre-colonial states matter for current development. Because these are cross-sectional analyses, they are limited in their ability to rule out omitted variables. Bockstette et al. (2002) show that a very early start in state formation predicts growth today. Their cross-country results are robust to the inclusion of several controls, among them ethnic fractionalization. Their results similarly hold if the sample is restricted to non-OECD countries. Bardhan (2005) finds that their state antiquity variable also predicts present rule of law and political rights. The relative lack of state antiquity in Africa explains part of its poverty relative to the rest of the world.

Gennaioli and Rainer (2007) show that pre-colonial state centralization is positively correlated with modern GDP. This relationship survives inclusion of GDP per capita in 1960, exclusion of outliers and North Africa, and addition of a number of other controls. They posit that rulers of more centralized pre-colonial states were better able to extract public goods from colonial authorities, and show that their measure of centralization predicts the presence of paved roads, immunization rates, infant mortality, adult literacy, and school attainment. Bolt and Smits (2010) add local structures to this analysis, and show that countries whose pre-colonial societies had well developed community hierarchies and were outward looking are better governed in the present. This is robust to the inclusion of a number of other controls.

Englebert (2000b) demonstrates that a dummy for state legitimacy – an indicator for countries that were never colonized, regained former status on independence, were not created by colonialism, were not reduced to insignificance by colonialism, or for which the post-colonial state did not do “severe violence” to pre-existing political institutions – can make the “Africa dummy” disappear in cross-country growth regressions. This result survives a
handful of other controls. Englebert (2000a) further explores the importance of pre-colonial legitimacy. “Vertical” legitimacy (congruence with pre-colonial states, as defined above) is more consistently significant than “horizontal” legitimacy, measured by the percentage of the population in ethnic groups that are not present in other countries.

States are not the only institutions with pre-colonial roots that have received empirical attention, though many of the studies that stress their importance use modern, rather than historical data. Four examples are land tenure, slavery, polygyny, and the nature of kinship. First, there is a large literature on the importance of African land tenure systems, for which the institutions that existed before colonial rule are important inputs. Goldstein and Udry (2008), for example, have recently estimated that the fear of expropriation prevents individuals in Ghana from leaving land fallow if they have received it through their matrilineage. The resulting loss in land fertility costs Ghana close to 1% of its GDP per year. Second, indigenous African slavery, despite its importance in the “old” economic history of Africa (e.g. Kopytoff and Miers (1977), Miers and Klein (1998)) has not received adequate attention. A promising start has been made by Bezemer et al. (2009), who show a strong negative OLS correlation between the fraction of each country in Sub-Saharan Africa that practiced slavery prior to colonial rule and GDP today. This is robust to a number of geographic controls and to the identity of the colonizer. They speculate that slavery inhibited the development of states conducive to economic growth. They do not, however, have a plausible strategy for eliminating the possibility that there are omitted variables (such as export slavery) that determine both indigenous slavery and contemporary economic performance.

Third, while polygyny rates in Africa are much higher than in the rest of the world, this institution has been largely ignored by econometric studies. Tertilt (2005) has argued that the purchase of wives, whose daughters generate bride price, is a profitable investment for men and can crowd out investment in capital. She relies on a calibration exercise to show that her results are plausible. Polygyny may also matter for fertility (e.g. Garenne and de Walle (1989)), child mortality (e.g. Strassmann (1997)), and HIV transmission (e.g. Reniers and Tfaily (2009)), though no research design has yet overcome selection into polygyny. Finally, pre-colonial kinship structures such as matriliney matter in the present. La Ferrara (2007) argues that matriliney in Ghana should encourage sons to make transfers to their fathers in order to elicit gifts of land within their lifetime. She shows that sons give greater transfers to their fathers when a male nephew exists as a rival heir, that this effect is confined to (matrilineal) Akan households, that a similar effect cannot be produced by replacing nephews with other boys not entitled to inherit, and that these transfers increase if land has been given away in the past year. Similarly, La Ferrara (2003) provides suggestive evidence that kinship links in Ghana make it possible for lenders to use the borrower’s child to punish default, either by having the child withhold transfers or preventing other lenders
from making loans to the child. Gneezy et al. (2009) use experiments to show that men among Tanzania’s patrilineal Maasai chose competition at twice the rate of women, while this result is reversed for the matrilineal (and less patriarchal) Khasi of India.

Not all causal historical studies of African institutions need necessarily trace out their implications for present development. Cazzuffi and Moradi (2010), Mariotti (2009) and Fafchamps and Moradi (2009) each study a the functioning of a particular institution (cocoa cooperatives, Apartheid, and the Ghanaian army, respectively) in historical context. Cazzuffi and Moradi (2010) study Ghanaian Cocoa Producers’ Societies during the 1930s. They argue that the effectiveness of these societies was non-monotonic in size; beyond a certain point, coordination problems outweighed economies of scale. Cocoa sales per member first increase and then decrease in group size, conditional on many controls. This result also survives society fixed effects and society-specific age trends. While this is the sort-of context-specific study using micro-data that Hopkins (2009) calls for, Cazzuffi and Moradi (2010) are motivated by general (and, hence, falsifiable) theories of cooperatives and collective action.

Mariotti (2009) turns her attention to the functioning of labor markets during Apartheid. She argues that rising educational attainment during the 1960s and 1970s lessened whites’ resistance to opening semi-skilled jobs to Africans. She supports her argument using summary statistics to indicate that whites gained education and moved into more skilled occupations over time, a Mincer earnings regression to show the white skill premium did not fall from 1970 to 1980, and industry-level data to demonstrate that the share of African production workers was rising, but that this was not related to economic conditions captured by the capital labor ratio, output, or African wages.

Fafchamps and Moradi (2009) use military personnel records to investigate the quality of referred soldiers in the Gold Coast Regiment before the end of the First World War. Enlistees referred by their fellow soldiers had better observable characteristics in terms of height and chest circumference, while those sent by traditional chiefs were observably worse. Both referrals and those sent by chiefs were more likely to desert or be dismissed. These effects were concentrated among referrals brought in by sergeants or company sergeant majors, who could not be punished for bad referrals with denial of promotion; Fafchamps and Moradi (2009) take this cut of the data as evidence of referee opportunism.

5. Other contributions of causal history

In this section, I outline the contributions of “new” economic history and causal history to our understanding of the roles played by geography, the slave trades, colonial rule, and trade in African development.

5.1. Geography. Birchenall (2009) shows that urbanization, a common proxy for economic development, was lower in Africa in 1500 than in other pre-industrial societies. Comin et al.
similarly, show that the state of technology across countries in 1000 BC strongly predicts modern income and technology. These results suggest a powerful role for very long-term factors, including geography, in explaining the continent’s status today. “At the root of Africa’s poverty,” charge Bloom and Sachs (1998), “lies its extraordinarily disadvantageous geography.” They argue that Africa’s climate, fragile soils, and human and plant diseases have saddled it with low agricultural productivity, high disease burdens, low levels of trade concentrated in primary commodities, short life expectancies, and unfavorable youth dependency ratios.

The burden of malaria has received particular empirical attention. Gallup and Sachs (2001) use a cross-country growth regression to estimate that reducing malaria risk by 10% would raise growth by 0.3%. Not only does adult exposure to malaria adversely affect health and labor outcomes, but exposure as a child inhibits later health and human capital accumulation. To argue for causality, they replace the current presence of malaria with its severity in 1965. Sachs (2003) uses instrumental variables to further this point, using an index of ecologically-based malaria ecology to predict the disease’s prevalence.

These cross-sectional analyses can never completely rule out omitted heterogeneity. Weil (2010) takes an alternative approach, using various estimates of the present-day prevalence of sickle cell disease and childhood death rates from non-malarial causes to estimate that, in the most malarious regions of Africa, malaria lowered the probability of surviving into adulthood by ten percentage points. An alternative approach to uncovering causality is to exploit the differential prevalence of diseases before and after eradication campaigns. Bleakley (2010), employing an argument originally adapted to study the impact of hookworm eradication in the US south (Bleakley, 2007; Bleakley and Lange, 2009), uses the elimination of malaria in the United States, Brazil, Colombia and Mexico to show that later labor market outcomes for children exposed to eradication campaigns improved faster in originally more malarious districts. Lucas (2010) uses a similar procedure to argue that malaria eradication increased schooling and literacy in Paraguay and Sri Lanka. Using Indian data, Cutler et al. (2009) find no effect of malaria eradication on literacy or educational attainment, but do find that it increased consumption for men. Though none of these eradication studies use African data, they give support to the Gallup and Sachs (2001) view.

The causal effects of other aspects of African geography have been tested empirically. Absence of the plough is a major theme of African history. Alesina et al. (2011) use geographic suitability for plough-based crops as an instrument for its historic use, and show that the historical exclusion of women from agriculture in plough-based societies also excluded them from public life. This is reflected in female labor force participation and attitudes towards women today, both in cross-country data and across immigrant groups within the United States. Rainfall has been declining in Africa since at least the 1960s; Barrios et al. (2010)
use a regression of growth rates on rainfall, a rainfall-Africa interaction, and country and year fixed effects to estimate that, if this had not happened, the income gap between Africa and non-African developing countries would be between 9% and 23% smaller today. Many of Africa’s natural endowments are “point resources,” such as oil, gas, diamonds, and precious metals; the literature on the natural resource curse argues that these endowments promote Dutch disease, corruption, conflict and instability rather than development (e.g. Isham et al. (2005)). The discovery of oil in 1997 in São Tomé and Príncipe serves as a natural experiment with which these theories can be tested. Vicente (2010) shows that perceived corruption rose in that country after 1997 relative to Cape Verde, another former Portuguese colony.

Nunn and Puga (2007) argue that, while ruggedness inhibits modern economic performance by raising transportation costs and lowering agricultural productivity, rugged terrain in Africa helped its inhabitants resist the slave trades and their long-term negative institutional consequences. They show that, while ruggedness negatively predicts contemporary income, there is a significant positive interaction between ruggedness and an Africa dummy in a cross-country growth regression. This is robust to several controls, alternative measures of income and ruggedness, removal of influential observations, and inclusion of interactions between ruggedness and their controls. The interaction is strongest for regions hardest hit by the slave trade, and disappears once slave exports are added.

5.2. The slave trades. Empirical studies have begun to explore the effects of the Atlantic, Red Sea, Saharan, and Indian Ocean slave trades on Africa, building on a much older historical literature. These earlier studies focused on four major themes – underdevelopment, states, institutions, and demographics. Rodney (1974) is the most notable proponent of the view that Europe “underdeveloped” Africa. Others, including Darity Jr. (1992), Inikori (1992, 2007), and Eltis (1987) have built a literature that discusses the extent to which the slave trade diverted African economies from production to exchange, drained productive capacity, or opened the door for imports that undermined African producers. Writers such as Fage (1955), Davidson (1961), Thomas and Bean (1974), Rodney (1974), and Manning (1983) have debated the extent to which the slave trade led to the creation, growth, or destruction of African states, and the extent to which the slave trade precipitated war. Rodney (1966), Inikori (1982), Lovejoy (2000) and others have studied the slave trade’s effect on the institution of slavery within Africa. Others, such as Davidson (1961) and Evans and Richardson (1995) have highlighted potentially beneficial institutional legacies, such as long-distance trading networks, credit arrangements, and market organization. Manning (1990) uses a simulation model to estimate the demographic impact of the slave trades, though his assumptions and the results that follow from them have been challenged by Eltis (1987).
Some of this older work was econometric, and the topics that were addressed with regression analysis were directed by the availability of data on slave prices, export volumes, and the characteristics of individual voyages. LeVeen (1975) regressed quantity on price in order to argue that supply was upwards-sloping, though he did nothing to control for the simultaneity of supply and demand. Eltis (1990) uses a multivariate regression to show a secular decline in the heights of Yoruba slaves following the collapse of the Oyo Empire. Behrendt et al. (2001) collect a sample of 479 slave ships to show that ships from Upper Guinea, those that experienced more sailor deaths on the coast, and those carrying a greater share of males faced a higher risk of slave revolt. Eltis and Richardson (1995) use a Cobb-Douglas production function to compute total factor productivity estimates for 1,610 slave voyages, and then regress these estimates on origin, region of trade, and a time trend to show the particular efficiency of Cuban and English traders, Bristol, and West-Central Africa, and to demonstrate rising productivity over the first third of the 1700s.

The causes of mortality on slave ships generated its own “new” economic history. Eltis (1984) uses bivariate regressions to show that slave mortality increased with voyage length in some cases, but that there was little relationship between slaves per ton (i.e., crowding) and mortality. Eltis (1989) adds that the general rise in mortality through the nineteenth century was due largely to the predominance of West-Central Africa and Southern Brazil as exporting and receiving regions. Steckel and Jensen (1986) run separate regressions for deaths of slaves and crew, during loading and during the voyage. They find a significant effect of crowding on slave mortality during the voyage, as well as significant regional differences, but that the restrictions on slave crowding under Dolben’s Act had little impact on mortality. Haines et al. (2001) use a larger sample of voyages to find that voyage length generally increased mortality. The number of slaves embarked and absence of a stopover on the way to the New World both have marginally significant positive effects on mortality, while crowding and seasonal dummies are not significant.

A causal literature has emerged recently to test some of these claims more rigorously. The instigating work was done by Nunn (2008). He constructs country-level estimates of slave exports and demonstrates that these negatively predict GDP per capita in the present. To address causality, he uses an instrumental variables approach, in which the distance of a country from the nearest port of each of the four major slave trades is used to predict the number of slaves exported. This has given rise to several other papers, and the result is now part of the macroeconomic mainstream; in a recent review of African growth since 1995, Sala-i-Martin et al. (2010) are careful to note that growth rates have risen in both high-slave-export and low-slave-export countries.

Nunn and Wantchekon (2008) use ethnicity-level estimates of slave exports to argue that the slave trade also produced lower levels of trust today. Taking individual-level responses
from the Afrobarometer survey on the extent to which respondents trust other ethnic groups, their neighbors, their relatives, and local institutions. Nunn and Wantchekon (2008) show that – even conditional on individual and ethnic group characteristics and country fixed effects – members of ethnic groups that were greater victims of the slave trade in the past are less trusting today. They address causality by predicting slave exports using the ethnic group’s historic distance from the coast.

Whatley and Gillezeau (2009) demonstrate that the supply of slaves from Africa was upward-sloping, evidence that the trade increased “the production of social death in Africa.” They also find support for the “guns for slaves” view of Inikori (1977). Their approach is to estimate simultaneous supply and demand equations for the British slave trade via first-differences. The slave supply function includes prices, gunpowder imports, and a time trend, while demand is a function of price, British exports, the price and quantity of sugar produced in British colonies, and the presence of war in Europe. The exclusion of some variables from each equation allows them to identify the slope of both functions. In a related work, Whatley and Gillezeau (2010) argue that the slave trade contributed to ethnic fractionalization, showing that slave exports positively predict the number of ethnic groups within a defined radius of points along the African coast. This is robust to different radiiuses, different atlases of African ethnic groups, and geographic controls. Like Nunn (2008), they use distance to regional slaving markets to instrument for slave exports.

5.3. Colonial rule. Early writers on the impact of colonialism on Africa such as Rodney (1974) and Amin (1972) focused on the “underdevelopment” of the continent – its relegation to the role of exporter of primary commodities. Later Marxists, including Leys (1975) and Brett (1973) added class analysis to these theories. More recent studies have focused on the institutional consequences of colonial rule. Three themes from this newer literature are the undermining of existing institutions, destruction of systems of environmental management, and creation of political disorder. Vansina (1990), for example, suggests the only concession Europeans made to the “equatorial way of life was to preserve some cultural flotsam and jetsam, and to erect a structure labeled customary law, which was utterly foreign to the spirit of the former tradition.” Environmental historians such as Fairhead and Leach (1996), Kreike (1996), Kjekshus (1977), Neumann (2001) and von Hellermann and Usuanlele (2009) have shown that colonial rule in general and forestry and game reservation policies in particular disrupted African agricultural and ecological systems, leading to poverty and ecological degradation. Political historians such as Mamdani (1996) and Young (1997), contend that current disorder has colonial roots.

While Mamdani (1996) rejects any substantive distinction between British and French colonies, other writers such as MacLean (2002) and Firmin-Sellers (2000) argue that these

---

4I am indebted to Charlotte Walker for suggesting these three themes to me.
differences created institutions that produce divergent outcomes in the present. Empirical studies have confirmed that there are considerable differences in colonial legacies. Price (2003) finds in a non-industrial sample of countries that there is a significant and negative interaction between an “Africa” dummy and an “ever colonized” dummy, as well as a significant negative interaction between colonial history and malaria. Colonialism, he concludes, was worse in Africa, particularly in malarious regions. Bertocchi and Canova (2002) find that the identity of the colonizer is significant in a cross-country growth regression, while Agbor et al. (2009) similarly show that former British colonies have grown faster than French ones. Grier (1999) finds that colonies that were held longer have performed better today; Olsson (2009) obtains the same result using democracy as an outcome. Acemoglu et al. (2009b) note that earliness of independence and constraints on the executive at the end of the colonial period predict modern democracy. All of these studies, however, are cross-sectional analyses that do not control for the possible endogeneity of their colonial variables or for the correlation of these measures with unobserved determinants of their outcomes of interest.

More recent works have used careful research designs to credibly demonstrate the heterogeneous impacts of different colonial experiences. Moradi (2008) uses anthropometric evidence to show that heights grew at different rates and in different periods across ethnic groups within colonial Ghana and Kenya. Bubb (2009) exploits a regression discontinuity design to show that, while schooling outcomes vary drastically across the Ghana-Côte d’Ivoire border, differing policies towards land tenure have been unable to influence institutions on the ground. Instead, he adopts the argument from Besley (1995) that tree crops create “Lockean” rights over land, and shows that an ecological cocoa-suitability index is a strong predictor of households’ rights to sell and rent out land. Berger (2008) shows that divergent forms of colonial rule continue to reach into the present even within countries. Between 1900 and 1914, the border between Northern and Southern Nigeria was a straight line. Without ports that could levy tariffs, direct taxes were used to create local bureaucratic capacity in the North. Using a regression discontinuity design, he shows that present-day communities just north of the now anachronistic border continue to show greater satisfaction with their local governments and receive higher levels of vaccination.

There is a substantial cross-country literature (e.g. Porta et al. (1997)) stressing that a legacy of common (as opposed to civil) law promotes property rights, quality of government, political freedom, and financial development today. The basic argument is that common law systems evolved largely to preserve individual rights, while civil law was historically used as a tool for expanding the Roman empire. The influence of lawyers under common law is a counterweight to state power. Common law courts can serve as a check on the executive, and are less dependent on an efficient bureaucracy. The colonial reliance on “customary”
law has received significant attention from historians – Berry (1992), Chanock (1985), Moore (1986), and Spear (2003) are notable examples.

Empirical studies have tested these theories of colonial legal heritage in Africa. Joireman (2002) finds that common-law African countries have performed better on the ICRG measures of rule of law, corruption, and civil liberties since 1990. Her analysis is not multivariate, however, and so she does control for other factors correlated with legal origin. Lange (2004) has shown that the proportion of “customary” court cases in all court cases in 1955 is a negative predictor of several measures of good governance today in a sample of 33 former British colonies. His small sample size limits what can be done with the data, but he does show that the result is robust to removing South Asia, and to inclusion of a handful of key controls. Marchand (2010) performs an exploratory regression to show that contemporary deforestation rates are correlated with expropriation risk and corruption, and that these correlations are moderated by legal origin and colonial history.

Educational policies also differed between French and British colonies. While British rulers provided primary vocational training and literacy, their fear of educated elites dissuaded them from encouraging education above the primary level. In French colonies, while there was also a suspicion of educated Africans, curricula were centralized and schools were mandated to teach in French (Bolt and Bezemer, 2009). Grier (1999), Cogneau (2003), Bolt and Bezemer (2009), and Huillery (2009) have all tested whether differences in colonial educational education have caused divergent outcomes in the present.

Grier (1999) regresses the percentage of the population attending school at independence on the year of independence and a British colonial dummy, finding a positive and significant coefficient on the latter variable. Because colonial powers may have divided Africa according to observable and unobservable characteristics that may also affect the provision of education, Cogneau (2003) uses a non-parametric matching estimator to test for the causal effect of the identity of the colonial power on post-colonial education outcomes in an African sample. He includes a set of arbitrary geographic variables in the selection model and a handful of additional controls. He finds that former British colonies outperformed French ones on years of schooling and enrollment rates in 1960 and 1990, though his results are undermined by the lack of a theoretical justification for the choice of variables in the selection equation.

Bolt and Bezemer (2009) are concerned with whether colonial education better explains African development than colonial institutions. In a sample of 24 former British and French colonies, they show that settler mortality does not predict modern institutional quality within Africa, but does predict school pupils as a fraction of the population in the late colonial period. Using this as an instrumental variables strategy, they find that colonial education predicts GDP per capita today. Huillery (2009) also shows that colonial investments in education matter today, using matching estimates to show that the number of teachers
per capita between 1910 and 1928 predicts differences present-day school attendance across districts in former French West Africa. She uses the same technique to show that colonial infrastructure and health investments similarly have had long-term effects.

The long-term consequences of colonial-era missionary efforts have also received empirical attention. Gallego and Woodberry (2010) take cross-section of African provinces and find that, while the presence of either Protestant or Catholic missions in the past predicts literacy and schooling today, the competition Protestant missions faced due to restrictive policies in Catholic colonies led them to improve the quality of their schools in those countries. Nunn (2010) shows that historical exposure to missions at both the ethnic group and village level is a positive predictor of whether a respondent in the Afrobarometer is Christian today, conditional on individual, ethnic, and village-level controls. Nunn (2009a) extends this analysis, finding that historical exposure to Protestant missions, as opposed to Catholic ones, positively predicts education today, particularly for women. Historical exposure to Catholic missions reduces respondents’ willingness to agree with the statement that women should have equal rights or participate in politics, as well as their support for democracy. Because he is concerned about endogenous placement of colonial missions, Nunn (2010) also controls for variables that helped determine mission location – temperate climate, high altitude, and prior contact by European explorers.

Other econometric studies have looked at the impact of European settlement directly. Angeles and Neanidis (2010) find a non-monotonic relationship between European settlement and corruption in a sample of ex-colonies, although they do not control for the possible endogeneity of European settlement. They hypothesize that greater European settlement initially increases the size of the rent-seeking elite, but that once Europeans become a majority of the population the settler community serves as a check on the elite. Huillery (2010) shows that the European share of the population in 1925 is a significant positive predictor of contemporary differences in literacy, schooling, child health, access to basic services, and housing quality across districts of former French West Africa. To account for endogeneity, she uses the number of events expressing hostility towards the colonial power as an instrument for the number of early settlers.

Other studies, while not uncovering causal mechanisms, have employed econometric techniques of the “new” economic history to uncover secular changes during the colonial period. Buelens and Marysse (2009) use stock returns to show that mining was highly profitable in the Belgian Congo to 1955, when country risk became a reality. Moradi (2009) demonstrates, using a sample of Kenyan army recruits and civilians from 1880 to 1980, that there was a large increase in heights from 1920 on. “However bad colonial policies were,” he concludes, “nutrition and health steadily improved from the beginning of the 20th century.” Austin et al. (2009), similarly, show that heights increased over time in colonial Ghana, with this
trend reversing only in the 1970s. Gould et al. (2008) find that the height increase from 1925 to 1960 in Ghana and Côte d’Ivoire was comparable to that in France and Britain over the years 1875 to 1975, a trend that reversed in the first twenty five years of independence.

5.4. Trade. “Underdevelopment”-type explanations of African poverty (Frank, 1969; Nkrumah, 1966; Wallerstein, 1974) have already been discussed above in connection with colonial rule and with the slave trades. The continued dependence of independent African countries on primary commodity exports has received widespread notice as a drag on growth. Economic growth and commodity price movements are strongly correlated in Africa, and these prices have either been trendless or downwards-trending over the long term (Deaton, 1999). The finding that natural resource exports impede growth (e.g. Sachs and Warner (1997)) does lend some support to the structuralist view. The consensus from the empirical literature on trade and growth, however, is that trade has a strong positive effect on GDP per capita. The correlation between openness to trade and GDP growth is strong; Sachs and Warner (1995) argue that it is the difference between convergence and failure to converge for poor countries. Because of possible reverse causation, the strongest evidence for a positive link between trade and growth comes from studies that have used instrumental variables. Frankel and Romer (1999) use geographic characteristics to predict trade as a share of GDP, and find that OLS estimates of the effect are not overstated.

Interest in underdevelopment was revised during the late 1990s and early 2000s by the rise of “globalization.” Two broad critiques have been raised against this latter wave of trade liberalization, with empirical support. First, openness increases exposure to external risks. Jones and Olken (2010), in an empirical specification that controls for both product-country and product-year fixed effects, note that climate shocks affect the agricultural and light manufacturing exports of poor countries, while heavy manufacturing and the exports of rich countries are not affected. Rodrik (1998) suggests that the positive correlation he finds in a cross-country growth regression between trade openness and government size exists because government spending can reduce terms-of-trade risk. He demonstrates that there is a positive interaction between openness and terms of trade volatility with the size of government as a dependent variable. Second, the impact of trade on governance may be negative. Statistically, democracy appears to promote trade openness, but not the reverse (Milner and Mukherjee, 2009). In fact, Rigobon and Rodrik (2005) use identification through heteroskedasticity to find a negative effect of trade openness on democracy.

While Africanists have participated in these contemporary debates (e.g. Bevan and Fosu (2003)), empirical history studies such as Fourie and von Fintel (2009), Boshoff and Fourie (2010) and Jedwab (2010) have looked at Africa’s past trade in historical context. Fourie and von Fintel (2009) use VOC tax records and principal components analysis to create wealth indices that show rising inequality in the Cape Colony through 1730. They argue
that, while the region was suitable for slave-based production that concentrated wealth and power in the hands of an elite, the climate could not initially support large plantations. The mercantilist policies of the VOC were needed for institutions to arise that supported inequality. Boshoff and Fourie (2010) test for co-movement in ship traffic, wheat exports, wine production, and stock farming in the early Cape Colony. They show that wheat exports spurred local demand in other sectors of the economy, supporting broader growth. Jedwab (2010) finds that cocoa exports drove urbanization in the Ivory Coast and Ghana. He collects statistics on urban populations and district-level cocoa production in both countries from 1921 (Ghana) and 1948 (Ivory Coast), and regresses urban population on cocoa production and year and district fixed effects. To control for endogeneity, he instruments for cocoa production using an interaction of cocoa suitability with distance to the cocoa frontier.

6. Conclusion

Hopkins (2009) concludes with three notes of caution. First, he worries that the “new” economic history is too “driven by the need to demonstrate a mastery of econometric techniques rather than by a concern to incorporate differences in the historical context that are essential if the problem under review is to be explained successfully.” This critique has been repeated in economics; Angrist and Pischke (2010) quote Scheiber’s (2007) piece for The New Republic, in which prominent economists argue that causal methods have diverted attention away from topics such as poverty and unemployment and towards behavior on game shows, health club membership renewals, or purchases of winter clothing. Ironically, while Hopkins (2009) worries that technique-driven research design leads economists to ask questions that are too broad and lack context, economists have charged the design-based literature with focusing on topics that are too narrow. Angrist and Pischke’s (2010) responses are equally applicable here: careful research designs on minor topics where the data is particularly good can validate the conclusions of broader descriptive studies. Evidence accumulates from multiple studies across several contexts, until a consensus emerges.

Second, he is concerned that this literature will lead to the resurgence of the view that societies outside the West “had little to contribute to economic development.” The “Rest” was not an undifferentiated mass before colonial rule, and he argues that it is “anachronistic to try to reduce the history of modern economic development to a question of determining the location of Western settlement.” Here, Hopkins (2009) is correct, but this is an invitation for future work and not a critique. The $R^2$ in Acemoglu et al.’s (2001) regression of contemporary institutions on settler mortality in only 0.31 (Column (9) of Table (3)). This leaves a substantial fraction of the institutional legacies of colonialism to be explained by later work. Progress in incorporating pre-colonial structures into explaining different colonial legacies within Africa has already been made by Huillery (2010) and Englebert (2000b).
Third, he warns historians to be skeptical of universal prescriptions derived from first principles, noting that the United States has thrived despite ethnic fragmentation, regionalism, political corruption, and difficulty in agreeing on public goods. Here, Hopkins (2009) validates the aims of causal history. The importance of each of these variables for development has a theoretical basis, but it is up to empirical work to test these theories. A robust, causal statistical relationship cannot be dismissed by referring to examples any more than it can be proved by them.

Hopkins (2009) argues that, in spite of his concerns with the “new” economic history, “there is still a powerful case for treating the new literature seriously.” His grounds for this, however, are misplaced. He laments that it has been “more than twenty years since historians themselves produced big arguments attempting to understand Africa’s long-run economic development and continuing poverty,” and that there is now a “prospect that history will be taken seriously by policy-makers.” He is counseling historians to engage the “big ideas” of economists’ recent work on Africa’s past, and not with their methods. A careful series of case studies that engage the big ideas of reversing fortunes, institutions, and ethnicity will not be taken seriously by policymakers unless they can credibly claim to uncover causal relationships that can be extrapolated beyond their specific contexts in order to inform current policy. This is exactly what causal studies of African history strive to achieve.

References


University of the West Indies.


Scheiber, N. (2007). Freaks and geeks. how Freakonomics is ruining the dismal science.


