The imprint of the European colonizers on contemporary African development and culture is multifaceted and far-reaching. This volume provides an overview of how economists have attempted to shed light on these legacies. The range of the topics covered is breathtaking, starting from the slave trades, continuing with the drawing of the colonial borders in European capitals, the extent of indirect rule, the use of bloody concessions in DRC, the investments in transportation infrastructure, the role of Christian missions, and ending with the type of the independence movements in Africa. The geographic focus of the volume then turns to exploring specific aspects of the colonizing experience in Asia and the Pacific looking at the role of direct versus indirect British rule in India, the economic trajectory of China’s treaty ports and the impact on current-day gender attitudes of the male-biased sex ratios of early colonizers in Australia.

To get to these fascinating questions a wealth of diverse sources including historical archives, anthropological maps, and satellite images are fruitfully combined and subjected to state of the art econometric techniques and theoretical models. The findings of this ambitious research agenda are novel highlighting the shadow of history on the various aspects of the economy and the polity. While there are many open issues and debates and although development is not deterministic, one message is clear: We are shaped by history (Martin Luther King Jr).
The long economic and political shadow of history -
Volume 2. Africa and Asia
The long economic and political shadow of history -

Volume 2: Africa and Asia

Edited by
Stelios Michalopoulos and Elias Papaioannou

A VoxEU.org eBook
February 2017
Centre for Economic Policy Research (CEPR)

The Centre for Economic Policy Research (CEPR) is a network of over 1,000 research economists based mostly in European universities. The Centre’s goal is twofold: to promote world-class research, and to get the policy-relevant results into the hands of key decision-makers.

CEPR’s guiding principle is ‘Research excellence with policy relevance’.

A registered charity since it was founded in 1983, CEPR is independent of all public and private interest groups. It takes no institutional stand on economic policy matters and its core funding comes from its Institutional Members and sales of publications. Because it draws on such a large network of researchers, its output reflects a broad spectrum of individual viewpoints as well as perspectives drawn from civil society.

CEPR research may include views on policy, but the Trustees of the Centre do not give prior review to its publications. The opinions expressed in this report are those of the authors and not those of CEPR.

Chair of the Board                Sir Charlie Bean
Founder and Honorary President    Richard Portes
President                        Richard Baldwin
Research Director                 Kevin Hjortshøj O’Rourke
Policy Director                   Charles Wyplosz
Chief Executive Officer           Tessa Ogden
# Contents

<table>
<thead>
<tr>
<th>i</th>
<th>Foreword</th>
<th>vi</th>
</tr>
</thead>
</table>
| 1  | Series Introduction: Historical legacies and contemporary development  
*Stelios Michalopoulos and Elias Papaioannou* | vii |
| 2  | Understanding the long-run effects of Africa’s slave trades  
*Nathan Nunn* | 32 |
| 3  | The contemporary shadow of the Scramble for Africa  
*Stelios Michalopoulos and Elias Papaioannou* | 37 |
| 4  | Indirect rule and local government reform in Africa today  
*Tristan Reed* | 50 |
| 5  | King Leopold’s ghost: The legacy of labour coercion in the DRC  
*Sara Lowes and Eduardo Montero* | 67 |
| 6  | How colonial railroads defined Africa’s economic geography  
*Remi Jedwab, Edward Kerby and Alexander Moradi* | 76 |
| 7  | The devil is in the detail: Christian missions’ heterogeneous effects on development in sub-Saharan Africa  
*Julia Cagé and Valeria Rueda* | 88 |
| 8  | Independence movements and democracy in Africa  
*Leonard Wantchekon and Omar Garcia-Ponce* | 99 |
| 9  | Direct versus indirect colonial rule in India  
*Lakshmi Iyer* | 111 |
| 10 | The legacies of forced freedom: China’s treaty ports  
*Ruixue Jia* | 121 |
| 11 | On the short- and long-run effects of male-biased sex ratios: Evidence from convict transportation to Australia  
*Pauline Grosjean* | 128 |
Foreword

Centuries of European colonisation have left widespread imprints on culture and society across the globe. In this eBook, an overview is provided of how economists have attempted to shed light on the legacies of colonial rule on the political and economic landscapes of the African continent, India and China.

This second volume of three covers a vast range of topics about Africa, including the slave trade, the artificial drawing of borders in European capitals, the extent of indirect rule, the investments in transportation infrastructure, and the roles of Christian missions, and independence movements. Issues related to the Asian continent are also included, such as the role of direct versus indirect British rule in India; the economic trajectory of China’s treaty ports, and the impact on modern gender attitudes of the male-biased sex ratios of early colonisers in Australia.

The fascinating findings of this ambitious series come from a wealth of diverse sources. Historical archives, anthropological maps and satellite images are combined with state of the art econometric techniques and theoretical models. They highlight the long-run impact of history on the economy and politics, and the need to understand it.

CEPR is grateful to Professors Elias Papaioannou and Stelios Michalopoulos for their joint editorship of this eBook. Our thanks also go to Sophie Roughton and Simran Bola for their excellent and swift handling of its production. CEPR, which takes no institutional positions on economic policy matters, is delighted to provide a platform for an exchange of views on this crucially important topic.

Tessa Ogden
Chief Executive Officer, CEPR
February 2017
The emergence of New Economic History

Those 18th and 19th century philosophers who shaped economic thought (David Ricardo, Adam Smith, Karl Marx, John Stuart Mill), in company with early 20th century scholars such as John Maynard Keynes and Joseph Schumpeter, believed in the blending of core economic ideas (value maximisation, incentives, market laws) with history. Political economy – as economics was referred to at the time – was a discipline that combined elements from a wide array of social sciences and interpreted them with its new tools. Yet, efforts to draw insights from history and related disciplines with which to shed light on economic questions lost steam during the latter part of the last century.

The neoclassical approach would leave little room for a deeper understanding of important historical events in the growth process, such as colonisation, institutional changes and cultural traits that most people would instinctively see as fundamental drivers of comparative development. Not modelling explicitly the role of history, geography-ecology, and culture (religion, beliefs, family ties, norms), led to these features being viewed empirically as a ‘residual’ of the growth process driven by capital deepening, either in the physical sense of tools and machines or in the human sense of education and health.
And, while new growth theories show how small differences in initial endowments may translate into large differences, when it comes to urbanisation, agglomeration, and well-being, they were agnostic on the origins of these initial conditions, naturally rooted in the historical record. But economists’ disregard for history was not confined to the growth literature or macroeconomics. The total number of published economic history papers in ‘top’ general-interest journals declined (McCloskey 1976, Abramitzky 2015), as economists moved from relatively simple, intuitive theories and applied methods to more complex, often esoteric, mathematical models and elaborate quantitative-econometric empirical approaches (see also Temin 2013). However, since the late 1990s there has been a revival; a ‘new economic history’ literature has emerged that studies important historical episodes, with the goal of tracing their consequences on contemporary outcomes. This new literature applies economic models and econometric techniques to examine the shadow that history casts over various aspects of the economy and the polity.

As is typical with history, the process of being ‘rediscovered’ among economists – learning from the past and uncovering its legacy – has been gradual; and clearly there were important works of economic history between the 1950s and 1990s, such as Douglass North’s *Structure and Change in Economic History* in 1981 and Avner Greif’s work on the interplay of culture and institutions (Greif 1993, 1994). Yet, the tipping point seems to be a trio of influential works that appeared in the late 1990s. Acemoglu *et al.* (2001, 2002, 2005) put forward the ‘colonial origins of comparative development’ thesis, where the type of colonial strategy and early colonial institutions influenced subsequent economic and political development. The ‘law and finance’ works of La Porta, Lopez de-Silanes, Shleifer, and Vishny (1997, 1998, 2006) showed that legal

---

1. The origin of this paradigm shift seems to be the publication of Paul Samuelson’s revolutionary textbook, *The Foundations of Economic Analysis* in 1947. While the use of mathematics goes back to Alfred Marshall, John Hicks, Edgeworth, and others, Samuelson is rightly credited for trying to synthesise and mathematically formalise economic reasoning (see Temin 2013).

2. The global financial meltdown of 2007-2009 and the associated deep recession renewed the interest of macro-economists working on business cycles, monetary and fiscal policy in economic history. Somewhat paradoxically, as “dynamic macro” didn’t have much of an interest in history, two of the main protagonists of policymaking in the aftermath of the crisis, Ben Bernanke and Christina Romer, were among the best scholars of the Great Depression and other historical crisis periods.
origins transplanted during colonisation have a significant bearing on contemporary contractual institutions and finance. Engerman and Sokoloff (1997, 2002) argued that colonial-era inequality, shaped by the type of colonial strategy employed (mining, plantations), explains the divergent development paths of Southern and Northern America. The analysis of historical events was also fuelled by important theoretical contributions; the development of ‘unified growth theory’ (Galor and Weil 1999, 2000, Galor 2011) which brought to the forefront the study of pre-industrial (Malthusian) and modern economies in a unified framework; and the game-theoretic modelling of both institutions and ‘class struggles’ (Acemoglu and Robinson 2000, 2001). For a textbook treatment see Acemoglu and Robinson (2006) and Besley and Persson’s (2011) that study jointly the co-evolution of economic and political development, through the lens of ethnic fragmentation, inequality and conflict, which, in turn, may be shaped by geographic, cultural or historical forces.

At the same time, the profession was greatly influenced by Jared Diamond’s magnum opus, Guns, Germs, and Steel. Blending insights from geography, biology (zoology), cultural anthropology, sociology, and even archaeology, Diamond provided a powerful thesis on economic and political development over the long run. While Diamond stressed the importance of geography-ecology in shaping institutional, political, and societal traits, which in turn impact development, in parallel work David S. Landes (1998) stressed the role of family ties, religion, beliefs and norms, arguing forcefully that “culture makes all the difference”. Besides their original contributions, these works also brought to the forefront old ideas that had been side-lined – for example, Max Weber’s thesis on the importance of Protestantism in the Industrial and Commercial Revolution, or Friedrich Hayek’s ideas linking legal-origin traditions to economic freedom and prosperity.

The new economic history follows an inter-disciplinary approach, integrating historical narratives and insights from other-than-economics social sciences (mostly sociology, political science and anthropology) with mathematical models, and formally tests long-standing, influential conjectures with econometric techniques. This eBook series summarises contributions aimed at uncovering The long economic and political shadow of history.  

3 Nunn (2009, 2014) provides a thorough overview of this body of research. See also Spolaore and Wacziarg (2014).
Some of the themes of these works transcend continents (like colonisation), some are regional (e.g., the spread of Protestantism in Europe or the slave trades in Africa), while others are country(ies)-specific (e.g., the French Revolution’s impact in France and Germany, the impact of the Holocaust in Russia, the Nazi occupation of Italy, and the impact of communism). The works covered in this eBook series differ in their focus. Some stress the role of cultural traits, such as religion (e.g., Becker and Woessman 2009) and beliefs (Voigtländer and Voth 2012, Grosfeld and Zhuravskaya 2013); others emphasise the importance of ‘early’ institutional structures, related to colonial practices (Naritomi, Soares, and Assunção 2012), colonial ‘indirect-rule’ (Acemoglu, Reed, and Robinson 2014, Iyer 2011), ‘forced labour’ systems (Dell 2010, Lowes and Montero 2016) and the forced placement of indigenous tribes in special reserves (Dippel 2013); others stress the impact of the environment, geography, and ecology (e.g., Hornbeck and Naidu 2015, Ashraf and Galor 2003); and Fenske and Kala provide, in Volume I, a synopsis of ‘Environmental Economic History’. Some studies analyse the long-lasting effects of colonial investments in roads and railroads (Jedwab, Kerby, and Moradi 2016) and ports (Jia 2014), or the effect of Christian missions on education, health, and beliefs (Cage and Rueda 2016, Valencia 2016) whereas others explore the legacies of relatively recent historical events, examining the impact of the Holocaust (Acemoglu et al. 2014), communism (Alesina and Fuchs-Schundlen 2007), and Nazi occupation (Fontana et al. 2016) not only on economic development, but also on inequality, beliefs and political participation. Some works analyse the long-lasting effect of specific migration movements – for example Moser et al. (2014) examine the impact of Jewish émigré scientists leaving Nazi Germany on US innovation, while Grosjean and Khattar (2015) and Grosjean and Brooks (2016) study the impact of convict transportation from England to Australia, during the 18th and 19th century, on contemporary culture.

While there have been many influential contributions, there are many open issues. As the literature has tackled important and highly controversial subjects, there is naturally a debate and therefore more research is needed to develop a clearer view of these topics. For example, while there is not much ambiguity that the type of colonisation, or the identity of the colonial power, was consequential for contemporary development, there is still debate on the exact mechanisms at work. This is a hard task, as such large-scale episodes affect not only the economy, but also beliefs and norms, institutions, and the distribution of economic, and political power.
Another open issue regards the potential interactions of historical events with critical junctures in the process of development; for example, the negative consequences of forced labour systems in Latin America and Africa may be large and long-lasting, since they prevented industrialisation at a time of massive technological innovation. In line with this, Pascali (2017) finds that the trade and growth benefits of the introduction of the steam engine, during the first era of globalisation, were especially large in countries with strong executive constraints and checks-and-balances. Likewise, the high levels of civicness in Medieval Italian city-states, which correlate strongly with beliefs and development (Guiso et al. 2016), may have been especially important during the Enlightenment and the subsequent Commercial and Industrial Revolution.

While the papers summarised in this eBook series all reveal considerable persistence, they do not imply that development is entirely determinist; leadership and policies promoting investment and human capital accumulation or policies and technologies enabling trade integration do matter. So, while on the surface, there may seem to be a tension between studies uncovering the shadow of history and the ability of policy to affect development (Banerjee and Duflo 2014), the works below can be viewed as showing that important policy decisions occurring in ‘critical’ times do have long-lasting effects (that subsequent policy-makers find hard to reverse). For example, as the chapter of La Porta, Lopez-de-Silanes, Shleifer and Vishny puts it, while laws often change, the associated legal environment and culture is much harder to modify, since they depend on judges and lawyers’ training, legal tradition, precedent, and social norms. The works summarised in this eBook series, as well as other important studies in the same genre, carry an obvious, yet often neglected, lesson. The success of any policy implemented crucially depends on the underlying institutional and cultural heritage of a given society, therefore it is vital for policymakers, development actors and investors to understand the historical context, so as to model incentives and design effective interventions.

Despite the wide-ranging topics that contributions to this eBook series cover, there are many, often not fully-appreciated, similarities across them.

4 See for example Jones and Olken’s (2005, 2010) innovative studies on the role of leadership on growth.
First, an element that quickly stands out across all chapters is their inter-disciplinary approach. Many entries, for example, carefully investigate the economic and cultural consequences of historical events and test influential conjectures put forward by political scientists and sociologists. For example, Nunn (2008) quantifies the legacy of Africa’s slave trades, testing long-established theories in political science. Going over the papers, summarised in the chapters of this eBook series, the reader learns a lot, not only about the economics, but also about the context, the protagonists, and the key issues at stake. This inter-disciplinary approach has been a major step forward and must be continued (Lamoreaux 2015). Related to this, most of the recent contributions come from scholars who cannot be identified as ‘economic historians’, since their contributions span several fields (e.g., Daron Acemoglu, Andrei Shleifer, Luigi Zingales, Guido Tabellini, Ross Levine, and Oded Galor, to name but a few).

Second, most papers use (typically, self-collected) high-quality historical data that are combined with contemporary proxies of economic, institutional, and financial development. Some of the data collection and compilation efforts are remarkable, as economists have cleaned and digitised hard-to-access historical archives: for example, Nunn (2008) has digitised records on the hundreds of shipments of slaves out of Africa; building on earlier work of Philip Curtin, Acemoglu et al. (2001) compiled data reflecting settler mortality rates during the colonial time, using newspaper articles from the 19th century; Voth and Voitalander (2012) digitised data on Jewish pogroms in German lands during the Black Death; Cage and Rueda (2016) and Valencia (2016) geo-located the Catholic and Protestant missions in Africa and Latin America; and Acemoglu et al. (2014) constructed a dataset tracing all paramount chiefs installed by the British during their indirect rule of Sierra Leone. Using a plethora of original sources and meticulous efforts, Easterly and Levine (2016) have unearthed data on the share of Europeans, during the colonial times, for more than 100 contemporary countries, in order to assess their initial presence and long-lasting effects. And many papers are based on digitised historical and anthropological maps, portraying, for example, the spatial distribution of African ethnicities at the time of colonisation (Nunn 2008, Nunn and Wantchekon 2011, Michalopoulos and Papaioannou 2016), or of Native American tribes (Dippel 2013).

---

5 This is an important step forward, because, as Mancur Olson famously argued it is hard conducting interdisciplinary research without losing discipline.
Many use data from anthropologists (e.g., George Peter Murdock’s mapping of ethnicities and codification of key economic, cultural, and institutional traits), as well as geo-referenced high-resolution data on geography (e.g., soil quality, terrain ruggedness, suitability of land for particular crops, among others). Some works employ data produced by population geneticists and evolutionary biologists (Spolaore and Wacziarg 2014 and Ashraf and Galor 2016). In addition, some of the papers summarised in this eBook series have compiled novel data for the most recent period. La Porta et al. (1997, 1998) have gone over the corporate and bankruptcy laws of many countries to construct quantitative measures capturing the strength of minority shareholders’ and creditors’ rights. Going over the historical record regarding the independence period for each African country, Wantchekon and Ponce (2016) have classified independence movements into rural and urban based.

Third, the works in this eBook series share an important feature: they are ambitious. Employing state-of-the-art econometric techniques and intuitive theoretical constructs, they aim to understand the consequences of events, such as the role of colonial strategies in Brazil (Naritomi et al. 2012), the impact of the Holocaust (Acemoglu et al. 2011), the artificial drawing of those colonial borders in Africa that survived independence (Michalopoulos and Papaioannou 2016), the role of Jewish immigrant scientists in US innovation (Moser et al. 2014), the role of the steam ship in advancing commerce and institutions (Pascali 2017), the long-lasting effect of communism on attitudes and values (Alesina and Fuchs-Schundlen 2007), or the contemporary legacies of city-states during Medieval times (Guiso et al. 2016). For each of these topics there are thousands of pages of excellent work in the areas of history, political science and sociology, along with a plethora of qualitative studies. To this corpus of excellence, economists bring along formal theorising and sound econometric tools, complementing and building upon the existing research. Naturally, ambition is at the same time a strength and a shortcoming. This is because, almost by default, it is impossible to perfectly explain and trace the legacy of history. Yet economists have been open to critique and in many instances they have gone back to the original archival data, re-examined and re-codified the entries, and improved upon the findings of these influential works.
Clearly, the massive amount of work in history and political science, that rightly studies detailed aspects of these events, should and will not be ‘replaced’ by some simplified economic model that (by necessity) abstracts from many relevant aspects to focus on some key channels; and the vast body of historical evidence should not be swept under the carpet because of some interesting econometric technique. But economics does offer a fresh new angle. Continuing and strengthening the ongoing cross-pollination between economists and scholars in other disciplines can only prove beneficial to the cause; that of understanding whether, how and why history matters in terms of shaping the observed variation in outcomes today across regions.

Fourth, as the revival of the ‘new economic history’ coincided with the ‘credibility revolution in empirical economics’ (Angrist and Pischke 2010), many of the papers rely on modern state-of-the art econometric methods, such as difference-in-difference estimation, ‘matching’ methods, regression discontinuity, instrumental variables, and ‘natural experiments’. These methods – while not perfect – attempt to account for unobserved or hard-to-measure confounding factors, or to exploit some form of ‘quasi-random’ variation in the historical record. For example, in her study on the impact of British direct rule of Indian states (as compared to indirect rule via local kings), Iyer (2010) exploits experimental variation stemming from the lapse rule, applied by the British between 1848-1856; under this system the British would take under direct rule only those Indian states where the dying native ruler did not have an heir. To identify the impact of colonisation, Feyrer and Sacerdote (2009) focus on small islands and use variation in wind patterns to obtain quasi-random variation on the timing and duration of European contact. History is full of idiosyncratic events, accidents and mistakes that economists (and more recently political scientists) exploit to credibly identify relationships. For example, scholars assess the effects of (unexpected) environmental shocks, as Hornbeck and Naidu (2014) examine the impact of the Great Mississippi Flood of 1927 on black out-migration and subsequent agricultural development in the US South. Likewise, Fenske and Kala (2015) link enslavement raids in Africa to temperature shocks. Even though historical variation entails often random elements, history is shaped by

6 There is a related large literature on environmental shocks and civil conflict (see Burke, Hsiang, and Miguel 2015, for a review).
people, leaders, technology, and the environment; so the impact of historical events
cannot be evaluated via formal experimental tools (such as lab experiments or
randomised-control trials), which are quite useful for the evaluation of various narrow
policy interventions. Yet this does not invalidate their strength; quite the contrary.
The studies reviewed in this eBook series exploit some unique aspect of the historical
landscape or try to account as carefully as possible for confounding factors to examine
the legacy of events that affect people’s views, beliefs, country institutions, and the
economy.

A fifth shared principle of these works is the conscious effort of the authors to make
their findings accessible not only to specialists but also to the general public. The papers
are widely accessible and fun-to-read (even when you disagree). Some authors, most
notably Acemoglu and Robinson (2012) and Mokyr (2016) have written important
books for the general public that build on this recent research agenda. And almost all
authors have written non-technical summaries and Op-Eds. We hope that this eBook
series will contribute to that ongoing effort.

Below we illustrate broadly the patterns of persistence over long periods of time. We
then briefly go over the original works summarised in the 33 chapters of the three
volumes of this eBook series.

## 2 Long-Term Persistence

### 2.1 The Evidence

Economic development appears quite persistent.\(^7\) Figure 1a provides a visualisation of the
correlation of the logarithm of GDP per capita in 2014 and the respective statistic in 1960.
The regression yields a close-to-unity elasticity (0.925), implying that countries that were 10%
percent richer than other nations in the 1960s are (on average), today, 9.25% more developed.

\(^7\) Clearly there is noise in the GDP numbers, especially when one goes back to the pre-1960 period. If measurement error
takes the classical form (i.e., is unrelated to country’s output), then the correlations will be attenuated, indicating lower
(than the true) persistence.
And one can explain half of the variation in contemporary performance with that of 1960 – the first year of (relatively) high-quality data from the Penn World Table (the elasticity and in-sample fit are similar if one uses Angus Maddison’s data). Clearly during the past 6 decades, there have been success stories of rapid structural transformation and industrialisation, most notably in East Asian (Japan, South Korea, Taiwan, and Singapore) and Southern Europe. There are also episodes of decline, such as those of Haiti, Venezuela, and of many African countries. Yet, as Figure 1a reveals, such episodes are the exception rather than the rule; output per capita is quite persistent. And, if anything, persistence is stronger, if one examines output in levels rather than in per capita terms.

Even if one examines the link between current GDP per capita (p.c.) and pre-WWI output p.c., the correlation remains quite high. Using data from Angus Maddison, Figure 1b illustrates the correlation of log GDP p.c. in 2008 and in 1913; the elasticity is 0.90 and the in-sample fit is considerable, with an R-squared of around 0.50. While there are some evident success cases (the East Asian countries) and some failure cases (Africa, Nepal, North Korea, and Argentina), most countries are close to the regression line. And, as Figure 1c illustrates, the pattern is similar when we associate contemporary log GDP p.c. (in 2008) with GDP p.c. before the first wave of globalisation and during the Second Industrial Revolution (in 1870). The elasticity is almost 1 and log GDP in 1870 again explains half of the variability in contemporary GDP p.c. (R-squared continues to be around 0.50). The correlation falls only when we go back to the 1820s, when most countries were still in the pre-industrial phase and per capita income differences were negligible compared to today’s large differences, even then the correlation exceeds 0.35
The Long Economic and Political Shadow of History—Volume I. A Global View

Figure 1a

REGRESSION FIT: \( \log(\text{GDP}_{\text{cap2014}}) = 1.8537 + 0.9242 \times \log(\text{GDP}_{\text{cap1960}}) + \epsilon \)  
R-squared = 0.50  
Source: PWT

Figure 1b

REGRESSION FIT: \( \log(\text{GDP}_{\text{cap2008}}) = 2.4735 + 0.9063 \times \log(\text{GDP}_{\text{cap1913}}) + \epsilon \)  
R-squared = 0.51  
Source: Maddison
The patterns in Figures 1a-1c summarise (in a crude way) a large body of research revealing considerable persistence in income, output, and urbanisation, during at least the post-colonial period. There seems to be a ‘structural break’ only when one goes back to the Columbian exchange and the early stages of colonisation, a pattern that Acemoglu et al. (2002) have eloquently coined ‘reversal of fortune’. While obtaining reliable data on income and output for the pre-colonial times becomes tricky and there are natural questions on data quality, many works reveal persistence, even when one takes a very long-run perspective (Comin et al. 2010, and Ashraf and Galor 2013). Perhaps more importantly, there is persistence on income per capita and population density, even when one looks across regions within the same countries. For example, Maloney and Valencia (2015) report strong correlations between pre-colonial and contemporary population densities across regions in 18 countries in North, Central and Latin America (population density is the appropriate metric for development in Malthusian pre-Industrial societies). For 12 (7) countries the correlation exceeds 0.5 (0.75). Similarly, Broda and Weinstein (2002) report high correlations on population densities across Japanese regions over a four-century period.

Chanda, Cook, and Putterman (2014) show that when one conducts the analysis at the population (rather than the territorial) level, then persistence increases.
Maloney and Valencia (2015) also uncover strong positive correlations between contemporary regional (log) income and (log) pre-colonial population density. Likewise, Tabellini (2010) finds sizable persistence between contemporary regional income across eight European countries and historical proxies of development.

The evidence on persistence hints that there are deep-rooted, historical factors shaping development, at least partly. This finding is, at first-glance, puzzling, as there have been watershed events in the past couple of centuries, including colonisation, two world wars, numerous regional conflicts, revolutionary regime changes, the Industrial Revolution, the tremendous rise in globalisation. During the past two centuries the world population experienced an unprecedented increase in standards of living, as incomes rose, poverty fell (especially since the 1990s), fertility rates declined, and life expectancy and health improved. Yet the gains have been highly asymmetric, and, if one takes a medium to long-run perspective, there is evident divergence. This is revealed in Figures 2a-2b where, using the historical GDP series of Maddison, we plot output per capita since 1870 for continents and some countries.

**Figure 2a**

![Graph showing long-term GDP per capita from 1870 to 2000 for different regions including Western Offshoots, Western Europe, Latin America, Asia, and Africa.](source: Maddison)
Understanding persistence appears crucial, therefore, as a combined reading of Figures 1a-1c with Figures 2a-b suggests that countries/regions with a small-to-modest head start managed to gain the most, both from the technological innovations of the 19th and 20th centuries and from the global spur of commerce.

2.2. Explanations

The literature has bundled the potential explanations of persistence and divergence into four broad (and quite often interrelated) categories (see Acemoglu and Robinson 2012):

i. those related to historical institutions, such as colonial extractive rules, pre-colonial political legacies and enslavement;

ii. those stressing the role of cultural traits, related to religion, trust, family ties, beliefs and norms;

iii. those related to geography-ecology, such as the impact of floods, temperature shocks, terrain ruggedness and isolation, soil variability, etc.;

iv. others that emphasise historical accidents, such as the artificial drawing of colonial borders in Africa.
The chapters of this eBook series summarise works that explain contemporary economic development (as well as political development and attitudes) via different mechanisms, some geographic, some related to cultural traits, some stressing the legacies of institutional features, and some emphasising the role of accidental events. We view these contributions as complementary and, as many of the chapters will reveal, culture and institutions are interlinked and quite often share geographic origins (see Alesina and Giuliano 2016, Ashraf and Galor 2016).

Volume 1 of the eBook series starts with chapters reviewing works on the spatial distribution of economic activity across the globe and its main correlates. Next, it includes chapters summarising works which explore watershed events that have global repercussions. The chapters in Volume 2 summarise research on the deep origins of African development as well as works on the legacy of colonial practices in India, China, and Australia. Volume 3’s chapters summarise studies on historical legacies in Latin and North America and then surveys works on the shadow of history in Europe.
References


About the authors

Stelios Michalopoulos is an Associate Professor of Economics at Brown University, a Faculty Research Fellow at the NBER and a Research Affiliate of the CEPR. A native of Argos, Greece, Stelios holds a B.A. from the Athens University of Economics and Business and a Ph.D. in Economics from Brown University. After completing his doctorate in 2008, he joined the Department of Economics at Tufts University as an Assistant Professor. In 2010-2011 he was the Deutsche Bank Member at the Institute for Advanced Study at Princeton.

His primary research interests intersect with political economy, growth and development, and the economics of culture. He has published in leading peer-reviewed economic journals including the American Economic Review, Econometrica, Journal of Political Economy and the Quarterly Journal of Economics.

Elias Papaioannou is Professor of Economics at the London Business School (London, United Kingdom). He is also a research fellow of the CEPR (Centre for Economic Policy Research) and the NBER (National Bureau of Economic Research).

He holds an LL.B. from the law school of the National and Kapodistrian University of Athens, Greece, a Master’s in Public Policy and Administration (MPA) with a concentration in international economics from Columbia University, and a Ph.D. in economics from the London Business School. After the completion of his doctorate in 2005 he worked for two years at the Financial Research Division of the European Central Bank (ECB) in Frankfurt, Germany. From 2007 till 2012 he served as Assistant Professor of Economics at Dartmouth College (NH, USA), while during the 2010-2011 and 2011-2012 academic years he was a Visiting Assistant Professor at the Economics Department of Harvard University (MA, USA).

His research interests cover the areas of international finance, political economy, applied econometrics, macro aspects of regulation, law and finance, and growth and development. He has published in many leading peer-refereed journals, such as Econometrica, the Quarterly Journal of Economics, the Journal of Political Economy, the American Economic Review, Journal of Finance, the Economic Journal, the Review of Economics and Statistics, the Journal of Development Economics, the Journal of the European Economic Association, the Journal of International Economics, and more. His work has also appeared in numerous edited book volumes.
His research has been recognized with the inaugural 2013 European Investment Bank Young Economist Award, 2005 Young Economist Award by the European Economic Association and the 2008 Austin Robinson memorial prize by the Royal Economic Association. Elias consultants for international organisations, investment banks, hedge funds, and institutional investors on macroeconomic developments in the EU and Greece.
1 Introduction

Stelios Michalopoulos and Elias Papaioannou
Brown University and CEPR; London Business School and CEPR

The first part of Volume II discusses works on the deep origins of African development, including the slave trades, the colonial experience, and independence movements (Chapters 1-7). The second part of Volume II goes over studies on the long-lasting legacies of some important aspects of colonial history in India (Chapter 8), China (Chapter 9), and Australia (Chapter 10).

In Chapter 1, Nathan Nunn summarises works on the legacy of Africa’s slave trades (1400-1900). Drawing on Nunn (2008) he first presents evidence revealing the lasting toxic effects of Africa’s four waves of slave trades on contemporary development. African regions that today include countries such as Angola, Nigeria, and Sierra Leone were profoundly affected by the slave trades, and are less developed when compared to those territories that, due to isolation, were shielded from enslavement, such as today’s region of Botswana. Nunn further shows that, if anything, the slave trades had more effect on the relatively more advanced parts of pre-colonial Africa, thus linking enslavement to a continental reversal of fortune. Second, Nunn reviews empirical studies that shed light on the mechanisms at play, suggesting that the enslavement-underdevelopment nexus operates via distrust, conflict, a culture of authoritarianism, and a higher degree of ethnic fractionalisation.

1 We need to express our gratitude to all the authors who have contributed to this ebook. A special thanks to Charles Wyplosz and Richard Baldwin for pushing us to undertake this project to collect non-technical summaries for some of the major contribution revealing the legacy of important historical events. Due to space and other constraints some important works are not covered by this eBook. We would like also to thank Sophie Roughton, Simran Bola, and Tessa Ogden from CEPR for helping us.
In Chapter 2, we review works that examine the impact of the artificial drawing of colonial borders, delineated by Europeans during the Scramble for Africa in the late 19th century. First, we summarise our recent work (Michalopoulos and Papaioannou 2016), showing that ethnic partitioning is associated with current conflict. Combining anthropological maps of the spatial distribution of African ethnicities, at the time of colonisation, with contemporary geo-referenced data on civil conflict and exploiting within-country (across ethnic region homeland) variation we find that the likelihood, duration and intensity of political violence are all higher in the ancestral homelands of split-by-the border ethnicities. We also show that ethnic partitioning is related to ethnic-based discrimination, and repression, often leading to major ethnic-tainted civil wars. We also show complementary results, revealing that individuals who identify with partitioned groups have worse educational outcomes and lower access to basic public goods. Second, we discuss other-than-ethnic-partitioning deleterious consequences of the Scramble for Africa, including the peculiar sizes and shapes of many African countries, the preponderance of landlocked states, and their ethnic composition.

In Chapter 3, Tristan Reed reviews works on the legacy of ‘indirect rule’ – a colonial strategy that involved keeping in place existing rulers and ruling through them. [This chapter is closely related to Laksmi Iyer’s chapter on direct and indirect rule in India (Volume II, Chapter 8)]. This chapter draws on Acemoglu et al. (2014), who examine the chieftaincy system in Sierra Leone from the colonial era till today. The authors first show that political power locally, in contemporary Sierra Leone, is in the hands of families whose ancestors were appointed by the British as paramount chiefs during the late 19th century. These families have effectively ruled the local communities, not only during the colonial time (as British representatives), but also throughout most of its post-independence period. Second, the authors show that political competition reflected in the number of ruling families in the chieftaincy is correlated with better development outcomes. Third, the authors document that in places with more powerful chiefs (and fewer ruling families) social capital is unexpectedly higher, either because a dominant local elite there controls civic society or because trust and civicness substitute for lower levels of political participation.
In Chapter 4, Sara Lowes and Eduardo Montero summarise their ongoing work on the legacy of forced labour practices and terror, focusing on King Leopold’s rule over Congo Free State. [This chapter is closely related to Melissa Dell’s chapter on the impact of the *mita* system in Peru (Volume III, Chapter 2)]. Lowes and Montero (2017) compare various development proxies (education, wealth, health), between two different sets of areas: those that were violently ruled via mercenaries, employed by the private companies that had concessions for rubber, compared to nearby territories, that did not have to produce pre-specified amounts of rubber. Living conditions today are considerably worse in areas where mercenaries were coercing the local population to produce rubber, as compared to neighbouring areas that were not part of Belgian concessions. They also show that traditional leaders (chiefs) are less accountable and of lower quality in regions that suffered a great deal during King Leopold’s rule, suggesting that dysfunctional local institutions may be a source of persistence. Interestingly, however, the authors find a higher level of trust and civicness in the concessionary areas, suggesting that, in this case at least, trust and local institutions are substitutes.

In Chapter 5, Remi Jedwab, Edward Kerby, and Alex Moradi explore the role of the effect of colonial investments in railroads on African economic geography, drawing mostly on their recent work (Jedwab *et al.* 2016, and Jedwab and Moradi 2016). Connecting large cities and promoting inter-city trade was not of interest to colonial powers. Military considerations and accessibility to mine-rich areas were more important factors and, as the authors document, quite often African colonial railroads were going from ‘nowhere to nowhere’. This initial railroad network nevertheless helped to shape population density and the location of economic activity during colonial times. Cities were formed close to the railroads and, over time, they developed faster than areas far from the railroads. But the most interesting finding is in regard to the most recent period. Even though, post-independence, many railroads fell out of use (because of conflict, mismanagement, and lack of resources) and the road system took over in importance, locations along the old railroad lines are still more populous and more developed. This finding, which is consistent with new economic geography theoretical models on inertia, reveals how early colonial policies (on infrastructure in this case) can have long-lasting effects.
In Chapter 6, Julia Cagé and Valeria Rueda discuss the role of Christian missions in Africa (see Woodberry 2012, on the lasting effect of missions on beliefs and institutions). During colonial times, Christian missions played a key role in the provision of education and health throughout the world (besides conversion). [See also the discussion of Felipe Valencia on the long-lasting effects of Christian missions in South America (Volume III, Chapter 3).] Drawing on Cagé and Rueda (2016a, b), they present evidence on the heterogeneous effects of Christian missions. First, they show that proximity to historical Protestant missions with a printing press is associated with higher levels of trust, higher readership of newspapers, and higher levels of social capital today. In contrast, proximity to Protestant missions without a printing press is unrelated to social capital proxies. This result conforms well with the important work of Dittmar (2011) on the quantitatively strong impact of the printing press on urbanisation and growth in Europe. Second, they report on ongoing research examining the effects of Christian missions on contemporary health, focussing on HIV prevalence in particular.

In Chapter 7, Leonard Wantchekon and Omar García-Ponce assess how differences in the way that the colonial era came to an end across African countries have left their imprint on subsequent institutional outcomes. Specifically, they show that contemporary political development in Africa is related to the type of independence liberation movements (Wantchekon and García-Ponce 2014). Countries that experienced rural insurgencies, where the liberation movement was especially strong in the rural areas and the leaders tried to weaken the colonial state with rural fighting (as in Kenya, Madagascar and Mozambique), have more autocratic regimes today. In contrast, African countries where the independence movement took the form of non-violent urban protests (such as in Ghana and Tanzania) have more democratic governments. To push on causation, the authors apply an econometric technique that isolates the component of rural-urban insurgency driven by geography.

In Chapter 8, the geographic focus switches from Africa to India where Lakshmi Iyer summarises her works on the long-run consequences of direct versus indirect colonial rule in the subcontinent (Iyer 2010, Banerjee et al. 2005). Iyer explores the fact that 45% of the total area and 23% of the Indian population were not under direct British colonial rule and to achieve identification she takes advantage of a historical policy change, according to which annexing of native states by the British was allowed for a certain period, in cases where the ruler died without a natural born heir.
Her findings suggest that areas that came under direct British rule have significantly lower public goods provision, reflected in the access to health centres and transportation infrastructure. Iyer offers a rationale for the uncovered pattern. Rulers of princely states were, in fact, deposed by the British for misrule, as opposed to the case for British colonial administrators, who did not face any significant penalty for delivering poor outcomes.

In Chapter 9, Ruixue Jia traces the economic performance of prefectures with treaty ports in China over time, uncovering how oscillating between periods of openness and closedness shaped development outcomes both in the short and long run. Jia constructs a prefecture-level dataset from 1776 and 2000 and finds that, when comparing prefectures with treaty ports to those with no-treaty ports, while those with treaty ports grew faster before the Communist Revolution of 1949, the two sets of regions grew at a similar pace when all parts of China were closed to foreign influence after 1949. Nevertheless, when China opened up after 1980, it was the prefectures with treaty ports which were among the first to take advantage of the new globalisation opportunities. Why was this the case? Since differences in geographical endowments and measurable institutional features are similar in both sets of prefectures, Jia suggests that human capital and social norms of doing business are more likely to have played a role in explaining the long-run impacts.

Finally, in Chapter 10 of this volume, Pauline Grosjean investigates the origins of present-day gender roles in Australia, reviewing her recent work with Rose Khattar and Roberto Brooks (Grosjean and Khattar 2015, and Grosjean and Brooks 2016). Grosjean starts with the following striking observation: in the late 18th and 19th centuries, the British policy of convict deportation to Australia resulted in heavily male-biased sex ratios. Men vastly outnumbered women among convicts, with more than five males for every female. Has the surplus of males in the past shaped cultural attitudes towards gender roles today? Grosjean shows that in the short-run the presence of male-biased sex ratios induced more women to get married and so reduced their labour supply. Today, in Australia, even though such gender imbalances have long ceased to exist, people have more conservative attitudes towards women working and women still do less work in the labour market in areas that were more male-biased in the past.
Between 1400 and 1900, the African continent experienced four sizeable slave trades. The largest and best-known was the trans-Atlantic slave trade where, beginning in the fifteenth century, slaves were shipped from West Africa, West Central Africa, and Eastern Africa to the European colonies in the New World. The three other slave trades – the trans-Saharan, Red Sea, and Indian Ocean slave trades – were smaller in scale and predated the trans-Atlantic slave trade. During the trans-Saharan slave trade, slaves were taken from south of the Saharan desert and shipped to Northern Africa. In the Red Sea slave trade, slaves were taken from inland of the Red Sea and shipped to the Middle East and India. In the Indian Ocean slave trade, slaves were taken from Eastern Africa and shipped either to the Middle East, India, or the plantation islands in the Indian Ocean. In total, close to 20 million slaves were taken from the continent (Nunn, 2008). According to the best estimates, by 1800 Africa’s population was half of what it would have been, had the slave trades not occurred (Manning, 1990).

Slaves were captured through kidnappings, raids, and warfare. A summary of the method of enslavement among a sample of 144 former slaves is provided in Table 1. Historical accounts suggest that the pervasive insecurity, violence and warfare had detrimental impacts on the institutional, social, and economic development of societies. There are numerous examples of the slave trades causing the deterioration of domestic legal institutions, the weakening of states, and political and social fragmentation (e.g., Inikori 2000, 2003 and Heywood 2009).
Table 1. The methods of enslavement of Koelle’s Informants

<table>
<thead>
<tr>
<th>Manner of Enslavement</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taken in a war</td>
<td>24.3%</td>
</tr>
<tr>
<td>Kidnapped or seized</td>
<td>40.3%</td>
</tr>
<tr>
<td>Sold/tricked by a relative, friend etc.</td>
<td>19.4%</td>
</tr>
<tr>
<td>Through a judicial process</td>
<td>16.0%</td>
</tr>
</tbody>
</table>

Notes: The data are from Sigismund Koelle’s Linguistic Inventory. The sample consists of 144 informants, interviewed by Koelle, for which their means of enslavement is known.

The most illustrative example of this is the experience of the Kongo Kingdom, which was contacted in 1493 by Diogo Cao. Initially, a diverse array of products was traded between the Kongo Kingdom and the Portuguese, including copper, textiles, ivory, and slaves. At first, the only slaves to be traded were prisoners of war and criminals. However, the Portuguese demand for slaves, the pervasiveness of slave traders and merchants, and competition for the throne within the Kingdom, all resulted in a dramatic and uncontrollable increase in slave capture and raiding throughout the Kingdom. By 1514, King Afonso (the Kongo King) had already written to the King of Portugal, complaining that the Portuguese merchants were colluding with noblemen to illegally enslave Kongoese citizens. In 1526, in an effort to end the trade, King Afonso requested the removal of all Portuguese merchants. In the end, his efforts were unsuccessful. Large scale slave-raiding continued unchecked into the 16th century, when it culminated in the Jaga invasion of 1568-1570. Large-scale civil war ensued from 1665-1709, resulting in the collapse of the once-powerful Kingdom (Heywood 2009).

An empirical literature has emerged that aims to supplement these historical accounts with quantitative estimates of the long-run impact of Africa’s slave trades. The first paper that attempted to provide such estimates was Nunn (2008). In the study, an empirical test was undertaken, with the following logic. If the slave trades are partly responsible for Africa’s current underdevelopment, then, looking across different parts of Africa, one should observe that the areas that are the poorest today should also be the areas from which the largest number of slaves were taken in the past.

To undertake this study, Nunn had to first construct estimates of the number of slaves taken from each country in Africa during the slave trades (i.e., between 1400 and 1900).
These estimates were constructed by combining data on the number of slaves shipped from each African port or region with data from historical documents that reported the ethnicity of over 106,000 slaves taken from Africa. Figure 1 provides an image showing a typical page from these historical documents. The documents shown are slave manumission records from Zanzibar. Each row reports information for one slave, including his/her name, ethnicity, age, etc.

After constructing the estimates and connecting these with measures of modern day economic development, the study found that, indeed, the countries from which the most slaves had been taken (taking into account differences in country size) were today the poorest in Africa. This can be seen in Figure 2, which is taken from Nunn (2008). It shows the relationship between the number of slaves taken between 1400 and 1900 and average real per capita GDP measured in 2000. As the figure clearly shows, the relationship is extremely strong. Further more, the relationship remains robust when many other key determinants of economic development are taken into account.

**Figure 1.** Excerpt from document taken from Zanzibar archives
An alternative interpretation of the relationship shown in Figure 2 is that the parts of Africa from which the largest number of slaves were taken were initially the most underdeveloped. Today, because those characteristics persist, these parts of Africa continue to be underdeveloped and poor.

Nunn also tested this alternative hypothesis by checking whether it was, in fact, the initially least developed parts of Africa that engaged most heavily in the slave trades. He finds that the data suggest that, if anything, it was the parts of Africa that were initially the most developed, not least developed, that supplied the largest number of slaves.

In the analysis, Nunn also used a statistical technique called instrumental variables estimation to identify the causal effect of the slave trade on economic development. The findings from the instrumental variables estimates suggested that increased extraction during the slave trades did, indeed, cause worse economic performance. Overall, the conclusion from the analyses is that the relationship shown in Figure 2 is most likely causal and not spurious.
The estimates from Nunn (2008) allowed us, for the first time, to understand, quantitatively, the impacts the slave trades had on economic development. In particular, using the estimates, one is able to calculate how much more developed Africa would be, if the slave trades had not taken place. These calculations were undertaken by Nunn (2010). As an initial step in the calculation, first note that the mean level of average per capita income of the countries in Africa is $1,834 (measured in 2000). This is significantly lower than the income for the rest of the world (which is $8,809), and it is much lower even than the income of other developing countries (which is $4,868), where a developing country is defined as one with less than $14,000 in per capita income. According to the estimates from Nunn (2008), if the slave trades had not occurred, then 72% of the average income gap between Africa and the rest of the world would not exist today, and 99% of the income gap between Africa and other developing countries would not exist. In other words, had the slave trades not occurred, Africa would not be the most underdeveloped region of the world and it would have a similar level of development to Latin America or Asia.

The findings from Nunn (2008) provided suggestive evidence that much of Africa’s poor performance can be explained by its history, which is characterised by over 400 years of slave raiding. Although his analysis stopped short of providing a definitive understanding of exactly how and why the slave trades were so detrimental to economic development, he did provide suggestive evidence that was consistent with historical accounts of the slave trades resulting in a weakening and underdevelopment of political structures, as well as in impeding the formation of broader ethnic groups, leading to more ethnic diversity. The study showed that the countries from which the largest numbers of slaves were taken were also the areas that had the most underdeveloped political structures at the end of the 19th century, as well as being the countries that are the most ethnically fragmented today.

A number of studies followed on from the research in Nunn (2008), by attempting to better understand the channels through which the slave trades affect economic development today. Green (2013) revisited the impact of the slave trade on ethnic fractionalisation. He extended the findings in Nunn (2008) by examining the entire world, and constructing estimates of the export of slaves in all countries in the world. Using these data, he also found a strong relationship between slave exports and greater ethnic fractionalisation.
According to his estimates, all of the difference in ethnic fractionalisation between Africa and the rest of the world can be explained by its experience with the slave trades.

In a series of studies, Whatley and Gillezeau (2011) and Whatley (2014) combine slave shipping records with ethnographic data and estimate the relationship between slave shipments and institutional quality and ethnic diversity in the locations close to the ports of shipment. Their analysis, consistent with Nunn (2008) and Green (2013), indicates that the slave trades did result in greater ethnic fractionalisation. In addition, their analysis also shows that the slave trades resulted in a deterioration of local ethnic institutions, measured in the late pre-colonial period.

Another subsequent study, undertaken by Nunn and Wantchekon (2011) asks whether the slave trades resulted in a deterioration of trust. Their line of enquiry was motivated by two facts. The first is that trust has been long-hypothesised to be an important foundation for economic prosperity (e.g., Algan and Cahuc 2010). The second was the fact that, during the slave trades, individuals frequently turned on one another, kidnapping, tricking, and selling each other into slavery. The existing historical evidence indicates that these forms of betrayal were common. For example, among the sample of slaves reported in Table 1, 20% became slaves because they were tricked by a family member or friend.

In their study, Nunn and Wantchekon extended the data construction efforts of Nunn (2008) and constructed estimates of the number of slaves taken from each ethnic group in Africa (rather than country). The ethnicity level estimates are displayed visually in Figure 3. The analysis combined the ethnicity-level slave export estimates with fine-grained household survey data, which reports individuals’ trust of those around them, whether neighbours, relatives, local governments, co-ethnics, or those from other ethnicities. The study documented a strong negative relationship between the intensity of the slave trade among one’s ethnic ancestors and an individual’s trust in others today.

The study then attempted to distinguish between the two most likely channels through which the slave trades could have adversely affected trust. One is that the slave trades made individuals and their descendants inherently less trusting. That is, it created a culture of distrust. In the insecure environment of the slave trade, where it was common to experience the betrayal of others, even friends and family, greater distrust may have developed, which could persist over generations even after the end of the slave trade.
Another possibility is that the slave trades may have resulted in a long-term deterioration of legal and political institutions, which are then less able to enforce good behaviour among citizens, and as a result people trust each other less today.

The study undertook a number of different statistical tests to identify the presence and strength of the two channels. They found that each of the tests generated the same answer: both channels are present. The slave trades negatively affected domestic institutions and governance, which results in less trust today. In addition, the slave trade also directly reduced the extent to which individuals were inherently trusting of others. They also found that, quantitatively, the second channel is twice as large as the first channel.

The ethnicity-level slave export estimates constructed by Nunn and Wantchekon spurred a second line of research that looked into other impacts of the slave trade, using the new finer-grained data. For example, Zhang and Kibriya matched the ethnicity-level slave export data from Nunn and Wantchekon (2008) with fine-grained data on the location and intensity of conflicts within Africa from 1997-2014. They find a strong positive relationship between the intensity of the slave trades and the prevalence of civil conflict today. Given the strong link between income and conflict, this likely represents an important channel behind the reduced-form relationship between the slave trades and income documented in Nunn (2008).

A number of subsequent studies have also explored other cultural consequences of the slave trades. For example, research by Edlund and Ku (2011), Dalton and Leung (2014), and Bertocchi and Dimico (2015) found that the trans-Atlantic slave trade resulted in a long-term increase in the prevalence of polygyny (i.e., the practice of men having multiple wives). This is due to the fact that it was primarily males who were captured and shipped to the Americas, resulting in a shortage of men and skewed sex ratios within many parts of Africa. Interestingly, Dalton and Leung (2014) found that there is no evidence of such an impact for the Indian Ocean slave trade, where there was not a strong preference for male slaves. This has led the authors to conclude that Africa’s history of the slave trades is the primary explanation for why today polygyny is much more prevalent in West Africa than in East Africa.
Figure 3. Maps showing the number of slaves taken from each ethnic group in Africa during the trans-Atlantic and Indian Ocean slave trades.

3a Indian Slave Exports

3b Atlantic slave exports
The study by Bertocchi and Dimico (2015) is particularly interesting, because it takes the analysis one step further and shows that the greater prevalence of polygyny, which arose due to the slave trade, has led to an increase in HIV rates. They document that women who are in a polygynous relationship are more likely to have sexual partners other than their husband. This is perhaps not surprising, since in polygamous societies wealthy older men are typically the ones who have multiple wives, leaving a large number of young able-bodied men, without wives, able and willing to engage in extramarital relationships. Thus, according to the authors’ findings, the trans-Atlantic slave trade is an important factor in explaining the high rates of HIV in Africa today.

A study by Teso (2016) also examined the consequences of the skewed sex ratio that resulted during the trans-Atlantic slave trade, but focused on the impacts on female labour force participation. The shortage of males in the population, during the trans-Atlantic slave trade, meant that women had to undertake many of the tasks traditionally performed by men, such as those in agriculture, or in the military, and to accept positions of leadership and authority. The Dahomean female ‘Amazon’ army, shown in Figure 2, is the best-known example of this. The female army was established by Dahomey in the 17th century at the height of the slave trade, and their existence continued until the end of the 19th century. Teso’s (2016) research shows that those parts of Africa that experienced the trans-Atlantic slave trade most severely have higher rates of female labour force participation today.

Figure 4. A photo of the female army of Dahomey, who were often referred to as ‘Amazons’
A number of studies have turned the line of research started by Nunn (2008) on its head by asking, not ‘what were the consequences of the slave trades?’, but ‘what were the causes of the slave trades?’. Fenske and Kala (2015) examined what role weather played by combining information on slave shipments with data on annual historical temperatures, measured at a 5-degree grid-cell level. They found that negative weather shocks (taking the form of abnormally cold years) increased the supply of slaves from the locations that experienced these shocks. They also find that locations that experienced such shocks during the height of the slave trade tend to have lower income levels today. In a separate study, the same authors examined how the suppression of the slave trade after 1807 affected slave shipments (Fenske and Kala 2016). Not surprisingly, the suppression led to a decline in slave exports, although this was focused in British-controlled parts of the continent. The authors also showed that the decline in slave exports occurred alongside an increase in violence and coercion within the continent. Thus, the evidence suggests that not only did the slave trade have detrimental effects, but possibly its abolition did too.

A study by Nunn and Puga (2015) has examined the extent to which geography affected the capture and export of slaves during the slave trades. The authors used the data from Nunn (2008) to document how uneven terrain, taking the form of cliffs, ridges, and escarpments, was used by societies to defend against and escape from slave raiders. The uneven terrain could be used to help build defensive fortifications and to provide places to escape, in the face of slave raids. They show that, as a consequence, those places in Africa that are more rugged were better able to escape the slave trade and, thus, are richer today.

Although research understanding the long-term impacts of Africa’s slave trades is still in progress, the evidence accumulated up to this point suggests that this historic event played an important part in the shaping of the continent, in terms of not only economic outcomes, but cultural and social outcomes as well. The evidence suggests that it has affected a wide range of important outcomes, including economic prosperity, ethnic diversity, institutional quality, the prevalence of conflict, the prevalence of HIV, trust levels, female labour force participation rates, and the practice of polygyny. Thus, the slave trades appear to have played an important role in shaping the fabric of African society today.
References


Bertocchi, Graziella and Arcangelo Dimico (2015), The Long-Term Determinants of Female HIV Infection in Africa: The Slave Trade, Polygyny, and Sexual Behavior, IZA DP No. 9102


Edlund, Lena and Hyejin Ku (2011), The African Slave Trade and the Curious Case of General Polygyny, MPRA Paper NO. 52735


About the author

Nathan Nunn is Frederic E. Abbe Professor of Economics at Harvard University. Professor Nunn was born in Canada, where he received his PhD from the University of Toronto in 2005. Professor Nunn’s primary research interests are in economic history, economic development, cultural economics, political economy and international trade. He is an NBER Faculty Research Fellow, a Research Fellow at BREAD, and a Faculty Associate at Harvard’s Weatherhead Center for International Affairs (WCFIA). He is currently a co-editor of the *Journal of Development Economics*.

One stream of Professor Nunn’s research focuses on the long-term impact that historical factors have on current economic development. In particular, he has studied how history shapes the evolution of institutions and cultures across societies. He has published research empirically examining the historical foundations of a wide range of current outcomes, including distrust, gender norms, religion, rule following, conflict, support for democracy, and effectiveness of foreign aid.

A second stream of Professor Nunn’s research examines economic development in a contemporary context. He has published research examining the effects of Fair Trade certification, CIA interventions during the Cold War, foreign aid, and industrial policy.

A third stream of Professor Nunn’s research focuses on the importance of contracting institutions for international trade. He has published research showing that a country’s ability to enforce written contracts is a key determinant of comparative advantage, as well as research examining how contracting frictions affects firms’ decisions to engage in FDI versus outsourcing.
3 The contemporary shadow of the Scramble for Africa

Stelios Michalopoulos and Elias Papaioannou
Brown University and CEPR; London Business School and CEPR

When economists debate the long-lasting legacies of colonisation, the discussion usually revolves around the establishment of those ‘extractive’ colonial institutions that outlasted independence (e.g., Acemoglu et al. 2001), the underinvestment in infrastructure (e.g., Jedwab and Moradi 2016), the identity of colonial power (e.g., La Porta et al. 2008) and the coloniser’s influence on early human capital (Easterly and Levine 2016). Following the influential work of Nunn (2008), recent works have explored the deleterious long-lasting consequences of Africa’s slave trades (see Nunn 2016, for an overview). Yet, between the slave-trade period (1400-1800) and the arrival of the colonisers at the end of the 19th century, the Scramble for Africa stands out as a watershed event in the continent’s history. The partitioning of Africa by Europeans starts, roughly, in the 1860s and is completed by the early 1900s. The colonial powers signed hundreds of treaties, which involved drawing on maps the boundaries of colonies, protectorates, and ‘free-trade’ areas of a largely unexplored and mysterious continent (see Wesseling 1996, for a thorough discussion). In this context it is perhaps not surprising that many influential scholars of the African historiography (e.g., Asiwaju 1985, Wesseling 1996, Herbst 2000) and a plethora of case studies,

1 In the African context, Europeans applied forced-labour practices in Congo (Lowes and Montero 2016), Rwanda (Blouin 2016), and other countries; and in many cases they established extractive indirect-rule colonies and protectorates, where small ethnic minorities and local chiefs controlled the interior using brutal methods (e.g., Acemoglu et al. (2014), Mandani 1996).

2 In Michalopoulos and Papaioannou (2016, 2017) we provide a detailed overview of the key events surrounding the Scramble for Africa. Pakenman (1991) offers a comprehensive account.
suggest that the most consequential aspect of European involvement in Africa was not colonisation *per se*, but the erratic border designation that took place in European capitals in the late 19th century.

At the time of the Scramble for Africa, Europeans had limited knowledge of the local political, economic and geographic conditions. Furthermore, they were in a rush to partition the continent, in order to finance industrialisation in Europe and the Americas. As such, they did not even wait for reports from the great explorers who were then mapping the continent. Whilst conspiracy theories abound, A. I. Asiwaju (1985) summarises the consensus view of the African historiography: “the study of European archives supports the accidental rather than a conspiratorial theory of the marking of African boundaries.” A key objective shared by the colonisers, with memories of the Napoleonic Wars fresh in their minds, was to prevent conflict amongst themselves over African territory. As British Foreign Secretary and Prime Minister, Lord Salisbury, famously put it:

“We have engaged in drawing lines upon maps where no white man’s feet have ever trod. We have been giving away mountains and rivers and lakes to each other, only hindered by the small impediment that we never knew exactly where the mountains and rivers and lakes were.”

While, upon African independence, policy makers, scholars and development experts argued for the need to redraw borders, departing colonisers did not want to open this perceived ‘Pandora’s box’; and many of the leaders of African independence harboured a firm belief that nation-building policies, industrialisation, and a notion of Pan-Africanism would eventually attenuate ethnic differences and animosity.

---

3 Frankema *et al.* (2015) point to the global developments during the Scramble for Africa, including the commodity price boom in the decades before the Berlin Conference of 1885 (the latter has come to symbolise the partitioning of the continent). Bolt (2013) discusses the key driving forces of the Scramble and the role of the steam boat, medical innovations, iron metallurgy, etc.

4 In line with this, Alesina, *et al.* (2011) estimate that around 80% of African borders follow latitudinal and longitudinal lines, more so than in any other part of the world.

5 Wantchekon and Garcia-Ponce (2016) discuss how differences in the type of independence movements have shaped subsequent political, institutional, and economic developments.
Jeffrey Herbst (2000) summarises: “for the first time in Africa’s history [at independence], territorial boundaries acquired salience...The boundaries were, in many ways, the most consequential part of the colonial state.”

Recent research shows that the Scramble for Africa has contributed to economic, social and political underdevelopment, via several channels. First, by partitioning numerous ethnic groups into more than one contemporary country\(^6\), it has spurred ethnic-tainted civil conflict and also promoted repression and ethnic-based discrimination (Michalopoulos and Papaioannou 2016). Second, the Scramble for Africa crucially shaped the ethnic composition of the newly independent states (leading to ethnic polarisation, fractionalisation, and inequality), which, in turn, has influenced institutional and economic development, public goods provision, and conflict. Third, the Scramble resulted in Africa having a large share of landlocked countries, a situation which crucially inhibits their access to global markets (Collier 2007). Fourth, as a consequence of the artificial colony/protectorate design, many African countries are either tiny or absurdly large, while others have peculiar shapes (Herbst 2000); these features limit the reach of a state beyond its capital, and thus weaken the enforcement of laws, which, in turn, impedes development (Michalopoulos and Papaioannou 2014).

In this chapter we summarise the key findings of recent empirical works that use high-resolution geo-referenced data and econometric methods to estimate the long-lasting impact of the various aspects of the Scramble for Africa.

1. Ethnic partitioning and civil conflict

1.1 Identifying partitioned ethnicities

Quantifying the impact of the Scramble for Africa requires the identification of partitioned groups in a systematic manner. In Michalopoulos and Papaioannou (2016) we project contemporary country borders onto George Peter Murdock’s Ethnolinguistic Map (1959), which depicts the spatial distribution of African ethnicities at the time of colonisation, during the late 19\(^{th}\) and early 20\(^{th}\) century (Figure 1a).

\(^6\) Alesina et al. (2011) estimate that partitioned groups constitute, on average, 40% of a country’s population.
We classify as partitioned those ethnicities with at least 10% of their ancestral landmass belonging to more than one country. As a result, out of the 825 mapped groups, 229 are classified as partitioned (Figure 1b). There is clearly noise in Murdock’s map and the map does not allow for ethnic overlapping, something that obviously was the case; moreover, Africans evidently could, and did, move. Yet, although not perfect, this simple and transparent procedure identifies most major partitioned ethnic groups, based on both qualitative evidence (Asiwaju 1985) and casual empiricism. For example, the Maasai have been split between Kenya and Tanzania, the Anyi between Ghana and the Ivory Coast, and the Chewa between Mozambique, Malawi, and Zimbabwe.

**Figure 1a.**
1.2 A primer on border artificiality

The African historiography provides ample case-study evidence on the artificial nature of colonial borders; but in some cases Europeans did try to accommodate local conditions. Hence, in Michalopoulos and Papaioannou (2016) we examine whether partitioned and non-split homelands differ with respect to various geographic, ecological, political, and other historical features, since this may shed light on whether these traits influenced the colonisers’ map making. With the exception of the land mass of the historical ethnic homeland and (weakly) the presence of lakes, there are no major differences in the regions of split and non-split groups across dozens of geographic-ecological aspects (e.g., soil quality, elevation, prevalence of malaria).
There are also no systematic differences across several pre-colonial, ethnic-specific, institutional, cultural, and economic features, such as the size of settlements, the type of subsistence economy, and proxies of pre-colonial conflict.

1.3 Ethnic partitioning and conflict

In Michalopoulos and Papaioannou (2016), we then quantify the impact of ethnic partitioning on civil conflict, using data from the Armed Conflict Location & Event Data Project (ACLED) This dataset reports georeferenced information on battles between government forces, rebels and militias; violence against civilians; foreign interventions; conflict between non-state actors; and riots/protests.

Figure 2a portrays the spatial distribution of violence in Africa over the period 1997-2013 (coverage increases considerably after 2010, but this does not much affect the results). The map plots 64,650 high-precision georeferenced conflicts. There is noteworthy heterogeneity. There are numerous events in Central Africa, mostly in Eastern Congo, Rwanda, Burundi, and Uganda. In Western Africa, conflict and political violence are concentrated in Nigeria and Sierra Leone. In Eastern Africa, violence is pervasive in Somalia. There is also considerable variation within countries. For example, while conflict in Tanzania is low, there are several violent events along the border with Kenya and Rwanda. Most of the conflict in Angola is close to the northern border with DRC and in the oil-rich Cabinda enclave.

Both early works, examining the long-run effects of historical factors on economic development, and the voluminous academic literature on the correlates of civil conflict, have relied on cross-country comparisons (see Blattman and Miguel 2010); yet this approach, while naturally the first step, is prone to various conceptual and econometric shortcomings, as countries differ along many dimensions, the features shaping warfare and underdevelopment are correlated and there are not that many observations. We, thus, conduct the analysis at the historic-ethnic-homeland level, comparing conflict intensity across ethnic regions in the same country. By doing so, we effectively account for national politics, colonial and post-independence country-wide institutions, and other country features that may affect conflict. Moreover, since we examine ethnic regions in the same country, unobserved heterogeneity is less of a concern. Figure 2b portrays the spatial distribution of all conflict events at the country-ethnic homeland level.
Figure 2a.

Type of Violent Events in ACLED

- African Boundaries
- Battle
- Non-violent activity
- Riots/Protests
- Violence against civilians
The econometric analysis reveals that civil conflict is significantly higher in the ethnic regions of partitioned, as compared to non-split, groups. This applies to conflict intensity (number of conflict events), the duration (in years) of conflict, fatalities, and the likelihood of conflict. The significant association between ethnic partitioning and conflict is present, when we restrict estimation to ethnic homelands close to the national borders, so as to account for any border effects (which nonetheless may be driven by partitioning). The estimates suggest that conflict intensity is approximately 40% higher and conflict lasts on average 55% longer in the homelands of partitioned ethnicities. Moreover, the likelihood of conflict is approximately 8% higher in the homelands of split ethnicities.
The econometric analysis reveals sizable spillovers; conflict intensity is approximately 30% higher and the likelihood of conflict increases by 7% in the homelands of groups where 50% of surrounding groups are split (as compared to groups whose neighbours are not partitioned).

We then examine what type of conflict is more likely to occur in partitioned ethnic homelands. First, we document that military interventions from adjacent countries are more common in the homelands of partitioned groups than in nearby border areas where non-split groups reside. Second, we find that partitioning is a key factor for conflict between government troops and rebel groups “whose goal is to counter an established national governing regime by violent acts” and, to a lesser extent, for one-sided violence against civilians. These patterns are corroborated using a different georeferenced conflict database (Uppsala Conflict Data Program Georeferenced Event Dataset) that records only deadly events associated with major civil wars. There is no link between ethnic partitioning and riots/protests, which are predominantly an urban phenomenon; and there is no association between partitioning and conflict between non-state actors.

1.4 Ethnic partitioning and political violence

We then dig deeper regarding the partitioning-repression-civil war nexus, using the Ethnic Power Relations (EPR) dataset (Wimmer et al. 2009), which offers an assessment of ethnic-specific political participation and discrimination since independence. This dataset also gives a classification of ethnic-based civil wars.

The within-country analysis reveals three key findings. First, partitioned ethnicities are significantly more likely (an 11%-14% higher likelihood) to engage in civil wars that have an explicit ethnic dimension. Out of the 234 split groups, 72 (31%) have participated in an ethnic civil war — with the corresponding statistic for non-split groups being 69 out of the 359 (19%) (Figure 3). Second, the likelihood that split ethnicities are subject to political discrimination from the national government is approximately seven percentage points higher when compared to non-split groups. Of split groups, 25% (58 of the 234) have been subject to discrimination by the government, while the respective likelihood for non-split groups is ten percentage points lower (Figure 3).
Third, when we examine the impact of ethnic partitioning, jointly on one-sided violence (repression) and two-sided violence (civil war), we find that, in the weakly institutionalised African countries, ethnic partitioning leads much more often to major two-sided conflict, than to discrimination.

**Figure 3**

![Graph showing comparison between non-partitioned and partitioned ethnicities](image)

We complement the group-based analysis with results linking individual-level access to public goods and education to the location and the ethnicity of roughly 85,000 households across 20 African countries (using the Demographic and Health Surveys, DHS). The econometric analysis reveals that individuals who identify with partitioned groups have fewer household assets, poorer access to utilities, and worse educational outcomes, as compared to individuals from non-split groups in the same country (and even residing in the same town/village). This pattern is not due to a generalised decline in the standards of living for all households residing in split homelands; rather, it is driven by the poorer economic performance of members of split ethnicities, irrespective of their residence.

The evidence from the EPR and DHS shows that the consequences of ethnic partitioning are not circumscribed by the contours of ancestral ethnic homelands, but have significant repercussions for the members of partitioned groups, irrespective of their whereabouts.
2. **Country level characteristics shaped by the Scramble and contemporary development**

Besides ethnic partitioning, scholars have emphasised the negative impact of various country-wide aspects stemming from the Scramble for Africa.

2.1 Access to the coast

The Scramble for Africa and the administrative splits within colonial powers led, upon independence, to the creation of many landlocked countries, with no access to the global shipping lanes. Out of 49 countries, 16 are landlocked, namely Botswana, Lesotho, Swaziland, Zambia, Zimbabwe, Malawi, Rwanda, Burundi, Uganda, Ethiopia, Central African Republic, Chad, Niger, Mali, Burkina Faso and South Sudan. And the largest country (in terms of land area), the Democratic Republic of Congo (the former Congo Free State), only has access to the sea (Atlantic Ocean) via a tiny 30-kilometre strip of land between the mainland of Angola (in the South) and the Cabinda enclave (that is part of Angola, in the North). Close to 300 million Africans (400 million if we add DRC) live in landlocked countries. Collier (2007) argued that being landlocked increases the likelihood that the country will be subject to 'poverty trap' dynamics, since it is associated with various costs\(^7\). First, transportation costs are higher and trade with the rest of the world is limited for landlocked countries (Storeygard 2015). The costs of being landlocked are higher for those African countries that depend heavily on export proceeds from agriculture (e.g., Burundi, Rwanda, and Uganda) or minerals (e.g., Chad, Zambia, Zimbabwe). Second, landlocked countries depend extensively on their neighbours for both exports and imports. Since the neighbours are themselves often unstable, African landlocked countries face serious challenges. For example, Zimbabwe exports to the rest of the world via the Beira corridor and port in Mozambique in the Indian Ocean. During the Mozambican civil war (1977-1992), access to Beira was severely limited or shut down altogether, impeding Zimbabwean exports from reaching world markets.

---

\(^7\) Nunn and Puga (2012) show that distance to the coast correlates negatively with GDP per capita. Within Africa, this correlation is statistically indistinguishable from zero, most likely because landlocked countries were shielded from the ‘slave trades’.
Likewise, almost all of Ugandan exports are channelled via the port of Mombasa in Kenya. The railroad connecting Mombasa to Uganda, the so-called Lunatic Express, has been blocked on many occasions by protests, or has stopped functioning because of conflict, harming Ugandan producers.

2.2 Country shapes and sizes

Herbst (2000) has argued that, by creating states with peculiar sizes and shapes, the Scramble for Africa has had sizable adverse long-run impacts. He also emphasises the ‘inappropriate’ location of capitals. Upon independence, the newly-minted entities, particularly the inefficiently large ones, found it difficult to broadcast power outside their capital centres. African countries have very low levels of state (fiscal and legal) capacity (Besley and Persson 2011). Because of the limited presence of courts and state authorities outside the capitals, many Africans settle disputes (especially over rural land) in tribal courts and seek the advice of traditional leaders and chiefs (Michalopoulos and Papaioannou 2015). Quite often the state cannot monopolise violence. In Michalopoulos and Papaioannou (2014), we provide indirect evidence along these lines; showing that differences in country-wide institutions, across the national border, map into differences in regional development across the border, only when we look at (ethnic) areas close to the respective capitals. Using survey level data, we further show that proxies of state capacity, related to tax collection and policing, rapidly decay as one moves away from the capitals. In line with this, Campante et al. (2016) show that countries with isolated capital cities have lower quality institutions.

2.3 Ethnic composition (fractionalisation, polarisation, and inequality)

The Scramble for Africa has also inadvertently shaped the ethnic composition of the newly independent states. While the link between ethnic fractionalisation (the likelihood that two randomly drawn individuals of a country do not come from the same ethnic group) and development and/or conflict is weak, there is a tight relationship between ethnic polarisation (which takes the maximum value when there are two groups of equal proportions) and conflict (Montalvo and Reynal-Querol 2005, Esteban et al. 2012).
Alesina et al. (2016) show that ethnic-specific differences in economic prosperity and geographic endowments, which are highest in Africa, map into overall underdevelopment (see also Huber and Mayoral 2015). Alesina and Zhuravskaya (2012) show that ethnic segregation, which again tends to be higher in Africa than elsewhere, is linked to low levels of institutional development.

3. Conclusion

Research quantifying the impact of ethnic partitioning, state artificiality, and other features of the Scramble for Africa with formal econometric techniques and theory, is still in its infancy. The preliminary empirical findings of the economics literature bode well, however, with the case-studies and narratives of the African historiography highlighting the multi-faceted consequences of the Scramble for Africa. More research is needed to better understand how those colonial borders and colonial artefacts that survived decolonisation affect the present. And, since border artificiality and ethnic partitioning are not phenomena exclusive to the African continent, subsequent works could focus on other parts of the world, such as the Middle East.

References


**About the authors**

**Stelios Michalopoulos** is an Associate Professor of Economics at Brown University, a Faculty Research Fellow at the NBER and a Research Affiliate of the CEPR. A native of Argos, Greece, Stelios holds a B.A. from the Athens University of Economics and Business and a Ph.D. in Economics from Brown University. After completing his doctorate in 2008, he joined the Department of Economics at Tufts University as an Assistant Professor. In 2010-2011 he was the Deutsche Bank Member at the Institute for Advanced Study at Princeton.
His primary research interests intersect with political economy, growth and development, and the economics of culture. He has published in leading peer-reviewed economic journals including the *American Economic Review*, *Econometrica*, *Journal of Political Economy* and the *Quarterly Journal of Economics*.

**Elias Papaioannou** is Professor of Economics at the London Business School (London, United Kingdom). He is also a research fellow of CEPR and the NBER (National Bureau of Economic Research).

He holds an LL.B. from the law school of the National and Kapodistrian University of Athens, Greece, a Masters in Public Policy and Administration (MPA) with a concentration in international economics from Columbia University, and a Ph.D. in economics from the London Business School. After the completion of his doctorate in 2005 he worked for two years at the Financial Research Division of the European Central Bank (ECB) in Frankfurt, Germany. From 2007 till 2012 he served as Assistant Professor of Economics at Dartmouth College (NH, USA), while during the 2010-2011 and 2011-2012 academic years he was a Visiting Assistant Professor at the Economics Department of Harvard University (MA, USA).

His research interests cover the areas of international finance, political economy, applied econometrics, macro aspects of regulation, law and finance, and growth and development. He has published in many leading peer-refereed journals, such as *Econometrica*, the *Quarterly Journal of Economics*, the *Journal of Political Economy*, the *American Economic Review*, *Journal of Finance*, the *Economic Journal*, the *Review of Economics and Statistics*, the *Journal of Development Economics*, the *Journal of the European Economic Association*, the *Journal of International Economics*, and more. His work has also appeared in numerous edited book volumes.

His research has been recognized with the inaugural 2013 European Investment Bank Young Economist Award, 2005 Young Economist Award by the European Economic Association and the 2008 Austin Robinson memorial prize by the Royal Economic Association. Elias consultants for international organisations, investment banks, hedge funds, and institutional investors on macroeconomic developments in the EU and Greece.
4 Indirect rule and local government reform in Africa today

Tristan Reed

An influential body of research – much of it summarised in the introductory chapters to this eBook – has argued that sub-Saharan Africa’s low growth and uneven development during the second half of the 20th century was driven, in no small part, by the legacy of those colonial institutions that limited the ability of citizens to hold their national leaders accountable.

Those who subscribe to this view have cause for optimism about the future: whereas in 1980 fewer than 5% of Africans lived in democracies, 40% do so today.1 But on a predominantly rural continent, where the reach of the central state can be short, much of politics, to paraphrase the saying, is local. The positive impact of innovations in national institutions may be limited by the extent that local governments remain unaccountable.

In many countries, the lowest level of government remains occupied by traditional leaders, or ‘chiefs’, whose institutions are shaped by laws developed under colonial indirect rule. Their regimes feature variously some traits of autocracy: executives are selected by hereditary succession, participation in elections is limited, and terms may last for life. 50% of sub-Saharan Africans report that these leaders have ‘some’ or ‘a great deal’ of influence over their daily lives: they raise taxes, control the judicial system, and allocate land, the scarcest resource in rural areas.

1 Author’s calculations using World Bank population data and the Polity IV Annual Time Series. If Nigeria’s institutions consolidate further towards democracy, as appears more likely after voters there oversaw, last year, the first ever peaceful transition of power between two political parties, this share could swing to 60%.
Attempts to make local government more accountable must necessarily involve these rulers. Over the past few decades, national governments – Mozambique, Uganda and Sierra Leone are salient examples – have decentralised expenditure decisions to elected local officials, who must campaign and govern alongside chiefs. The World Bank has allocated millions of dollars in support of community driven development programmes, which place greater control over aid expenditure in the hands of local leaders, elected or otherwise. The hope has been that, by delegating authority to local leaders, who may have better information about local needs, government as a whole will become more accountable. But, as described in detail in Bardhan and Mookherjee (2006), theory is ambiguous as to whether this hope will prevail. Decentralisation may also open up greater scope for the capture of public funds by chiefs, especially if they remain unaccountable themselves due to the legacy of colonial law.

Two competing views exist about how chiefs may shape these initiatives. The first, summarised by Logan (2013), is that of ordinary Africans. In responses to the Afrobarometer survey, chiefs enjoy widespread popular support, and the majority of citizens believe that the influence they have should increase. Taken at face value, this finding suggests that chiefs’ influence may be complementary to local government reform. This view, however, is in stark contrast to a second, untested but widely-held, view among historians that the scope for chiefs to abuse their authority – expanded and entrenched by the legal structures of colonial indirect rule – has slowed the pace of development (see, most recently, Mamdani 1996).

**Economic development and elite control of civil society in Sierra Leone**

In work with Daron Acemoglu and James Robinson, I use the colonial organisation of the chieftaincy in Sierra Leone to study the impact of the absence of constraints on chiefs’ power on economic outcomes, citizens’ attitudes, and social capital today (Acemoglu et al. 2014). Our results provide lessons of caution for local government reformers on the continent today.
Indirect rule was a colonial strategy that involved keeping in place existing rulers and ruling through them. Lord Lugard (1922), who elaborated the model during the pacification and control of northern Nigeria, explained in his manual, *The Dual Mandate in British Tropical Africa*, how chiefs, despite their freedom to govern their people as they chose, would derive their legitimacy entirely from the colonial government: “*The chief himself must understand that he has no right to place and power unless he renders his proper services to the state*”. Chiefs were accountable to their occupiers, but the system did not require them to be accountable to their people.

In 1896 British colonial authorities established indirect rule in the newly created Sierra Leone Protectorate by empowering a set of paramount chiefs as the sole authority of local government. Only individuals from the designated ‘ruling families’ of a chieftaincy – the elite, created, and given exclusive right to rule, at the initiation of the system – are eligible to become paramount chiefs. Chiefs remained effectively the only institution of local government until the World Bank sponsored the creation of a system of elected local councils in 2004.2

We argue that the fixed number of ruling families allowed to put forward candidates for the chieftaincy is a useful measure of political competition, and, thus, a chief’s accountability to his or her people. This is consistent with the arguments of historians such as Burrows (1976), who describes an analogy between ruling families and political parties. Using this measure, we show that, for those born in places where there are fewer ruling families, development outcomes are significantly worse. Figure 1 shows the distribution of the numbers of families across Sierra Leone, along with the other key variables in our study.

Clearly, the number of ruling families in a chieftaincy may be correlated with omitted variables that influence current outcomes through other channels. Three observations though should alleviate this challenge to a causal interpretation of our results. First, using field and archival research on the histories of the ruling families we document that their origins are highly heterogeneous and often the result of historical accident, such as the availability of a male heir or the number of leaders in an invading war party.

---

2 An excellent summary of Sierra Leone’s post-war local government reforms and their impact is provided by Casey, Glennerster and Miguel (2012).
Figure 1.

1a

Legend
- District Headquarters
- 1896 Trade Routes
- 1907 Railroad
- Navigable River

Number of ruling families

1
2
3 - 4
5 - 6
7 - 12

Freetown colony

0 25 50 100 Kilometers

1b

1900 tax assessment per 1963 population

0.003 - 0.010
0.010 - 0.015
0.015 - 0.023
0.023 - 0.035
0.035 - 0.188

No data
Second, as can be observed in Figure 1, chieftaincies with many families are not clustered in any particular area of the country, nor around navigable rivers, trade routes or railway lines; all our regression specifications are robust to the inclusion of these and other geographic controls that may be correlated with the economic development potential of a chieftaincy. Third, as we verify quantitatively, and as one may observe by comparing the top two panels in Figure 1, the number of ruling families is uncorrelated with the level of development before the creation of paramount chiefs, as measured by tax assessments, per chieftaincy, produced by the British colonial government in the late 1890s.

Overall, we find a significant positive relationship between the number of ruling families and development outcomes. Quantitatively, the effects are substantial. Moving from the bottom quartile to the top – from 1.8 ruling families to 7.7 – corresponds to an increase in literacy and in primary and secondary school attainment of about seven percentage points, off a low base of 37%. This relationship can be observed visually in Figure 1. We also find that this same change in the number of ruling families increases a child’s height for weight Z-score by nearly a third of a standard deviation, and decreases the likelihood of child anaemia by 13 percentage points. Where institutions of indirect rule created varying lower levels of accountability in local government, development is worse today.

An influential line of argument in political economy maintains that autocratic power among leaders both results from and leads to low social capital and civic participation (for the seminal work, see Putnam, 2000). That is, where we see unaccountable chiefs, we should also see a disengaged society. To our surprise, in our data, we found the opposite: places with fewer ruling families exhibit greater social capital on dimensions such as attendance of community meetings, participation in social groups and the undertaking of collective action. The inverse relationship for meetings can be observed visually in Figure 1, and is estimated to be substantial: moving from the top to bottom quintile of the number of ruling families increases the likelihood of attending such meetings by 12 percentage points. Further, we find that in places with fewer ruling families, individuals have more favourable attitudes towards the paramount chief’s authority, being more likely to agree that “we should have more respect for authority,” and that “only older people are mature enough to be leaders.”
This somewhat puzzling finding, we suggest, arises because more dominant chiefs have been better able to mould civil society and institutions of civic participation, for their own benefit and continued dominance, in a way that appears consistent with the case-study literature on local government reform in Sierra Leone (e.g. Fanthorpe 2001, 2005, and Sawyer, 2008). As a consequence, relatively high measures of civic participation in areas with less-accountable paramount chiefs are not a sign of a vibrant civil society disciplining politicians, but rather of a dysfunctional civil society captured by elites.

Applied across Africa, these results may also explain why chiefs maintain so much popular support: If civil society has been captured, citizens will typically still find it valuable to interact with elites. In places where chiefs are less accountable, citizens are more dependent on the patronage and favours of their chiefs, and thus may find it useful to make specific investments in the system. Having made these investments, they may prefer the system to persist.

**Lessons for the future**

What does this tell us about local government reform today? Local government reform initiatives seek to give local civil society more authority over expenditure, in the hope that doing so will improve public goods provision and accountability.

Our results suggest caution is needed with such initiatives. If civil society is captured by chiefs, efforts to strengthen it without freeing it from the control of traditional elites might just strengthen the power of chiefs and expand their opportunities for capture of public goods. It is an open question whether interventions that strengthen civil society organisations, within a given institutional structure, improve governance, or further bolster existing institutions, even if they are dysfunctional. Local government reform should be monitored closely to ensure it is not captured.

Our work also suggests that attempts to reform local government should be judged by their effects on final and not intermediate outcomes. Indicators of an active civil society and verbal support for the chief’s participation in public life may be consistent with elite capture. The real test of whether local government is truly effective will be in health, educational and economics outcomes, and not in whether people are attending meetings.
References


Bardhan, Pranab and Dilip Mookherjee (2006), Decentralization to Local Governments in Developing Countries: A Comparative Perspective, Cambridge: MIT Press.


Lugard, Frederick (1922), The Dual Mandate in British Tropical Africa, London: William Blackwood and Sons.


About the author

Tristan Reed is research director at Crime Lab New York, an UChicago Urban Lab. His current research interests are the political economy of local government, and the design of urban social policy, both in low and high income countries. Previously he worked as a management consultant for McKinsey and Company in Lagos, Nigeria, serving clients across sub-Saharan Africa, focusing primarily on large-scale transformations of public sector organizations. He holds a BA in economics from the University of California, Los Angeles and a PhD in economics from Harvard University.
King Leopold’s ghost: The legacy of labour coercion in the DRC

Sara Lowes and Eduardo Montero
Harvard University

The Democratic Republic of Congo (DRC), located in the centre of Africa, is one of the world’s poorest countries. Prior to becoming the DRC, Zaire, or even the Belgian Congo, it was the Congo Free State (CFS), the personal colony of King Leopold II of Belgium. Between 1885 and 1908, it was the site of some of the worst human rights abuses in Africa’s colonial history. In fact, historians Hochschild and Vansina estimate that 10 million people, approximately half of the population of Congo, died between 1880 and 1920 (Vansina 2010 and Hochschild 1998). Motivated by a desire to profit from soaring natural rubber prices, Leopold granted private companies concessions to extract rubber in the upper Congo basin. The concession companies, with the support of the Force Publique, and with their own militias, forced people to collect rubber as a form of taxation. Those who did not meet the quotas for rubber collection were subjected to severe punishment and violence.

In an academic working paper, Lowes and Montero (2016) examine the long-term effects of labour coercion on economic development in the DRC, using the well-defined boundaries of the two largest historical rubber concessions from the CFS-era, ABIR and Anversoise. This work relates to several literatures on understanding Africa’s comparative development. First, a growing literature in economics documents that Africa’s colonial history has important implications for its economic development. For example, Nunn (2008) demonstrates that African countries with a greater exposure to the Atlantic slave trade are less developed today. Michalopoulous and Papaioannou (2016) study how the country borders that were arbitrarily drawn under colonialism have led to more conflict and lower development, by comparing partitioned to non-partitioned ethnic groups. Second, during the colonial era labour coercion was widely prevalent.
There is evidence that historical labour coercion affects present day development. Dell (2010) examines the long-term impacts of the mining mita in Peru that lasted over two centuries. The labour coercion in the CFS was extremely violent, but also of relatively short duration compared to the mining mita or the slave trade. Finally, Lowes and Montero (2016) address the literature on the effects of indirect rule within Africa. Mamdani (1996) argues that the strategy of indirect rule, during the colonial era, may have undermined the accountability of local leaders. Lowes and Montero (2016) examine the effects of the co-option of chiefs during the rubber era on the quality and accountability of local leaders.

**History of the CFS rubber concessions**

European colonisers had claimed much of Africa by the mid-1870s, yet central Africa remained largely unexplored. King Leopold II leveraged the mistrust between the British, French and German governments, regarding their colonial aspirations, to negotiate the creation of the Congo Free State, as his own personal colony, in 1885. In return, Leopold promised to end the slave trade and to maintain the Congo as a free-trade zone for individuals of all nationalities (Nzongola-Ntalaja 2007). A condition of his control over the CFS was that he needed to demonstrate state presence within the Congo. This proved to be costly, and by 1890 Leopold had spent almost the entirety of his father’s fortune in the Congo (Van Reybrouck 2014). His financial troubles coincided with the invention of the pneumatic tyre, which greatly increased the demand for natural rubber. Congo had an abundance of natural rubber, offering Leopold an opportunity for profits.

In 1892, because the CFS had limited capacity and manpower, Leopold granted concessions to private companies to extract rubber. The two largest concessions that focused primarily on rubber were ABIR and Anversoise. As most of the interior of DRC was unexplored at the time, the boundaries of these two concessions were determined using salient geographic characteristics - in this case, major rivers and their tributaries. ABIR was situated between the Maringa and Lopori rivers and Anversoise occupied the Mongala river basin. Figure 1 shows an historical map illustrating the concession borders and Figure 2 presents a map of the digitised concession borders and the modern DRC border.
Leopold was a majority stakeholder in both companies and the CFS shared in the profits of the companies (Harms 1975).

**Figure 1.** Historical map of the rubber concessions

*Notes:* The Anversoise rubber concession is the northern concession and the ABIR concession is the southern concession in the figure above. The original map is from (Waltz 1918).

In the sparsely populated concessions, access to labour was a key constraint on the rubber production process. The rubber companies set up systems to ensure control over and access to labour. European agents were sent to the interiors of the concessions to establish posts. They would survey surrounding villages to create a census of adult men and quotas were then set, based on these population counts. Generally, male villagers were required to deliver around 4 kilos of dried rubber to the European agents every two weeks. Additionally, villages were required to provide food and supplies to nearby posts (Harms 1983).
Figure 2. The Democratic Republic of Congo and the rubber concession borders

Notes: The map illustrates the present day boundaries of the Democratic Republic of Congo and the digitised rubber concession boundaries.

To collect rubber, villagers would travel into the jungle to tap *landolphia* vines. The process could take days, and, as natural rubber supplies dwindled, villagers would take almost the entire two weeks to fill the mandated quota. The punishment for not meeting the quota was severe – individuals could be imprisoned and subjected to various forms of physical violence, including burning, whipping by the *chicotte* (a whip made of hippopotamus hide), or death. Figure 3 and Figure 4 provide contemporaneous illustrations of the rubber collection process and the *chicotte* punishment, respectively.
Figure 3. Illustration of the rubber collection process

The soldiers from the *Force Publique* and sentries from the concession companies’ private militias were primarily responsible for carrying out these violent tactics, although European agents also imprisoned and tortured villagers. Sentries armed with rifles were assigned to each post. To prevent waste, the sentries were required to provide a human hand for every bullet used.

Robert Casement, a British consul, was sent to Congo to investigate accusations of atrocities. First hand African accounts document the intensity of the labour coercion:

“When I was still a child, the sentries shot at the people in my village because of the rubber. My father was murdered: they tied him to a tree and shot and killed him, and when the sentries untied him they gave him to their boys, who ate him. My mother and I were taken prisoner. The sentries cut off my mother’s hands while she was still alive. Two days later, they cut off her head.” (Janssens 1904)

The primary responsibility of the sentries was to ensure rubber production. The head sentry for a village, called the kapita (which today is a Lingala word used to denote ‘village chief’), would assert his authority over a village by undermining men in positions of esteem or authority. According to accounts by Morel (1904), the village headmen were “shamelessly degraded in the eyes of their people, made to fetch and carry for soldiers, cast into chains and flung into prison”.

Though the CFS government objected publicly to the violence, in practice it encouraged and benefitted from it (Hochschild 1998). The concession companies made exorbitant profits, particularly as rubber prices increased from 6.20 francs per kilo in 1894 to over 10 francs per kilo in 1898. Meanwhile, the estimated cost incurred by the concession companies to ‘purchase’ a kilo of rubber in CFS was approximately 1.35 francs (Harms 1983). The magnitude of profits earned by the concession companies led one contemporary observer to note that “ABIR has in a single fiscal year made a net profit that represents more than twelve times the initial capital investment. Such a result is perhaps without precedent in the annals of our industrial companies” (Plas and Pourbaix 1899).

The intense extraction efforts by the rubber concessions and the destruction of the delicate landophilia vines meant that the natural rubber supplies were nearly exhausted in the Upper Congo Basin by 1905. This coincided with the release of the Casement report, documenting the human rights abuses in the CFS. Facing dwindling rubber supplies and condemnation of their labour practices, ABIR and Anversoise left the CFS in 1906. In 1908, the CFS became a Belgian colony (Harms 1975).
The regime of labour coercion had disastrous effects for the local population, including famine, spread of sleeping sickness, and low fertility rates. The brutality of the rubber collection tactics earned the policy the nickname ‘Red Rubber’ (Vangroenweghe 1985).

The long-run effects of labour coercion

To examine the long-term effects of the rubber concessions, Lowes and Montero (2016) use Demographic and Health Survey (DHS) data from 2007 and 2014 to test whether the former concessions have worse development outcomes today. The well-defined concession boundaries allow them to use a geographic regression discontinuity design – essentially comparing villages just inside the former concession borders to those just outside the former concession borders – to causally identify the effect of labour coercion on three primary outcomes of interest: years of education, wealth and health. Similar to the borders studied by Michalopoulous and Papaioannou (2016), the concession boundaries were determined at a time when there was little knowledge of the geography of the interior of Congo. Thus, Lowes and Montero (2016) compare areas that were geographically similar and were culturally identical at the time of colonial contact but differed only with respect to whether they became part of the rubber concession.

Lowes and Montero (2016) find that former concession areas have significantly worse education, wealth and health outcomes relative to places just outside the concessions. Individuals inside the concessions have on average 1.5 fewer years of education, are 20% less wealthy, and have lower height-to-age percentiles, which is a measure of cumulative health status. Additionally, they compare older and younger age cohorts within the DHS and find no evidence that outcomes are improving over time. These results suggest that a 14-year event, occurring over 100 years ago, still matters for understanding the relative development of these former concession areas.

This naturally leads to the question of why this historical event still affects present day development. One potential explanation is that the subsequent Belgian government invested less in infrastructure inside the former concessions. Alternatively, since missionaries were the primary providers of schooling during the colonial era, if there were fewer Catholic and Protestant missions inside the former concessions, then perhaps this can explain the observed results.
However, using maps on colonial investments in health centres, roads, and mission stations, Lowes and Montero (2016) find no evidence that there was differential investment in the region by the Belgian colonial administration nor that there was differential missionary presence.

Previous work has highlighted the importance of institutions for explaining differences in development (Acemoglu et al. 2001). Using survey data collected near the border of the Anversoise concession, Lowes and Montero (2016) test whether villages within the former concessions have lower quality local institutions. The historical co-option of village chiefs to support the rubber regime may have undermined long-term local institutional quality. They measure institutional quality by the extent to which the chief provides various public goods for the village and how chiefs are chosen. They find that village chiefs within the former concessions are less likely to provide critical public goods, such as road maintenance and conflict arbitration. Additionally, they are less likely to be elected to their position and are more likely to be hereditary. These results provide important evidence that even a relatively short-term exposure to violence and extractive colonial institutions can affect present day economic development, if it undermines institutional quality.

These findings are related to the work by Mamdani (1996), who argues that indirect colonial rule created unaccountable chiefs in Africa, and work by Acemoglu et al. (2014) on how the denomination of different numbers of ruling families by the colonial government has affected development in Sierra Leone. Acemoglu et al. (2014) find that districts with fewer ruling families, and therefore less political competition, experience worse development outcomes but higher levels of social capital. Interestingly, Lowes and Montero (2016) also find evidence that individuals from former concessions trust their local chiefs more than those individuals outside the former concessions, even though the chiefs inside the former concessions provide fewer public goods today.
Conclusion

Lowes and Montero (2016) examine an important and brutal episode of colonial labour coercion. Despite the rubber regime having lasted only 14 years, Lowes and Montero (2016) demonstrate that the labour coercion has undermined long-run development in the DRC. Individuals in former rubber concession areas have lower levels of education, are poorer, and have worse health outcomes than individuals just outside the concessions. Using data collected along the boundary of one of the former concessions, Lowes and Montero (2016) provide evidence that the rubber concessions have affected local institutional quality. Studying the legacy of this historical event – the Congo Free State rubber concessions and their aftermath – highlights the role of history for understanding the process of economic development. Even historical events of short duration can have long-lasting effects on development and local institutional quality.

References


About the authors

Sara Lowes is a Ph.D. candidate in the Political Economy and Government program at Harvard University. She graduated from Middlebury College in 2007 with a B.A. in Economics and Political Science. She studies economic development, political economy and economic history. Her most recent work is in the Democratic Republic of Congo. She also has research experience in Uganda, Zambia, Ghana, Nigeria, Bolivia and Vietnam.

Eduardo Montero is a Ph.D. candidate in the Economics Department at Harvard University. He is from San Jose, Costa Rica. He graduated from Stanford University in 2010 with a B.A. in Economics, and in 2011 with an M.S. in Statistics. Eduardo’s interests are in development economics, political economy, and economic history. His research focuses on Central Africa (in particular, the Democratic Republic of Congo) and Central America.
6 How colonial railroads defined Africa’s economic geography

Remi Jedwab, Edward Kerby and Alexander Moradi
George Washington University; Stellenbosch University; Sussex University

Infrastructure and development in poor and remote regions

Cities are the main engines of growth (Lucas 1988). By agglomerating people, cities facilitate the exchange of goods and ideas, increasing aggregate productivity (Glaeser and Gottlieb 2009), and potentially promoting growth in developing countries (Duranton 2015). The study of why cities emerge, and persist, has become crucial to our understanding of the growth process, and of the evolution of spatial inequality both across and within countries.

At the turn of the 19th century, sub-Saharan Africa was the least urbanised region in the world, with only about 50 cities of more than 10,000 inhabitants. By 2010, the number of cities had increased to almost 3,000. Our research explores whether colonial railroad investments transformed Africa’s economic geography. Were railroads so instrumental that without them economic outcomes would have been different and development delayed?

It is commonly acknowledged that transportation infrastructure is needed in order for countries to develop. However, it is less clear whether transportation infrastructure has the influence to determine where and what cities develop, and whether they permanently change the economic landscape of a country. This comes down to a ‘the chicken or the egg’ problem. At the root of city development could be physical geography: the fact that there is a river, a fertile plain, a natural harbour, etc. nearby. Thus, geography could give locations a natural advantage and railroads and roads would therefore be built as a response to this. If geography is the main driver, then there is little scope for history (i.e., man-made advantages) to determine the spatial distribution of economic activity (i.e., cities).
Davis and Weinstein (2002), Bosker et al. (2007) and Miguel and Roland (2011) studied the bombing of Japan, Germany and Vietnam respectively, and showed that many of the destroyed cities subsequently recovered their population and their importance within the national economy. Davis and Weinstein (2002) attributed this to the effect of their geography (which was not altered). However, there is also a literature which finds that localised historical shocks, or ‘accidents’, can have permanent spatial effects. For example, Redding et al. (2011) showed how the division of Germany, despite later reunification, permanently changed the location of its main airports. Likewise, Bleakley and Lin (2012) showed that many US cities originally formed at such obstacles to water navigation as attracted commerce and supporting services. These cities still exist today, in spite of the fact that the importance of water navigation has become marginal over time.

The African context is different: when colonial railroads were built, African countries were poor and non-urbanised. The overwhelming majority of people lived through agriculture - which has less scope for economies of scale. In Jedwab and Moradi (2016) and Jedwab et al. (2016) we studied railroads and development in Ghana, plus Africa and Kenya, respectively.

**Colonial railroads in Africa – historical accidents?**

Why did colonisers build railroads? Through studying the history of each railroad we identified three motivations:

i. Military domination: the line was built to exercise ‘effective control’ in the Scramble for Africa, or to dispatch troops for better control of the native population;

ii. Mining: the line was built into the interior to further European mining interests, and;

iii. Cash crops: the colonial power built the line to connect agriculturally rich areas.

With limited budgets, colonisers expected the railroads to pay for themselves, hence they connected areas of high economic potential.

So what room is there for historical accidents? Plenty.
While the origin and the endpoint of a line had some advantages, none of the localities in-between may have been considered worthwhile to connect. Railroad placement was essentially an engineering problem – figuring out the cheapest (typically the shortest) route from a point A to a point B. The Kenya-Uganda railroad is a case in point (see Figure 1).

Figure 1. Southern Kenya: The railroad followed the least cost path from Mombasa to Kisumu

This railroad was built between 1896 and 1901 to connect landlocked, and relatively prosperous Uganda to the coast (i.e., Mombasa). Kenya was merely transit territory. We verified that this railroad indeed followed the route that minimised construction costs. Using engineering reports, we fed into the computer geographic information and historical cost parameters associated with earth formation, clearing, permanent way, and sidings and culverts. We also noted that the best locomotives at that time could only overcome gradients below 5-10%. The overlap of the cheapest-to-construct route, as identified by the computer, and the actual line is striking. Tellingly, contemporary critics called the Kenya-Uganda railroad the ‘Lunatic Express’, ‘going from nowhere
to utterly nowhere’. In fact, the railroad bypassed highly populated areas *en route* to Kisumu (Lake Victoria) and Uganda. Hence, locations on or near the route were lucky to have received access to the railroad, but others, with similar or better geography missed out.

In Ghana, five railroad routes were proposed, but, due to random events, never materialised (see figure 2). First, the Cape Coast–Kumasi line (1873) was proposed, to link the capital of Cape Coast to Kumasi, sending troops to fight the Ashante. This project was dropped because the war came to an abrupt end in 1874. Second, Governor Griffith wanted a central line from Saltpond to Kumasi (1893), in order to tap the palm oil areas and link the coast to Kumasi. When he retired in 1895, he was replaced by Governor Maxwell, who gave in to the gold-mining lobbies and instead built the Western Line, which connected European interests to the coast. For the third and fourth proposed routes, Maxwell thought that a second line was needed, and projects were suggested with two different sets of termini: Apam–Kumasi and Accra–Kumasi. A conference was to be held in London to discuss the proposals, but Maxwell died before reaching London. Following him, Governor Hodgson favoured Accra, but thought that the line should instead be built to Kpong. He retired in 1904, before work began, and was replaced by Governor Rodger, who finally built the Eastern Line.

One of our strategies was, therefore, to compare the locations actually connected to locations hypothetically connected by those alternative routes that were equally likely to have been built but – by the toss of a coin – were not.
Economic development during the colonial era

In Kenya, the railroad placement led to a curious situation, whereby lines traversed sparsely settled areas with no Kenyan freight to transport. In the hope of creating an agricultural export industry, land was alienated and offered to European settlers, with plots near the railroad the most sought after. In our analysis we found that between 1901 and 1962 – a year before Kenya became independent – the railroad led to a concentration of both European and African populations (see Figure 3). The cultivation of coffee and tea expanded along the railroad lines. The effects decrease when moving away from the railroad, and are zero after 30 km. In addition, we found a strong effect of the railroad, within a distance band of 10km, on skilled European, Asian (mostly traders) and urban populations, whereby the former two were concentrated in the latter.
For Ghana, the story is similar. Railroads decreased transportation costs, thereby making cocoa production for export markets profitable. Ghana became the world’s largest exporter of cocoa by 1911. The rural population increased along the railway lines because cocoa cultivation required more labour, thus creating villages. In turn, cities also emerged, as trading stations along the railroad, servicing the villages (see Figure 4).

Lastly, we expand these findings to the rest of Africa circa 1960, when most countries became independent. Using data for 39 sub-Saharan African countries, Figure 5 shows that railroads built during the colonial period strongly predicted the current location of cities.
Figure 4. Railroads, city locations and urban growth in Ghana, 1931-2000

Source: Jedwab and Moradi (2016)

Figure 5. Railroads, city locations and urban growth in sub-Saharan Africa, 1960-2000

Source: Jedwab and Moradi (2016)
Development in the long run: Stability despite the demise of railroads

Railroads gave an initial advantage to the cities they created. However, this advantage disappeared over time. First, railroad systems collapsed after independence, as a result of mismanagement, lack of maintenance, and the adoption of a new transportation technology, namely motor roads. The new roads were not necessarily built along the colonial lines, which should have led to the decentralisation of economic activity, away from the railroads. Second, in the case of Kenya, the European population and a large part of the Asian population left the country after independence. In Ghana, the ageing crop of cocoa trees along the colonial railroad declined in productivity.

Despite all this, we find that for Kenya (2009, see Figure 3), Ghana (2000, see Figure 4) and most of sub-Saharan Africa (2000, see Figure 5): (i) locations along the old railroad lines are still more developed and urbanised today, and (ii) railroad cities, i.e. cities that emerged with the construction of the railroads, are still wealthier today than non-railroad cities of similar sizes.

Sources of path dependence

What is the reason for the strong persistence in spatial economic and urban patterns?

For Kenya we contrasted four different explanations based on: (i) institutional persistence, (ii) technological change, (iii) sunk investments, and (iv) spatial coordination failures.

First, we find that institutional persistence, for example the fact that the new African political elite may have purposely tried to preserve the colonial distribution of population in order to better control the resources of the country, does not account for path dependence.

Second, we verified that path dependence is not explained by changes in transportation infrastructure, as measured by the roads that may have replaced deteriorating railroads.

Third, we find colonial railroad cities were better endowed with non-transportation infrastructure (e.g., hospitals and schools) at independence, and that these colonial investments partially explain path dependence.
Fourth, we argue that persistence is also explained by the fact that the early emergence of the railroad cities served as a mechanism to ‘coordinate’ locational decisions and spatial investments in subsequent periods. For example, cities, once they have emerged and reached a certain size, indicate to potential rural-to-urban migrants the locations to which they can move, if they want to be better able to exchange goods and ideas with other individuals. We, therefore, contend that expectations about where people will live in the future may matter as much as history (i.e., colonial investments), which suggests that regional policies may be ineffective if they do not also change expectations.

For Ghana we came to the same result. Sunk investments during the colonial period and the resolution of these spatial coordination failures both explain path dependence.

**Lessons for infrastructure investments in Africa today**

Africa still suffers a massive infrastructure deficit (Calderón & Servén 2010). During the last decade, many countries have invested in modernising their infrastructure networks. Various giant transportation projects are on their way, such as the Tanzania–Gabon railway ($33 billion), the Mombasa–Kampala–Kigali railway ($14 billion), the Trans-Kalahari railway ($9 billion), and the Abidjan–Lagos motorway ($8 billion), and there are many more ambitious plans, such as the proposed rail/road corridor connecting South Sudan to the Indian Ocean. Chinese firms are playing a leading role, with striking similarities to the European colonial endeavours at the turn of the 20th century, seeking to unlock the continent’s resource potential.

In the light of our findings, what lessons does history provide to inform policy? Firstly, infrastructure investments can produce economic change by reducing trade costs and integrating markets. Second, the impact will, however, depend on how much these transportation investments decrease existing trade costs. The construction of the railroads in colonial Africa constituted a ‘transportation revolution’, because trade costs were extremely high before. However, the construction of roads in post-colonial Africa did not constitute a transportation revolution, since railroads had already significantly decreased trade costs and cities had already been created and grown to significant sizes. This implies that transportation investments will have smaller effects in more developed regions.
Conversely, new infrastructure could transform the economic landscape in poor, remote regions with high trade costs. Finally, the effects of new investments in transportation infrastructure in poor, remote regions will likely depend on a region’s intrinsic economic potential. This is because, at the beginning of the development process, city growth is fuelled by the agricultural activities and rents surrounding them. In both Ghana and Kenya, we found positive effects of transportation investments on agricultural development and structural change.

References


About the authors

Rémi Jedwab is an Associate Professor of Economics and International Affairs at the Elliott School and the Department of Economics of George Washington University. He received his Ph.D. in Economics from the Paris School of Economics. His main field of research is urban economics, though his work also has strong development economics, political economy/public economics, economic history, trade and environmental economics themes.

Edward Kerby is a Postdoctoral Research Fellow at Stellenbosch University in the Department of Economics and at LEAP (Laboratory for the Economics of Africa’s Past). He researches African Economic History, with a special interest in economic geography, development economics and business history. His PhD (LSE) focused on trade, investment and migration in apartheid South Africa. Current research projects include the rise and decline of manufacturing in Tanzania, colonial railways and disease burden in Malawi, and quantitative business history of the Cape Colony.

Alexander Moradi is a Senior Lecturer at the University of Sussex and a Research Associate at the Centre for the Study of African Economies (CSAE) at the University of Oxford and Stellenbosch University. His research primarily focuses on topics related to the economic development of Africa over the long run.
7 The devil is in the detail: Christian missions’ heterogeneous effects on development in sub-Saharan Africa

Julia Cagé and Valeria Rueda
Sciences Po Paris and CEPR; Pembroke College, Oxford University

Following Acemoglu, Johnson, and Robinson’s seminal contribution on the long-term determinants of comparative development (Acemoglu et al. 2001), there has been extensive applied work in economics, aimed at quantifying the extent to which colonisation, and the institutions established by colonisers, has affected the development of African countries in the long-run. This book itself illustrates the diversity of ways in which colonial institutions persistently shaped African development: in their legal design (Vol.1 Chapter 8), through extractive institutions like slavery and forced labour, (Vol. 2 Chapters 1 and 4, and Vol. 3 Chapter 2), the consequences of the arbitrary drawing of political borders (Vol. 2 Chapter 2), and colonial investments (Vol.1 Chapters 7 and 10, and Vol.3 Chapter 1). In this chapter, we study the effects of another form of European settlement that expanded throughout African countries, from the middle of the 19th century: Christian missionary activity.

In general, it was not the aim of Christian missionaries to establish any form of political rule in foreign lands, although they did have a central role in shaping some institutions, in particular through the abolition of slavery. Their primary objective can be interpreted as a form of spiritual imperialism, achieved by bringing Christianity into all the newly colonised territories. The conversion mission was relatively effective; today, in countries such as Angola,
DRC, Kenya, Malawi, South Africa, Zambia, or Zimbabwe, Christians represent 80% or more of the population. However, throughout history, religious change is known to have brought about significant economic change in many countries – conversion to Christianity in Africa was no exception.

The long-term consequences of religious incentives on economic success have been widely studied in the social sciences, the most well-known theory being Max Weber’s ‘Protestant Ethic’ (Weber 1930). According to Weber, the “intrinsic character of the religious [Protestant] belief” is a possible, although not unique, explanation for the development of ‘economic rationalism’ (Weber 1958, p. 49). Weber’s simplified argument is that, for the more ascetic branches of Protestantism, devotion to one's work is a form of religious piety because, in the dogma, salvation is predetermined by God. Continuous hard work and asceticism, as a sign of having a calling, were thus interpreted as signs of salvation. In Weber’s work, the more ascetic branches of Protestantism incentivised continuous hard work, which in turn explains the relatively higher economic performance of Protestant regions compared to Catholic ones. Chapter 7 of Volume 3 also presents the modern economic research on the economic change brought by the Protestant Reformation in Europe.

Another fascinating example of the long-lasting economic effects of religion is the case of Judaism. In their book *The Chosen Few*, Botticini and Eckstein claim that the over-representation of the Jewish population in high-skilled and high-earning jobs results from the reforms of Judaism after the destruction of the Second Temple in 70 CE and the beginning of the Talmud era 150 years later. These reforms centred the practice of religion on the study of the Torah at a young age and on the codification of Jewish law in the written Talmud. According to the authors, this new practice of Judaism is in fact the central explanation for the Jewish population shifting out of agriculture towards more skilled and urban professions (Botticini and Eckstein 2012).

Religions affect individuals’ decisions because they establish a set of values that have an impact on individual decision making. In many cases, religions also impose constraints on economic decisions, such as the possibility of accumulating assets or education. In this chapter, we will show that Christianity in sub-Saharan Africa had long-lasting consequences on development. In particular, we show that the diversity of investments brought by Christian missionaries to the region had different, and sometimes conflicting, effects on long-term development.
Protestantism and development: Human capital rather than the Protestant ethic?

Religious beliefs can affect economic outcomes through what they actually are — a set of beliefs and values; however, religions also affect individual choices such as education and investment in public goods, because they have become institutions with a certain form of central power. Recent work in economics has studied these types of channels to understand the role of religion on growth and development, as opposed to focusing on values per se. This strand of research questions Weber’s thesis that work ethic alone explains fully the better economic performance of Protestant regions in Europe. In the same vein as Botticini and Eckstein’s work on Judaism, these studies claim that Protestantism increased incentives for education.

In the Protestant doctrine, every Protestant should be able to read the Bible, a principle also known as Sola Scriptura. The Sola Scriptura, therefore, led to an increase in human capital in regions that adopted Protestantism, which in turn fostered economic growth (Becker and Woessmann 2009, Bai and Kung 2014, McCleary and Pesina 2013, Woodberry 2012, Cantoni 2015). In other words, human capital rather than work ethic would explain economic development (see in particular Vol.3 Chapter 7, Protestant Reformation’s Impact in Europe, of this eBook series). In our work, we focus on a different aspect of conversion to Christianity: we study the lasting effects of missionary investments – in particular the printing press – on development.

The long-term effects of the printing press: evidence from the European Renaissance

The printing press was a major technological change – a media revolution – since it drastically decreased the costs of diffusing and replicating information. Dittmar has shown extensively that the introduction of the printing press was a crucial catalyst for economic growth during the Renaissance in Europe. He shows that those European cities where printing presses were established between 1450 and 1500 grow 60% faster between 1500 and 1600 than, otherwise similar, cities which were not early adopters. Why? According to Dittmar, cities that first adopted print media benefited from localised spill-overs in human capital accumulation and technological change; in other words, the print media eased the diffusion of skills that were valuable to merchants.
One example of this is the emergence of textbooks on ‘commercial arithmetic’, to teach merchants the basic arithmetic skills necessary for their businesses (Dittmar 2011).

The nature of industrial organisation and competition also influenced religion, demonstrating that, as with any other institution, religion is also affected by technological change (Dittmar and Seabold 2015). They show that the ideas central to the Protestant Reformation were more likely to be adopted in cities where there was more competition in the printing market. These cities were also more likely to experience higher levels of growth (Dittmar 2015).

**Protestant missions and the printing press in sub-Saharan Africa**

While the aforementioned studies of the effects of religion on development consider Protestant conversion as a whole, our approach has focused on exploring the heterogeneity of missionary activity in sub-Saharan Africa. More precisely, we estimate the long-term effects of specific investments on development. Dittmar’s research provides a detailed overview of how the printing market had long-lasting effects on growth in Europe. Our first strand of research investigates the role of the missionary printing press in sub-Saharan Africa, using a new geocoded dataset and several previously unexploited historical archives (Cagé & Rueda 2016).

In 19th century sub-Saharan Africa, missions invested in numerous activities, amongst which were education, health care and printing. Protestant missionaries pioneered the development of a written tradition for sub-Saharan African languages. Wherever they went, Protestants quickly formalised indigenous languages and printed Bibles and educational material in these languages. They facilitated access to the printing press, acting as intermediaries for its diffusion. Therefore, most of the first indigenous newspapers were printed and sponsored at mission centres.
The first newspaper intended for black readers, the *Umshumayeli Wendaba* (‘Publishers of the News’), written in Xhosa, was published as an irregular quarterly in 1837 and printed at the Wesleyan Missionary Society in Cape Colony. *Isigidimi samaXhosa* (‘The Xhosa Messenger’), the first African newspaper edited by Africans, was first released in 1876 and printed at the Lovedale Mission Press in South Africa. In 1884, the English/Xhosa weekly *Imvo Zabantsundu* (‘The African Opinion’), the first black-owned and controlled newspaper in South Africa, was published. On the contrary, in regions where Protestant missions were less active, the first newspapers appeared only at the beginning of the 20th century, and no indigenous newspapers were created before WWI. The first paper in Ivory Coast to be owned and edited by an African, the *Éclaireur de la Cote d’Ivoire*, only appeared in 1935.
We investigate empirically the effects of such an historical lag in the creation of the first indigenous newspapers, and more specifically of proximity to a printing press on newspaper readership today. In order to do so, we geocoded the maps and the databases of the Geography and Atlas of Protestant Missions, an atlas published in 1903 that located all the missions in the world and inventoried all their investments (Beach 1903). From maps, such as the one shown in Figure 1, we construct a geo-referenced database – for example, the one shown in Figure 2, which gives the exact location of missionary investments – and we merge that with additional geographic information (climate, elevation, etc.) and modern economic data. Our sample of sub-Saharan African missions includes a total of 723 Protestant missions, of which 27 were equipped with a printing press in 1903.

**Figure 2.** The Geocoded Atlas of Protestant Missions
Using contemporary individual-level data from the Afrobarometer, we find that proximity to the closest location of a mission with a printing press has a robust, positive, and statistically significant impact on civic attitudes and social capital, such as media consumption (newspaper readership) and trust (consistent with the hypothesis formed by historians of sub-Saharan Africa, e.g. Omu 1978 and Tudesq 1995).

**Figure 3.** Distribution of trust and newspaper readership
A one-standard deviation increase in proximity to a mission with a printing press increases the probability of reading the news on a monthly basis by around 5.3% of a standard deviation, depending on the specifications. In contrast, proximity to a mission without a printing press – whether Protestant or Catholic – has no significant impact on newspaper readership (Cagé and Rueda 2016). Figure 3 shows the distributions of newspaper readership and trust in regions close to (below the median) and far from (above the median) the printing press. Regions close to the printing press tend to exhibit a lower density of towns with low trust and newspaper readership. Other investments, in turn, do not tend to affect social capital and media consumption in a systematic way.

**Christian missions were different, and thus had different effects on development**

By looking at the detail of missionary activity, our investigation sheds a new light on previous literature about the role that missions play in development. And the devil lies in the detail: missionary activity was extensively heterogeneous and analysing the investments conducted by missions is a way to capture this heterogeneity. Depending on the type of investments conducted, missionary activity had different effects on different dimensions of development. Another way to capture heterogeneity is to look at differences between missionary organisations, which is Valencia-Caicedo’s approach: he compares human capital accumulation around Jesuit and Franciscan missions among the Guaraní populations in South America (Valencia-Caicedo 2015).

Our more recent work (in progress) provides further evidence on the intricate development effects of the Christian influence in Africa (Cagé & Rueda 2016b). We study the role played by early missionary investments in sub-Saharan Africa to explain HIV/AIDS prevalence nowadays.

On the one hand, missionaries were the first to invest in medicine in a number of countries. The history of modern medicine in sub-Saharan Africa is indeed closely linked to the development of missionary activity. According to the *World Missionary Atlas* (Beach and Fahs 1925), there were 150 missionary physicians in Africa in 1925, and more than 235 nurses working with nearly 500 trained native nurses in 116 hospitals and 366 dispensaries.
Moreover, the early Christian provision of health care persisted after colonisation and is particularly influential in the design of health care in poor countries (Idler 2014).

On the other hand, health investments are not the only way through which missionary activity may have affected the propagation of HIV/AIDS. Christian values also affect sexually transmitted diseases and there is quantitative and historical evidence that missionaries actively changed sexual behaviours (Vaughan 2007, Doyle 2013, Mantovanelli 2014).

**Figure 4. HIV prevalence and distance to any mission**

![HIV prevalence and distance to any mission](image)

We show that the net effect of proximity to historical mission settlements on HIV prevalence is negative. As seen in Figure 4, regions far from missions tend to have less HIV prevalence today. This general correlation cannot be fully captured by the fact that missions tend to be located in what are today more densely populated areas, or by any other geographical determinant of missionary activity. However, this negative effect can be captured by negative perceptions of condom use and contraception. Among regions historically close to missionary settlements, proximity to a health investment is associated with lower prevalence rates, more acceptance of contraception, and lower exposure to risky behaviours, such as buying the services of sex workers.
More research is needed to better understand the long-run effects of the colonial period, and in particular the persistent impact of missionary activity. Our work shows that, while Protestant missions may have affected economic growth in Europe through literacy more than through values \textit{per se}, values may also matter for development. The devil is in the detail of whichever aspect of religious activity we focus on, like the type of investments conducted or values preached. Moreover, the type of outcome considered does not deliver the same conclusions regarding development. Missionary activity is in general associated with better social capital and education outcomes, but the same cannot always be said for health outcomes, in particular when considering STDs.

\textbf{References}


About the authors

Julia Cagé is an Assistant Professor of Economics in the Department of Economics at Sciences Po Paris, and a Research Affiliate at CEPR in Public Economics and Industrial Organisation. She completed her PhD at Harvard University in 2014. Her research interests focus on political economy, economic history, industrial organization and development. She is particularly interested in the media, especially the questions of how media competition affects the provision of information and political attitudes, and of the production of information in an online world. She is a Board member of the Agence France Presse (third largest international news agency in the world) and a Member of the Commission Economique de la Nation (non partisan Council of Economic Advisors to the French Finance Minister). Her work has been published in the Journal of International Economics, the American Economic Journal: Applied Economics, the European Economic Review, the Journal of Globalization and Development, as well as in several handbook chapters. She is the author of Saving the Media. Capitalism, Crowdfunding and Democracy, Paris, Le Seuil, 2015, translated in ten different languages (English translation: Harvard University Press, 2016).

Valeria Rueda is the Rokos Career Development Fellow at Pembroke College, Oxford University. She obtained her PhD from Sciences Po Paris. Her research focuses on understanding the long-term determinants of health, social capital, and political development outcomes using a quantitative historical perspective. In particular, she is interested on studying the roles of religion and culture as persistent determinants of development. She was born and raised in Colombia before moving to France and the UK to obtain her higher education degrees.
Despite cross-country similarities in economic development, Africa is the region of the world with the greatest variation in political regimes. While a number of its countries have experienced major democratic reforms following the end of the Cold War (e.g., Benin, Ghana, and South Africa), others have either remained autocratic or become unstable democracies, plagued with political violence (e.g., Cameroon, Congo, and Zimbabwe). What explains this cross-country variation in democratic development? We argue that the drastic divergence in democratic trajectories between these two sets of countries is largely explained by the crucial choices they made on their road to independence – in particular, the differing ways they chose to resist and fight colonial rule.

Our explanation challenges the conventional wisdom that political development in the form of democratic change is a byproduct of economic prosperity (Barro 1999, Lipset 1959, Przeworski and Limongi 1997). Following other influential work that underscores the lasting impact of choices, made during critical junctures, on political institutions (e.g., Acemoglu et al. 2008), we show that current levels of democracy in Africa are linked to the nature of its independence movements. Countries that experienced rural insurgencies tend to have autocratic regimes, while those that faced urban protests tend to have more democratic institutions.
Brief historical background

The decade following the end of World War II is perceived as a foundational moment for African political development (Cooper 2002, Mamdani 1996). Resistance movements against colonial rule evolved into Pan-African social movements, which were well integrated into the international socialist and labour movement and, as such, reflected its internal ideological divide. One wing was composed of Western European-style socialists (e.g., Kwame Nkrumah in Ghana, and Julius Nyerere in Tanzania). The other wing was composed of the more radical Maoist leaders (e.g., Frantz Fanon in Algeria, Dedan Kimathi in Kenya, and Ruben Um Nyobé in Cameroon). These two sets of leaders advocated radically different paths towards independence. While Nkrumah and Nyerere advocated mass mobilisation and nonviolent strategies, Fanon, Kimathi, and Um Nyobé encouraged violent rebellion. For example, in a May 1958 address to his party, Nyerere stressed the importance of a non-violent opposition to the colonial administration:

We shall wage a relentlessly determined battle against [colonialism] until we are free. We shall use no violence. We shall stoop to no dishonest methods. We shall be as clean in our methods as we are in our aims. We shall publicly declare our methods as we publicly declare our aims (Nyerere 1967, pp. 59–60).

In contrast, Fanon colourfully advocated the use of violence as a necessary strategy of emancipation. He wrote:

[At the national level] insurgents’ violence unifies the people [...] At the level of individuals, [it] is a cleansing force. It frees the native from his inferiority complex and from his despair and inaction; it makes him fearless and restores his self-respect (Fanon 1961, p. 94)

Geography dictated, to a large extent, the choice between rural insurgency and urban protest, with enormous consequences for post-independence political institutions. Consider the case of Guinea Bissau and Cape Verde. Despite the Maoist ideological leaning of the African Party for the Independence of Guinea and Cape Verde (PAIGC), the leaders of the movement chose the urban protest strategy in the flat terrain of Cape Verde.
By contrast, armed resistance occurred in the dense jungle of Guinea-Bissau. As of today, Cape Verde is considered an example of democracy in the region. Guinea-Bissau remains undemocratic.

These historical accounts illustrate very well the theoretical foundations of our study. Mass protests enabled participants to develop norms of peaceful political expression and compromise, which facilitate the emergence of democracy. In contrast, armed rebellions generated a culture of political exclusion that tends to perpetuate the use of violence as a form of political expression and conflict resolution.

**Our theoretical argument in a nutshell**

We argue that the divergence of political development paths in Africa is linked to the type of independence movement experienced by each country. The type of independence movement significantly shaped both current institutions and norms of behaviour. On the one hand, mass protests enabled participants to develop norms of peaceful political expression and compromise, which facilitate the emergence of democracy. On the other hand, armed rebellions generated a culture of political exclusion that tends to perpetuate the use of violence as a form of political expression and conflict resolution.

The logic of the argument is as follows. The organisation of an opposition movement – whether in the form of a mass protest or an armed rebellion – is a risky collective action. Individuals or groups decide to participate in the movement, depending on their assessments of the likelihood that others will join them. In other words, independence movements can be regarded as coordination games between groups and individuals. There are at least two possible mechanisms to solve the coordination problem. One is through the use of violent actions to manipulate citizens’ beliefs about the unpopularity of the government and to induce participation in the rebellion (Bueno de Mesquita 2010). The other is through organising a peaceful revolutionary mass movement. The type of public signal chosen – ‘armed rebellion’ or ‘mass protest’– may depend on geography, demography, and economic factors. For instance, members of the independence movement may choose the use of violence or armed rebellion if the country is covered by mountainous terrain (Fearon and Laitin 2003). In contrast, they may be more likely to choose mass protests or peaceful demonstrations, if they are located in urban settings with flat terrain.
Now the question is: why do these two types of movements leave opposing institutional legacies? Mass protests are social movements with a relatively horizontal organisation structure. By their very nature, they enable participants to learn the values of peaceful political participation and expression as well as those of political compromise and openness. As such, mass protests and demonstrations provide the cultural and perhaps the institutional basis for civil liberties and liberal democracy (Dahl 1971). In contrast, rural armed rebellions are violent social movements with relatively vertical, hierarchical organisational structures. If they fail, they are likely to generate self-censorship, such as that in Madagascar following the 1947 rebellion (García-Ponce and Wantchekon 2016). If they succeed, they can result in tighter state censorship and limits to free expression – e.g., restricting access to the state only to those who ‘fought in the bush’. As a result, armed rebellion may facilitate the emergence of autocratic regimes

**Empirical evidence**

To empirically assess the relationship between the type of independence movement and democracy in Africa, we conducted an in-depth review of historical events, to code each country as either having a legacy of rural rebellion or of urban protest. While these two forms of struggle are not necessarily mutually exclusive, we found that all African independence movements were characterised by the adoption of strategies and tactics of political dissent that were either mostly rural (armed rebellion) or mostly urban (mass protest). Figure 1 shows the dominant type of movement experienced by each country.

Figure 2 displays the relationship between the type of independence movement and democracy levels over time, as measured by the Polity IV and Freedom House scores. The data indicate that countries exposed to a legacy of rural insurgency tend to be less democratic than their counterparts. This trend seems to run parallel to the so-called ‘third wave’ of democratisation and is very clear after 1990, that is, after the end of the Cold War. Note, however, that some interesting patterns can be identified before the 1990s.
Figure 1. Types of independence movements in Africa

Notes: This figure shows countries where independence movements relied heavily on rural insurgency strategies (purple) versus countries that relied mostly on urban protests (light blue).

Figure 2. Democracy levels by type of independence movement

Notes: This figure shows annual changes in the average level of democracy (on a 0–1 scale) in rural insurgency versus urban protest countries, based on data from Polity IV (left) and Freedom House (right).
The relationship between the type of independence movement and level of democracy is shown in regression form in Figure 3. For each decade, we estimate a cross-country OLS regression of the average level of democracy on a rural insurgency dichotomous indicator. The point estimates show that the effect of rural insurgency on democracy is negative and statistically significant (at the 5% level) during the 1990s and 2000s. We hypothesise this post-Cold-War effect is due to the fact that it was not until the collapse of the Soviet Union that African countries became relatively free from the influence of international geopolitical factors, and, as a consequence, domestic political actors started playing a more decisive role in shaping local institutions. In other words, democracy levels in Africa tended to be lower during the Cold War, for reasons that provisionally nullified the effect of the type of anti-colonial movement.

Figure 3. The effect of rural insurgency on democracy by decade

Notes: Black dots represent point estimates from OLS regressions of the average democracy score by decade on the rural insurgency dummy. Vertical bars indicate 90% confidence intervals.

The statistical association between the type of independence movement and democracy that we document in this study is robust to a number of potential confounding factors, which include: time-invariant geographic features and natural resources in each country before independence; social and institutional changes induced by colonialism; and a host of post-independence controls, including income per capita, population size, ethnic cleavages, and religious fractionalisation. Our baseline results suggest that the average level of democracy among rural insurgency countries is about 0.2 points lower (on a 0–1 scale) than the average level of democracy achieved by urban protest countries during the post-1990 period.
Since the type of anti-colonial movement could be endogenous to past democratic or quasi-democratic institutions or experiences, we provide evidence for causality in this relationship by employing two different strategies. First, we use an instrumental variables approach, relating the degree of terrain roughness to the level of democracy, through its impact on the probability that a country experienced an anti-colonial rural insurgency. We rule out alternative accounts by showing that rough terrain does not affect income, violent conflict, ethnic diversity, or religious fractionalisation after independence. Additionally, we perform a sensitivity analysis that relaxes the exclusion restriction assumption, which confirms that our estimated treatment effect remains significant, even when the degree of violation of this assumption is high.

The second strategy exploits the structural break in the democracy data observed after the end of the Cold War. We argue that the collapse of the Soviet Union served as a plausibly exogenous shock that allowed domestic political actors in Africa to play a more decisive role in shaping local institutions, without much international pressure. We then employ a differences-in-differences design, with country fixed effects, to test whether democracy levels changed differentially, after the end of the Cold War, in rural insurgency versus urban protest countries. Our findings confirm that the average level of democratic development is significantly lower in rural insurgency countries than in urban protest countries in the post-1990 period.

We formally test potential channels of causality, and find evidence that anti-colonial movements affected post-independence political systems through the persistence of specific forms of political dissent. Urban mass protests led to non-radical forms of political expression, such as demonstrations or workers strikes, which facilitated peaceful transfers of power, political compromise and ultimately the consolidation of democratic reforms after the Cold War. The reverse is true where rural armed rebellion was the dominant strategy: armed rebellions created norms of violent collective action and repressive forms of government, which hindered the development of democratic institutions. Furthermore, using survey data from the Afrobarometer, we find that respondents from rural insurgency countries are significantly more likely to support the use of violence and one-party rule than respondents from urban protest countries.
Conclusion

This study contributes to a growing literature on the impact of historical events and political culture on democratic change. We show that colonial history matters for African political development, not only because of the policies enacted by the colonial administration, but also because of the way African pro-independence leaders chose to oppose them. Rural insurgencies generated autocratic governments in the long run, while urban movements enabled the development of democratic institutions. Our focus on past political events to explain current institutional outcomes does not imply that structural factors such as economic development, inequality, ethnic diversity, and education are not important in explaining political change. Instead, we argue that historical events – in particular, social movements at critical junctures – might mediate the relationship between structural variables and institutional change.

References


**About the authors**

**Leonard Wantchekon** is Professor in the Politics department and associated faculty in the Economics department at Princeton University. His research is broadly focused on political and economic development, particularly in Africa. His specific interests include the political economy of infrastructure provision, education and human capital externalities, democratization, clientelism and redistributive politics, the resource curse, and the long-term social impact of historical events. He is a member of the American Academy of Arts and Sciences and the founder the African School of Economics (ASE), which opened in Benin in September 2014.

**Omar García-Ponce** is an assistant professor in the Department of Political Science at the University of California, Davis. Previously, he was a Post-Doctoral Fellow at the Center for Global Development. His research focuses broadly on political and economic development, with emphasis on topics related to political behavior, crime, and conflict. He has two main lines of ongoing research. One examines how violence affects behavioural outcomes and the development of institutions. His second line of research is on the political economy of organised crime, with a regional focus on the U.S.-Mexico drug trade. His work has been published in the American Political Science Review, Electoral Studies, and the Journal of the European Economic Association.
India is a land of disparities. Economic and human development outcomes vary considerably across the regions of India. In 2004, annual per capita incomes ranged from 32,000 rupees in the state of Haryana to 6,000 rupees in the state of Bihar. The 1991 census recorded 95% of villages in Kerala as having a primary school, but only 39% of villages in Bihar had one. Similarly, the percentage of villages with paved roads varies across districts from 10% to 100%. These large and persistent disparities mirror the large disparities observed across countries. In 2004, GDP per capita varied from $40,000 in the United States to $780 in Niger. Over the past two decades, a substantial body of research has linked such persistent disparities to the long-run effects of historical institutions, often introduced by colonial powers.¹

An obvious question in this regard is whether the outcome is the long-run effect of colonial rule itself, or the persistent effects of historical governance arrangements. Many cross-country studies find that those governance arrangements under which native rulers retain a substantial degree of autonomy (‘indirect’ colonial rule) result in better long-run outcomes.²

India provides an excellent venue for comparing the long-run consequences of direct versus indirect colonial rule, since about 45% of the total area and 23% of the population were not under direct British colonial rule (see map in figure 1).

¹ See the chapters by Acemoglu and Robinson, Easterly and Levine, Feyrer and Sacerdote and La Porta et al. in this eBook. Nunn (2009) provides a detailed survey of this literature.

They were ruled by Indian kings, who had considerable autonomy over most internal matters, including administration and revenue collection systems, land tenure, legal codes, currency, education policy and health initiatives. These places were called native states or princely states by the British. Annexation of areas into direct British rule was halted after the 1857 mutiny of Indian soldiers against their British army officers. After the end of colonial rule in 1947, all of these areas were integrated into independent India and have since been subject to a uniform administrative, legal, and political structure.

Comparison of different areas of the same country has several advantages, relative to cross-country comparisons: the range of potential omitted variables is limited, we are comparing areas with similar postcolonial political and administrative arrangements, and the historical record provides exogenous policy changes, by means of which we are able to obtain causal estimates of the long-run effects of historical institutions.

Even within-country comparisons of places with different types of colonial rule may yield biased results, because there might be many characteristics that make an area more likely to be administered directly by a colonial power (such as a location on the coast, or the production of particularly desirable commodities), and these characteristics might also determine long-term outcomes for the area, irrespective of what kind of colonial rule was imposed. A look at the map in Figure 1 reveals a strong British colonial preference for places with better agricultural potential: the desert areas of western India and the forested areas of central India were not annexed, while almost all areas of the fertile Gangetic plains were.

To control for such selectivity concerning direct British government rule, I take advantage of a historical policy change. Lord Dalhousie, Governor-General from 1848 to 1856, enacted a policy whereby any native states where the ruler died without a natural born heir could be annexed (the so-called ‘Doctrine of Lapse’). Since Dalhousie’s predecessors and his successor Governors-General recognised adopted heirs and did not follow this policy, the Doctrine of Lapse constitutes a purely circumstantial means

---

3 The foreign and defense policies of these states were controlled by the British: native states could not declare war and peace, or even communicate with other states, without British permission, and they were limited in the size of the armies they could maintain. Most native states also had tribute obligations to the British.
of being subjected to direct colonial rule, \textit{independent of the characteristics of the local area}. This means that comparing states which experienced ‘Lapse’ (death of a ruler without a natural heir in 1848-56) to those which did not will give us a causal estimate of the impact of direct British rule. Of course, this entails restricting our analysis to areas that were not already annexed by 1848.

\textbf{Figure 1.} Geographical distribution of native states in India

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{map.png}
\caption{Geographical distribution of native states in India}
\end{figure}

\textit{Source: Iyer (2010)}
We should note that Lapse was indeed a significant determinant of annexation into British India during the tenure of Lord Dalhousie. During 1848-56, a ruler died without heir in eight native states, covering 20 modern districts; 80% of these districts were annexed into British India. Of the areas where such a death did not occur, only 13% were annexed (by conquest or other means) during the same period.

**Table 1.** Access to public goods and human development outcomes across districts where ‘lapse’ did and did not occur

<table>
<thead>
<tr>
<th></th>
<th>Lapse happened</th>
<th>Lapse did not happen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction of districts annexed</td>
<td>0.8</td>
<td>0.13</td>
</tr>
<tr>
<td>1991 fraction of villages having</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary school</td>
<td>0.81</td>
<td>0.81</td>
</tr>
<tr>
<td>Middle school</td>
<td>0.21</td>
<td>0.29</td>
</tr>
<tr>
<td>Primary health center</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>Primary health subcenter</td>
<td>0.04</td>
<td>0.1</td>
</tr>
<tr>
<td>Canals</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>Paved roads</td>
<td>0.29</td>
<td>0.45</td>
</tr>
<tr>
<td>Human development outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literacy rate 1991 (%)</td>
<td>41</td>
<td>40</td>
</tr>
<tr>
<td>Infant mortality 1991</td>
<td>106</td>
<td>82</td>
</tr>
</tbody>
</table>

Using data from the census of 1991, I compare the extent of public goods provision in the two types of areas, by computing the fraction of villages in each district having access to schools, medical facilities, and infrastructure such as canals and paved roads (Table 1). We find that areas where Lapse occurred have significantly fewer villages with access to public goods, such as middle schools, primary health centres and sub-centres, canals and paved roads. The magnitudes of these differences are large: areas where Lapse did not happen have twice as many villages with access to health centres and canals, 55% more villages connected to roads and 38% more villages supplied with middle schools.
(see Banerjee et al. 2005, for an analysis of a more detailed list of public goods). These differences in public goods access have implications for human development: infant mortality rates are 23% lower in non-Lapse areas compared to Lapse areas.

Interestingly, we do not see any differences in access to primary schools (and consequently literacy rates) in the 1991 census. This is directly attributable to the explicit policies of the post-colonial Indian state aimed at equalising access to public goods (Banerjee and Somanathan 2007). If we look at data from the 1961 census, we find that only 39% of villages in the Lapse areas had access to primary schools, compared to 46% in non-Lapse areas. Post-colonial policies in this case have been able to gradually erase the impact of history, although the effects can be discerned more than three decades after the end of colonial rule.

In Iyer (2010), I document three further findings that make us more confident of the conclusion that these differences are purely because of direct colonial rule, rather than other characteristics of Lapse areas. First, I show that the differences between Lapse and non-Lapse areas are statistically significant (i.e. unlikely to occur by chance), and continue to hold, even after controlling for geographic characteristics of the areas. Second, I use the fact that all annexation was stopped in 1858, so, if kings died without heirs after this date, their states would not get annexed. When I compare areas where kings died without heirs in 1858-1884 with other native states, there are no large differences to be seen in access to middle schools, health facilities or roads – thus lending greater credibility to the argument that it is actually the institution of direct colonial rule which makes these areas lag behind. Third, I perform several statistical robustness checks, including propensity score matching, adjusting standard errors through resampling and including state fixed effects.

Finally, I consider the question of why indirect rule may lead to these areas having a long-run advantage. I find that differences in the levels of taxation, political participation or land tenure systems are unable to explain the gap between directly and indirectly ruled areas. The key difference lies in the incentives for the administrators. Although the British gave up outright annexation of territory, they reserved the right to depose Indian kings if their states were misgoverned. Several rulers of natives states were in fact deposed by the British for misrule.
On the other hand, British colonial administrators did not face any significant penalty for delivering poor outcomes, and their average tenures in a given area were much shorter than for the typical Indian ruler. I find that native states where the rulers had been deposed for misrule by the British tend to have public goods outcomes similar to those for directly ruled British areas. In other words, areas which were poorly governed in colonial times tend to have worse public service delivery in postcolonial times as well – and rulers in indirectly ruled areas overall had better governance incentives in the colonial period.

**References**


---

4 In this sense, it may be unfortunate that independent India has opted for the British system of administration, with most policy implementation under the control of career bureaucrats with very short tenures in specific posts (Iyer and Mani 2012).
Direct versus indirect colonial rule in India
Lakshmi Iyer


About the author

Lakshmi Iyer is an Associate Professor of Economics at the University of Notre Dame. Her primary research fields are development economics and political economy, with a special emphasis on property rights and the distribution of political power within societies. Her research has examined many dimensions of the distribution of political power within emerging market countries, including the legacy of colonial rule, the division of authority between politicians and bureaucrats, and the determinants of conflict. She has also studied historical and current property rights institutions in several developing countries including India, Vietnam, China and the Philippines. Her work has been published in leading academic journals in economics. She has also authored a book of case studies Institutions, Institutional Change and Economic Performance in Emerging Markets.

She holds Bachelors and Masters degrees from the Indian Statistical Institute, and a PhD from the Massachusetts Institute of Technology (MIT). Her work can be found at: http://www3.nd.edu/~liyer/.
An examination into the case of China’s treaty ports provides a useful testing ground for studying the interconnection between history and development. First, the treaty ports were forced to open as a result of treaties between China and Western countries, following wars, and this opening itself establishes a quasi-natural experiment with which to study whether history is of importance for development. Second, between the 19th century and today, China has experienced alternate phases of closedness and openness; this dynamic environment might enable us to better understand some of the mechanisms through which history has a bearing on understanding economic factors of today.

The treaty port system dates back to the late Qing dynasty of China, which is usually described as a dark and shameful time in Chinese history. One important feature of this period is that the Qing government signed many ‘unequal treaties’ with Western countries. Along with these treaties, China conceded more than 40 cities, called ‘treaty ports’, to Western countries from the 1840s to the 1910s. The Westerners established municipal authorities, factories, schools, police and judiciaries in these ports. A hundred years later, in January 1943, China signed treaties with Britain and the United States that abolished extraterritoriality, and so the treaty ports system ceased to exist.

If we simply compare the economic performance of treaty ports with other locations today, it seems obvious that the treaty ports are better developed: indeed, many prosperous cities, such as Shanghai, Guangzhou and Tianjin, were treaty ports in the past. However, due to a lack of data and some selection concerns, little is known about the impacts of the treaty ports system on development.

In Jia (2014), I construct a prefecture-level dataset from 1776 to 2000 to examine the development paths of treaty ports. To relieve the selection concern, I restrict my attention to prefectures along the coast or along the Yangtze river and compare the development paths of prefectures with treaty ports and a control group of prefectures with similar characteristics.
Figure 1 plots the growth of the treaty ports system, where the shaded areas indicate the prefectures in the subgroup I focus on.

**Figure 1. Growth of the treaty ports system**

Notes: Wave 1 was opened in the 1840s; Wave 2 was opened in the 1860s; Wave 3 was opened during the 1870s and the 1880s; Wave 4 was opened in the 1880s and later. The dark areas are the prefectures in the subgroup I focus on.

**Documenting the development paths**

I find that prefectures with treaty ports grew faster in terms of population sizes after the ports had become open. After the Communist revolution in 1949, when all parts of China were closed to foreign influence, the two types of areas grew at similar rates. This finding is not surprising, as the economy was heavily regulated and reducing regional differences was one of the most important political goals of the newly established government. However, when China once more began to open up, after 1980, it was the places with treaty ports which were among the first to take advantage of the new globalisation opportunities, as shown by a higher population growth as well as faster growth of GDP per capita.
Figure 2 visualises the trends of population size over time for the treaty ports group and the comparison group. On average, the population growth rate of the treaty ports group is about 30% higher, except for the period between 1949-1978 when the advantage of the treaty port group disappears. The magnitude is similar in terms of the growth rate of GDP per capita available since the 1980s.

**Figure 2.** Trends in population size, from 1776 to 2000

![Figure 2: Trends in population size, from 1776 to 2000](image)

*Notes:* Population size measured in 10,000. The figure combines different waves of treaty ports together. Here the treaty ports group includes all prefectures that finally became treaty ports.

**Understanding the long-run impact**

In this study, I denote the effect until the 1950s as the short-run impact and the effect after the 1980s as the long-run impact. Many mechanisms can explain why the short-run impact was that population sizes grew faster in treaty ports after their opening. For instance, Westerners built modern hospitals in treaty ports, which provided better public health services. Additionally, Westerners also established factories, which provided more economic opportunities and, hence, attracted migrants. I do not attempt to disentangle all these channels that might explain the short-run impact.
Instead, I focus on mechanisms that might explain the long-run impact. As China once more began to open up, after the 1980s, why were the treaty ports among the first to take advantage of the new globalisation opportunities?

Many studies in this line of literature have examined why history is of importance for long-run development outcomes. Existing explanations include geography and natural resources (Diamond 1997, Rappaport and Sachs 2003), specific institutions such as legal systems (La Porta et al. 1998) and property rights (Acemoglu et al. 2002, Banerjee and Iyer 2005), human capital (Glaeser et al. 2004) as well as social norms (Tabellini 2008, and Nunn and Wantchekon 2001).

In this context, the places studied share similar geographical features (along the coastal line or along the Yangtze River) and all of them are ‘accessible’ to trade, which suggests that geography is unlikely to be the main mechanism. During the period 1949-1978, formal Western institutions were basically abolished but, once the country had been reopened, the former treaty ports still developed better. Hence, geography and tangible institutions are not likely to be the key mechanism. In contrast, human capital and social norms of doing business are more likely to have played a role in explaining the long-run impacts.

It is challenging to find comparable measures of human capital and social norms from history to today. Instead of examining these factors separately, I focus on more explicit measures of demographic and economic activities. I employ modern data on birth rates, death rates and migration. I find the birth rates and death rates to be similar for prefectures in the subsample. However, migration into the treaty ports areas from other provinces plays an important role. This finding suggests that better economic opportunities in the treaty ports areas attract more people, not from their neighbouring prefectures in the control group but rather from other areas of China. In terms of economic activities, I examine the growth of different sectors. I find the commerce and service sector to be the main driver of the long-run advantage of the treaty ports group.
Summary

The development paths of the treaty ports, in the alternate phases of closedness and openness experienced by China from the 19th century until today, provides a laboratory to understand the importance of history for current economic development. Jia (2014) focuses on development outcomes at the macro level and opens new avenues for future research, such as analysing the dynamic paths of institutions in these ports and understanding trade patterns using micro-level data.

References


**About the author**

**Ruixue Jia** is an assistant professor of Economics at the School of Global Policy and Strategy, UC San Diego, where she focuses her research on the political economy, development economics, and economic history. She received her PhD from the Institute for International Economic Studies, Stockholm University in 2013.

One stream of Jia’s research focuses on understanding elite formation and elite influence, in both historical and modern contexts. She has published research on how open access to elite status matters for political stability. She has also studied how competence and loyalty jointly determines who becomes a top politician in China. Her recent work is to understand how China’s national college entrance exam affects elite formation and social mobility. A second stream of Jia’s research examines the deep historical roots of economic development. She has published research on the long-run impact of the treaty ports system and the determinants of peasant rebellion. A third theme of Jia’s research focuses on the evolution and the impact of social norms. She has studied the impact of social norms on ethnic choice in China and elderly suicide in South Korea.
Despite significant improvements in the last century, gender disparities persist. In the workplace, women continue to work less, earn less, and are less likely to reach positions of leadership. At home, women still do most of the housework and remain the primary caregivers. Gender disparities manifest themselves even more severely in regions where heavily male-biased sex ratios prevail, such as China and India, where an estimated 80 million women are missing today. This chapter traces present-day gender roles back to the marriage market conditions that prevailed in the past and shows the impact of male-biased sex ratios. In particular, we show that the effects of male-biased sex ratios in 19th century Australia have persisted until today, well after sex ratios have reverted back to normal. A surplus of males in the past has shaped cultural attitudes towards gender roles, and still affects labour supply decisions, occupations, and marital satisfaction, but in ways that are not necessarily negative for – all – women.

Introduction

There are an estimated 100 million women missing in the world today, 80 million in China and India alone, due to sex-selective abortion and differential gender mortality (Hesketh and Xing 2006). An important question is how male-biased sex ratios will further affect female outcomes, and whether these effects will persist in the long run.
Biased local sex ratios can have a variety of effects on marriage markets, relationship outcomes and on labour supply decisions (Becker 1973, 1974). These effects could persist in the long run, even after the imbalance itself is corrected. Indeed, as is shown repeatedly in this publication, behaviours shaped by short-run factors can become practices that persist in the long run. For example, the Neolithic revolution and the use of plough cultivation still influence gender roles and female outcomes today (Alesina et al. 2011, 2013, Hansen et al. 2015).

Answering the question of how male-biased sex ratios will affect the relative welfare of men and women is very difficult. Male-biased sex ratios are generally the product of poorer opportunities for women (Qian 2008, Carranza 2014) and of cultural preferences for sons (Almond et al. 2013). The employment opportunities of women are limited in many economies that place a high premium on physical strength, for example because of specialisation in agriculture or other extractive industries (Chung and Das Gupta 2007). In other cases, where only sons are able to carry the family name (Hesketh and Xing 2006), or where there is a financial burden imposed by the payment of a dowry upon a daughter’s marriage (Rao 1993), this can lead to cultural preferences for having a son. Be it driven by economic or cultural factors, a lower value placed upon having daughters can lead to sex-selective abortion and female infanticide (Qian 2008, Almond et al. 2013). In these contexts, it is impossible to know for certain whether worse female outcomes are the cause of male-biased sex ratios, or their consequence. The ideal natural experiment would consist of placing a larger number of men than women on an isolated island, with these men and women being of a similar cultural background and operating in the same institutional environment, and then observe female outcomes from that point on.

Convict transportation to Australia and male-biased sex ratios

We exploit such a natural experiment in a series of recent papers (Grosjean and Khattar 2015, Grosjean and Brooks 2016).

In the late 18th and 19th centuries, the British policy of convict deportation to Australia resulted in heavily male-biased sex ratios. Men vastly outnumbered women among convicts.
Between 1787 and 1868, 132,308 and 24,960 convict men and women were transported to Australia, more than five males for every female convict. This overall figure hides heterogeneity over time, with the initial waves of convicts even more male biased.

Convicts were not confined to prisons, but were allocated to work, first under government supervision, and then, as the number of free settlers and emancipists (ex-convicts) grew, under the direction of private employers. Convicts were generally freed after seven years.

Even among free migrants, men vastly outnumbered women, as it was mostly men who sought economic opportunities in Australia, which consisted chiefly of agriculture and, later, after the discovery of gold in the 1850s, mining. As a result of sustained convict transportation and male-biased free migration, male-biased sex ratios endured in Australia for more than a century, although less severely after the end of convict transportation, as illustrated in Figure 1.

**Figure 1.** Sex ratio in Australia: number of men to every woman, 1830-2011

Notes: This figure plots the number of men over the number of women in the whole population, that is to say: convicts, emancipists (ex-convicts), and free migrants, as well as people born in the colony. It does not include Aboriginal or Torres Strait Islanders, who were not counted in the Census until the 1960s. Only very rough estimates are available for these populations.

Source: Australian Bureau of Statistics.
We collect information on the historical sex ratio (number of men for every one woman), as well as other historical characteristics, such as total population and the structure of the economy, from the first Census taken in each of the six Australian states: the 1836 New South Wales Census (which also included what is now the Australian Capital Territory), the 1842 Tasmanian Census, the 1844 South Australian Census, the 1848 Western Australian Census, the 1854 Victorian Census, and the 1861 Queensland Census. Convicts were only present in New South Wales and Tasmania. We then match our historical data to present-day data from the 2011 Census and a nationally representative household survey (HILDA) in order to study the long run effects of male-biased sex ratios. In particular, we are interested in the effects of male-biased sex ratios on female labour supply in the labour market, time spent on household work and taking care of children, and on the nature of female occupations, as well as on how much leisure women enjoy, and marital and life satisfaction.

Prior to this work, digitised shapefiles on Australian historical Census boundaries did not exist. We collected electronic copies and digitised hard copies of maps from the National Library of Australia and from State Libraries in order to construct these boundaries and match historical counties to present-day boundaries.

**The short-run effects of male-biased sex ratios**

We first study the short run effects of male-biased sex ratios. The limitations of our historical dataset restrict our focus on marriage rates and female labour force participation. In the presence of male-biased sex ratios, we find that women married more and reduced their labour supply. We also find that the quality of female occupations is negatively affected.

These results, contemporaneous to sex ratios, are in line with previous literature and the predictions of bargaining models of intra-household decision making (Grossbard-Schechtman 1984, Chiappori 1988, Chiappori *et al.* 2002), and with simple supply and demand models of the marriage market (Becker 1973, 1974). More interesting is to study whether the effects of male-biased sex ratios have persisted until today, even after sex ratios have reverted back to normal.
The long-run effects of male-biased sex ratios

Today, 150 years after the historical sex ratio that we measured, and even though sex ratios have long returned to parity, we find that people have more conservative attitudes towards women working and that women still work fewer hours in counties that were more male-biased in the past. As a probable consequence of the reduction in their labour supply and of the conservative attitudes held, women are still less likely to reach high-ranking occupations. However, we do not observe that women spend more time on household chores, or taking care of children. If anything, they spend less time. As a result, since they work less in the labour market but do not work more at home, women today consume more leisure in areas that were more heavily male-biased in the past.

In terms of magnitude, the long-run effects of the historical sex ratio on attitudes towards gender roles is comparable, at the mean, to 30% of the effect of being a female versus a male. Moreover, a one standard deviation increase in the historical sex ratio is associated with a 3.5% reduction in working hours supplied by females and a decrease in the share of women employed in high-ranking occupations by 0.13 standard deviations. Historical circumstances explain more than 3% of the variation in the share of women employed in high-ranking occupations that is left unexplained by traditional factors, even when accounting for the share of men employed in similar professions.

In Grosjean and Brooks (2016), we also study the long-term effects of male-biased sex ratios on the quality of relationships. We find that both men and women are more satisfied with their relationship in areas that were more male-biased in the past. In light of women experiencing more leisure, the result that women are more satisfied is perhaps not surprising. What is more surprising is that men also experience higher marital satisfaction. This result is surprising because it is in contrast with the theoretical prediction that scarcity of women will increase the bargaining position of women, but should reduce the bargaining position of men, meaning that we should observe a negative effect of men’s welfare. An interpretation of this result is that because women have higher bargaining power, they are more selective and search for a better match, and, as an indirect effect, men also benefit from this better match quality. Other possible interpretations are that men directly benefit from having happier wives, or are more satisfied with a ‘traditional’ division of labour within the household and ‘traditional’ values.

1 The average historical sex ratio is 2 to 1. Women are 46% more likely to hold progressive attitudes compared to men.
Conclusion

To the best of our knowledge, we are the first to document that sex ratios can leave a large and persistent imprint on labour supply decisions, occupations, and marital satisfaction, even long after sex ratios have returned to parity.

We want to emphasise that the outcome that women consume more leisure, but face limited opportunities in the workplace, may be welfare-improving for some – but not necessarily all – women.

Although our results may be specific to a certain technological context – work opportunities for women were very poor in 19th Century Australia – and although the average deviation from a balanced sex ratio that we study is larger than deviations observed today in the world, a noteworthy implication is that a temporary imbalance in the sex ratio can have significant consequences on society that endure well beyond the imbalance itself.

References


Grosjean, Pauline and Robert Brooks (2016), *The Persistent Effect of Sex Ratios on the Quality of Relationships*, mimeo UNSW.


**About the author**

Pauline Grosjean is an Associate Professor in the School of Economics at UNSW. Previously at the University of San Francisco and the University of California at Berkeley, she has also worked as an Economist at the European Bank for Reconstruction and Development (EBRD). She completed her PhD in economics at the University of Toulouse in 2006 after graduating from Ecole Normale Supérieure. Her research interests include the economics of institutions and culture, development and transition, political economy and economic history.


**About the author**

**Pauline Grosjean** is an Associate Professor in the School of Economics at UNSW. Previously at the University of San Francisco and the University of California at Berkeley, she has also worked as an Economist at the European Bank for Reconstruction and Development (EBRD). She completed her PhD in economics at the University of Toulouse in 2006 after graduating from Ecole Normale Supérieure. Her research interests include the economics of institutions and culture, development and transition, political economy and economic history.
The imprint of the European colonisers on contemporary African development and culture is multifaceted and far-reaching. This volume provides an overview of how economists have attempted to shed light on these legacies. The range of the topics covered is breathtaking, starting from the slave trades, continuing with the drawing of the colonial borders in European capitals, the extent of indirect rule, the use of bloody concessions in DRC, the investments in transportation infrastructure, the role of Christian missions, and ending with the type of the independence movements in Africa. The geographic focus of the volume then turns to exploring specific aspects of the colonising experience in Asia and the Pacific looking at the role of direct versus indirect British rule in India, the economic trajectory of China’s treaty ports and the impact on current-day gender attitudes of the male-biased sex ratios of early colonisers in Australia.

To get to these fascinating questions a wealth of diverse sources including historical archives, anthropological maps, and satellite images are fruitfully combined and subjected to state of the art econometric techniques and theoretical models. The findings of this ambitious research agenda are novel highlighting the shadow of history on the various aspects of the economy and the polity. While there are many open issues and debates and although development is not deterministic, one message is clear: We are shaped by history (Martin Luther King Jr).