Cash Crops, Colonialism and Legacies of Spatial Inequality: Evidence from Africa

Philip Roessler  William & Mary
proessler@wm.edu

Yannick I. Pengl  ETH Zurich
yannick.pengl@icr.gess.ethz.ch

Robert Marty  World Bank
robmarty3@gmail.com

Kyle Sorlie Titlow  William & Mary
kyletitlow@gmail.com

Nicolas van de Walle  Cornell
nv38@cornell.edu

July 1, 2020

Abstract. Spatial inequality is a key cause of conflict and barrier to economic growth. What accounts for this phenomenon and why is it highest in Africa? Bridging geographic and institutional theories, we attribute the continent’s severe subnational disparities in wealth to the interactive effects of the spread of the cash crop revolution and economic imperialism from the end of the 18th century onward. This confluence shifted the loci of economic production to smallholder farmers in areas suitable for cash crops, but it also fueled the vertical integration of agricultural zones with forward production linkages in Europe. The result was localized agglomeration, but with limited or negative spillovers to nearby areas. Using agro-climatic suitability scores and historical data on the source location of more than 95 percent of all exports across 38 African states, we find that colonial cash crop production had positive but locally concentrated long-run effects on urbanization, road infrastructure, nighttime luminosity and household wealth. These effects rival or surpass other geographic and historical forces. Path dependence due to colonial infrastructure investments is the more important causal mechanism than continued advantages in agricultural productivity. Yet, we also find that this agglomeration came at the expense of surrounding areas, leaving a legacy of spatial inequality and ethno-political conflict.
Colonialism almost never exploits the entire country. It is content with extracting natural resources and exporting them to the metropolitan industries thereby enabling a specific sector to grow relatively wealthy, while the rest of the colony continues, or rather sinks, into underdevelopment and poverty. In the aftermath of independence the nationals who live in the prosperous regions realize their good fortune and their gut reaction is to refuse to feed the rest of the nation. The regions rich in groundnuts, cocoa and diamonds stand out against the empty panorama offered by the rest of the country.

— Frantz Fanon, *The Wretched of the Earth*, 1963

**Introduction**

Few forces have been more important in the creation of the modern world than European imperial expansion from the end of the 15th century onward. Spurred by technological change, geopolitical competition and the pursuit of natural resources, over the next four centuries European powers would control wide swathes of the world’s territory. The long-run effects of European colonialism have been the focus of a large and influential body of scholarship in international relations and the social sciences more broadly—ranging from, among other important lines of inquiry, state and international-boundary formation; globalization and international economic

---

1 Jackson and Rosberg 1982; Herbst 2000; Clapham et al. 1996; Alesina, Easterly, and Matuszeski
structures; the causes of interstate and, especially, intrastate conflict; institutional change; and inequality.

We build on this literature, focusing on the colonial origins of inequality, in particular spatial inequality. In contrast to a classical literature in economics that conceived of “unbalanced growth” as a positive lever for long-run development, new scholarship finds that severe disparities in subnational wealth—especially when organized along ethnic lines—negatively affect economic performance. Such ethnoregional disparities are also found to elevate intra-state conflict risk as horizontal inequalities increase grievances, worsen bargaining problems and undermine ethnic powersharing. Spatial inequality thus may represent one of the root causes of the vicious poverty-conflict trap that affects many low-income countries, particularly in Africa—the region with the highest levels of spatial inequality in the world.

What are the roots of spatial inequality? Existing global studies tend to emphasize the primacy of underlying geography. Colonialism, if addressed, is treated as epiphenomenal. Africa-specific scholarship takes the spatial implications of colonialism much more seriously. But existing research also tends to fall along the geography-versus-colonialism fault line. One important stream emphasizes continuity and persistence in African political topography and economic geography from the pre-colonial period until the present, focusing on the recurring and path dependent effects of

2Lenin 1999; Wallerstein 2011; Cardoso and Faletto 1979.
4Acemoglu, Johnson, and Robinson 2001; Lange, Mahoney, and Vom Hau 2006; Lange 2009; Acemoglu and Robinson 2012.
5Engerman and Sokoloff 1997; Sokoloff and Engerman 2000; Angeles 2007.
6Hirschman 1958.
8Stewart 2008; Ostby 2008; Cederman, Gleditsch, and Buhaug 2013.
9See Table A5.
10See Alesina, Michalopoulos, and Papaioannou 2016; Henderson et al. 2018; Achten and Smarzynska Javorcik 2019. As we discuss further below, Henderson et al. also model geography’s interactive effects with market and technological change.
the region’s unique biogeographical fundamentals.\textsuperscript{12} In contrast, other prominent studies underscore the exogenous impact of imperialism—from the slave trades to colonialism—and their disruptive effects on spatial equilibria of development.\textsuperscript{13}

We contribute to bridging these research streams. To do so, we advance an integrated framework that accounts for the interactive effects of geography and colonial extraction on long-run development across Africa. Our framework enables us not only to address the economic geography of colonialism—that is, why the colonial state invested where it did—but it also sheds light on its potential distributional consequences—whether if, as Fanon suggests in the epigraph, colonial extraction’s development benefits for some regions came at the expense of others, entrenching spatial inequality.

Our historical starting point is Africa’s structural transformation away from the slave trades to export agriculture beginning from the end of the 18th century. We conceive that this exogenous change in international trade shaped the contours of the modern African state via the dynamic and interactive effects of three factors: a.) the shift in the value of geographic fundamentals to favor agro-climatic suitability for cash crops (such as, oil palm, groundnuts, cocoa, coffee, and cotton); b.) trade costs; and c.) imperial institutions.

With the slave trade’s abolition in the early 1800s and in response to demand from industrializing states for agricultural commodities, the cash crop revolution finally took off in Africa, starting in West Africa.\textsuperscript{14} Led in many places by smallholder farmers, the revolution stimulated agricultural innovation; boosted farmers’ incomes; induced migration to farming areas; and propelled investments in trans-

\textsuperscript{12}Such as, the relative paucity of domesticable plants and animals, the vast continental interior with few inlets or navigable rivers, and disease ecologies. Among other studies on this broader theme, see Diamond 1997; Gallup, Sachs, and Mellinger 1999; Sachs, Mellinger, and Gallup 2001; Marshall and Hildebrand 2002; Olsson and Hibbs 2005; Alsan 2014; Fenske 2014; Alesina, Michalopoulos, and Papaioannou 2016; Mayshar et al. 2018; Michalopoulos, Puterman, and Weil 2019.

\textsuperscript{13}Among other studies, see Nunn 2008; Nunn and Puga 2010; Nunn and Wantchekon 2011; Huillery 2009, 2011; Lee and Schultz 2012; Michalopoulos and Papaioannou 2016; Jedwab and Moradi 2016; Jedwab, Kerby, and Moradi 2017.

portation infrastructure. Thus, in contrast to the plantation economies of the Americas, West African cash crop agriculture tended to resemble the “yeoman farm system” of temperate staples. Areas suitable for cash crop cultivation emerged as loci of localized economic agglomeration as it spurred backward and forward linkages to, respectively, support cultivation and production, and processing, marketing and transport.

However, as the cash crop trade lifted up smallholder farmers, it also pulled European merchants deeper into the region. Occurring at the height of imperialism, European traders leveraged mercantilist practices and institutions—such as gunboat commerce, trading oligopolies, foreign-financed infrastructural investments, and colonialism—to dominate Africa’s agricultural markets. The spread of imperial institutions profoundly shaped the developmental effects of Africa’s smallholder farming revolution. European-financed transportation infrastructure, especially railways, fueled the cash crop trade, leading to significant improvements in farmers’ living standards. But designed to vertically integrate agricultural zones with European markets, this new transportation infrastructure—combined with protective trade policies (e.g., limiting imports to bulk, raw commodities)—dislocated forward production outside of Africa, thwarting local manufacturing and stymieing the domestic economic differentiation that otherwise might have resulted. Beyond the prized cultivation zones and trading hubs, other areas were either neglected or relegated as a labor reserve and intentionally underdeveloped to create a cheap supply of labor to work in cash crop and mining enclaves.

Several important implications follow from our conceptual framework and this

---

17 Hirschman 1977.
18 Rodney 1972.
19 Jedwab and Moradi 2016; Moradi, Austin, and Baten 2013.
20 This is a central line of argument of dependency theory (Frank 1966; Santos 1970; Rodney 1972; Cardoso and Faletto 1979). See also Hirschman (1977) who suggests that imperialism led to underdevelopment via its negative effects on production and fiscal linkages—from which our argument draws inspiration.
21 Amin 1972, 1974; Rodney 1972; Cordell, Gregory, and Piché 1996.
historical context. One, we would expect agro-climatic suitability for cash crops and historical cash crop production zones to be significant predictors of contemporary patterns of economic development across the continent—with colonial infrastructure investments serving as one of the primary channels. Two, while colonial extraction boosted economic agglomeration in highly suitable areas, in the absence of strong domestic production linkages we would expect limited positive spillover effects. And, if anything, given the exploitative nature of colonialism, other peripheral regions may have been left worse off than they otherwise would have been.

To test this argument, we build a comprehensive dataset of historical African economic geography, including detailed geospatial information on sites of cash crop production and mining. The latter data was extracted from a map published in 1961 depicting the source location of more than 95 percent of all exports across 38 African states, standardized in 1957 U.S. dollars ($). Constructed by a team led by renowned geographer of Africa, William Hance, the map draws on “hundreds of sources...including maps, articles, agricultural yearbooks, reports of commodity boards, and product and regional studies.” As far as we know, the Hance dataset is the most exhaustive and granular representation of the spatial diffusion of the cash crop revolution across Africa, but has never been systematically analyzed.

We validated Hance’s map by independently collecting administrative data from colonial records and other historical sources on the volume and value of the most important agricultural commodities and minerals produced by each state (i.e., those that make up at least 10 percent of total exports at the end of colonialism) standardized in 1960 US$ and aggregated at the level of subnational administrative units. Despite employing different units of analysis and our less comprehensive coverage of agricultural and mineral commodities, overall and resource-specific correlations between our dataset and Hance’s are high (see Figure A1). Given the greater coverage and better granularity of Hance’s data, we employ his spatial data in our analysis, measuring production at the extensive and intensive margins for 0.25×0.25 degree

\[22\text{Hance, Kotschar, and Peterec 1961.}\]
\[23\text{For methods and sources, see Appendix B “Sources of Historical Primary Commodity Production by Country.”}\]
To preview our results, we find that areas of historical cash crop production are significantly better off today on a set of infrastructural and wealth measures than comparable cells within the same African country. Colonial cash crop cells had a 16 percentage points higher probability of having a quality road in 1998; close to a 20 percentage points higher likelihood of emitting nighttime lights in 2015; a 19 percentage points higher likelihood of having a city in 2015; and 14 percent of a standard deviation greater household wealth. Historical cash crop production exhibits a comparable effect on contemporary roads, electrification and cities as colonial mineral extraction, despite the more capital-intensive nature of mining.

To address endogeneity concerns, we use propensity score weighted regressions; employ a randomization inference-type analysis that samples 1000 plausible spatial equilibria of colonial resource extraction; and run spatial 2SLS-IV models instrumenting observed cash crop production with the mean agro-climatic suitability score across the nine most important African export crops (cocoa, coffee, cotton, groundnuts, oil palm, tea, sugarcane, tobacco, bananas). This suitability measure predicts colonial production well and is unrelated to pre-colonial development outcomes. Results from these additional analyses are in-line with the baseline OLS models. In terms of the relative substantive significance of colonial cash crop agriculture on long-run development, we show that, next to distance to the coast, cash crop suitability surpasses or rivals other factors, such as caloric suitability or disease ecology. We show that the persistence of the cash crop revolution’s agglomerating effects operates more through the path dependency induced by colonial infrastructural investments in roads, railways and power generation than continued agricultural production.

We further find that commercial export agriculture generally had moderate spillover effects of up to 50 kilometers. Yet, in line with our argument that economic agglomeration from colonial extraction arose only as it deepened spatial inequities, spillover effects on electrification and urbanization are negative and significant after only 75 km. This suggests that under colonial extractive institutions cash crop agriculture’s developmental benefits came at the expense of other areas—which are worse off today
than would be expected based on their underlying characteristics.

Finally, we show that the legacies of inequality left by colonial extraction have important implications for contemporary patterns of ethno-political conflict. Those neglected from the colonial economy were not only left worse off but more likely to rebel in the post-independence period.

Overall, these findings mark an important contribution to the debate on the long-run effects of colonialism on development. At the country-level extractive colonialism is found to leave a legacy of weak institutions, poor property rights and underdevelopment. On the other hand, extractive processes may generate positive local agglomerating effects via investments in infrastructure and manufacturing. One important question in terms of the local developmental effects of extractive systems, however, is whether these processes are generating positive or negative spillovers. If the latter, then benefits for some are coming at the expense of others. Our spillover analysis provides at least suggestive evidence that colonial extraction generated negative spillovers—making neighboring areas worse off than we would expect based on underlying characteristics consistent with the mercantilist logic of colonialism laid out above. This analysis helps to reconcile the seeming differences between national-level and local-level impact of extractive colonialism. Rather than offsetting negative institutional effects, the subnational extractive processes likely made them worse. Colonialism left much of African trapped in a negative feedback loop of weak institutions and spatial inequities.

**State-building and Development in Africa**

Existing scholarship on African political economy revolves around the path dependent effects of underlying geographic factors or the legacies of the region's extractive institutions. In this section we briefly review the existing literature before developing our conceptual framework.

---

Biogeography

In accounting for the prevalence of weak, low-income states across Africa, a prominent body of scholarship links it to the region’s unique biogeography and the enduring consequences of its late Neolithic transition to sedentary agriculture. A combination of unfavorable locational fundamentals—erratic rainfall in the savanna; the relative paucity of domesticable plants and animals; endemic diseases such as trypanosomiasis and malaria; tropical climate; productivity advantage of roots and tubers over cereals; vast continental interior with few inlets or navigable rivers; and its north-south axis—not only delayed the uptake of farming but, once adopted, trapped these societies in a low agriculture-weak state equilibrium. Subsistence and shifting agriculture along with pastoralism predominated, limiting population growth and contributing to the fractal nature of the region’s state system. Consequently, compared to Eurasia, the Neolithic Revolution did not leave Africa on a path to deep economic and political stratification. In fact, unlike the rest of the world, in Africa the relationship between agricultural potential, population density and historical state formation is weak. Instead, the centralized polities that emerged and endured tended to lie astride ecological boundaries and arise from long-distance trade rather than intensive agricultural cultivation.

Outside of Africa’s trade-based states, which appear to have some persistence as centers of economic development, politics tended to revolve around less hierarchical kinship-based groups shaped by local differences in geographic and environmental conditions. This social differentiation is posited to have led to variation in

---

26 Diamond 1997; Olsson and Hibbs 2005.
29 Coquery-Vidrovitch 1975; Osafo-Kwaako and Robinson 2013.
30 See Osafo-Kwaako and Robinson 2013. Moreover, across the region, caloric suitability and cereal advantage are, if anything, negatively correlated with pre-colonial political centralization. See Table A6.
32 Fenske 2014; Michalopoulos and Papaioannou 2013.
33 Michalopoulos 2012; Enke 2019; Cervellati, Chiovelli, and Esposito 2019.
human capital formation and income disparities across ethnic lines.\textsuperscript{34}

The persistence of Africa’s biogeography on uneven patterns of development is proposed to operate through various channels: 1.) path dependency—e.g., differential returns that arise from variation in land endowments and human capital formation;\textsuperscript{35} 2.) institutions—e.g., variation in political centralization;\textsuperscript{36} 3.) culture—e.g., the adoption and transmission of different economic lifeways\textsuperscript{37} or ethnogenesis;\textsuperscript{38} and 4.) the recurring effects of locational fundamentals on subsequent economic forces, such as globalization\textsuperscript{39} or the spread of Christian missionaries.\textsuperscript{40}

**Extractive Institutions**

An alternative stream of research emphasizes the importance of institutions over geography, especially the impact that extractive institutions have played. Two of the most pernicious forms of economic extraction have dominated Africa over the last five hundred years: the slave trades and colonialism. Both are shown to have persistent negative general equilibrium effects on development via their institutional legacies.\textsuperscript{41} Yet, in terms of its spatial implications, colonialism appears to have had more substantial effects than the slave trades.\textsuperscript{42}

On the spatial effects of colonialism, Huillery finds that, independent of underlying pre-colonial characteristics, such as political centralization, trade and population density, the distribution of colonial public investments in health and education

\textsuperscript{34} Alsan 2014; Michalopoulos 2012.
\textsuperscript{35} Alesina, Michalopoulos, and Papaioannou 2016.
\textsuperscript{36} Gennaioli and Rainer 2007; Alsan 2014; Michalopoulos and Papaioannou 2013; Fenske 2014.
\textsuperscript{37} Michalopoulos, Puttersman, and Weil 2019.
\textsuperscript{38} Michalopoulos 2012; Cervellati, Chiovelli, and Esposito 2019.
\textsuperscript{39} Henderson et al. 2018.
\textsuperscript{40} Jedwab, Meier zu Selhausen, Moradi, et al. 2019.
\textsuperscript{41} Nunn 2008; Nunn and Wantchekon 2011; Acemoglu, Johnson, and Robinson 2001, 2002.
\textsuperscript{42} One paradoxical spatial implication of the slave trade is that in Africa—unlike the rest of the world—areas with rugged terrain tend to be richer not poorer, despite significant barriers to trade and other economic production, because of the protections it offered from slave raiding. Nunn and Puga 2010. However, as we show in Table A19, those areas most affected by the slave trades do not appear to have permanently fallen behind in terms of infrastructural development within countries.
produced significant path dependent effects in French West Africa.\textsuperscript{43} Perhaps even more important in changing the contours of African economic geography were colonial investments in transportation infrastructure.\textsuperscript{44} Historically, high transportation costs have proven one of the most significant barriers to trade in the region.\textsuperscript{45} The colonial rail revolution in Africa—in which, in the span of a century 58,716 kilometers of railroad tracks were built—significantly reduced transportation costs.\textsuperscript{46} In doing so, it opened up previously underdeveloped areas to trade, spurring an increase in economic activities, agricultural production and mining, which in turn increased labor demands. Strikingly, it is estimated that outside of Africa’s largest cities, 55 percent of aggregate change in urbanization between 1900 and 1960 occurred within a 10-kilometer radius of colonial railroads.\textsuperscript{47}

Jedwab and Moradi emphasize the role strategic considerations played in determining the siting of colonial railroads as the imperial powers competed for different territories and sought to consolidate internal control.\textsuperscript{48} Yet according to their data, military domination was given as a motivation in only 35.5\% of the cases. The rest were linked to primary commodity extraction (mining and cash crop agriculture)\textsuperscript{49}—and between minerals and agricultural commodities, most colonies were dependent on the latter. (See Appendix B “Sources of Historical Primary Commodity Production by Country.”)

In line with an influential stream of scholarship in economic history,\textsuperscript{50} this underscores the importance of cash crop agriculture in structuring colonial state-building and contemporary spatial inequality. Yet, there have been few systematic studies

\textsuperscript{43}Huillery 2009. European settlements represented an important vehicle for local development as they drew greater levels of colonial investments in public goods or economic production. Huillery 2011.
\textsuperscript{44}Jedwab and Moradi 2016; Jedwab, Kerby, and Moradi 2017.
\textsuperscript{45}Chaves, Engerman, and Robinson 2014.
\textsuperscript{46}Jedwab and Moradi 2016.
\textsuperscript{47}Jedwab and Moradi 2016.
\textsuperscript{48}Jedwab and Moradi 2016.
\textsuperscript{49}See an earlier version of their paper Jedwab and Moradi 2012.
along this line of inquiry. In the next section, we develop our conceptual framework that explains how the spread of the cash crop revolution and colonial economic extraction shaped the modern African state before employing new, comprehensive data to estimate their effects vis-à-vis other factors.

**Conceptual Framework and Historical Context**

Our framework focuses on how the cash crop revolution transformed state-building and spatial patterns of development across Africa through the dynamic, interactive effects of ‘new’ geographic fundamentals, trade and the spread of imperial institutions.

**Geographic Change**

As discussed above, locational fundamentals have had a powerful effect on long-run development across the continent. The value of different geographic characteristics is not fixed, however; it can change as technology and market forces change. These economic shocks can interact with locational fundamentals to give rise to new centers of agglomeration. This is an important dimension in accounting for the spatial impact of the cash crop revolution in Africa from the 19th century onward.

As is well-known, prior to the 19th century the global cash crop trade had differential effects in Africa relative to other tropical regions, despite their similar ecological conditions. Whereas the Columbian Exchange spurred sugar production in the Caribbean and Latin America from the 16th century and opened European markets for other agricultural commodities, such as tobacco and cocoa, European imperial powers, led by Portugal, turned to Atlantic Africa as a labor reserve for plantations in the New World—leading to the enslavement of millions of people. The slave

---

54 Manning 1990.
trade largely crowded out cash crop production for three hundred years across the continent.\textsuperscript{55}

By the end of the 18th century, however, the African cash crop trade began to increase, starting with oil palm in West Africa, in response to rising European demand for the commodity for soap-making, candle-making and as an industrial lubricant.\textsuperscript{56} The abolition of the slave trade in the early 19th century accelerated West Africa’s economic transformation. Beyond oil palm, European demand for other oleaginous crops, such as groundnuts, as well as cotton, cocoa, coffee, tea, and tobacco fuelled the cash crop revolution that spread from West Africa to across the continent over the next 150 years.

Outside of a few large-scale irrigated projects, such as the Gezira cotton scheme in Sudan, agro-climatic suitability would condition the spread of cash crop agriculture. (See Figure 1 below). Groundnuts, cotton and tobacco, which require rainfall seasonality, clustered in tropical and subtropical grasslands, savannas, and shrublands, whereas oil palm, cocoa and robusta coffee, which need warmer temperatures, higher levels of rainfall and higher humidity, thrived in tropical and subtropical moist broadleaf forests. These ‘new’ geographic fundamentals would underlie the spatial distribution of the commercialization of agriculture across the continent.

Trade

A second important dimension shaping the spatial distribution of the cash crop revolution was the geography of markets and trade costs. Henderson et al. show that the globalization of markets and technological advances in long-distance trade over the last 150 years led to the concentration of economic activity in coastal regions of late-developing countries.\textsuperscript{57} With cash crop demand driven by industrializing states in Europe, commercial agricultural zones were integrated with export markets. Thus, in West Africa, conditional on suitability, cash crop cultivation tended to concen-

---

\textsuperscript{55}See Rodney 1972; Inikori 2014.
\textsuperscript{56}Hopkins 1973; Lynn 1997; Frankema, Williamson, and Woltjer 2018.
\textsuperscript{57}Henderson et al. 2018.
trate in areas closer to the coast first. (These areas also tended to be the first exposed to the diffusion of cash crops and new agricultural techniques.

Trade costs are not fixed either, however. Infrastructural investments could significantly lower them, opening up suitable areas in the hinterland for cash crop production. As noted above, in many cases demand for cash crops endogenously increased investments in transportation infrastructure.\(^5\) One of the most prominent examples was the British Cotton Growing Association’s (BCGA) lobbying of the British government for a railway line to Northern Nigeria which the BCGA identified as the “salvation of Lancashire” to reduce British cotton manufacturers’ dependence on American cotton.\(^6\) Despite strong local pressure from colonial authorities and their agents, farmers in Northern Nigeria spurned cotton cultivation for groundnuts, which gave a higher return, required less land and labor, and were consumed locally.\(^7\) Nonetheless, the railway fueled Northern Nigeria’s ascendancy as one of the leading groundnut production zones in the world.\(^8\)

This points to a corollary to Henderson et al’s elegant model of spatial development in late developing countries.\(^9\) The rise of commercial export agriculture and connective infrastructure potentially fuelled agglomeration in suitable areas beyond coastal regions.

**Institutions**

The emergence of new spatial equilibria and their developmental consequences are concurrently shaped by a third dimension—the institutions in which markets and

---

\(^5\) As one of the leading French traders in West Africa, Georges Borelli, put it in lobbying the French government to build a railway in Dahomey, it would help "syphon [sic] off immense quantities of products, up to the present time not used, towards the coast and towards Europe" (cited in Daumalin 2004).

\(^6\) Hyam 1968.


\(^8\) This vignette illustrates the top-down and bottom-up dynamics driving the cash crop revolution. Externally-financed infrastructural investments were transformative but the agency of local producers also were important. See Hogendorn 1969.

\(^9\) Henderson et al. 2018.
trade are embedded. In states where regimes do not possess an encompassing interest in society and power stems from monopolizing rents, elites are likely to structure the economy to enable their capture of surplus arising from economic production—likely diminishing broader societal benefits and increasing inequality. In such systems, extractive, enclave economies tend to flourish; production is organized around the harvesting and export of raw primary commodities (such as, minerals and agriculture) from a few geographically-concentrated zones that possess weak forward or backward linkages with other domestic sectors.

The link between extractive enclave economies and natural resources points to an endogenous relationship between institutions and factor endowments. In accounting for institutional variation in the Western Hemisphere, Engerman and Sokoloff posit that, whereas Latin America’s agro-climatic suitability for sugar gave rise to higher levels of political and economic inequality via sugar’s economies of scale and the legacies of slave-dependent plantation production, North America’s suitability for grain cultivation and production in small family farms contributed to greater economic equality and competitive markets.

In Africa, however, the historical mode of cash crop production diverged from Latin America. Given land abundance and labor scarcity, extensive agriculture proved more efficient than more labor- and capital-intensive plantations, which generally failed. Consequently, the locus of the economic revolution resided in smallholding farms—especially the farming that emerged organically in West Africa before colonialism.

Yet, despite this dispersion of economic production in the hands of family farms, Africa diverged from North America’s development trajectory. Two differences stand out. West African smallholder agricultural production was concentrated in com-

---

64 Hirschman 1977; Cardoso and Faletto 1979; Leonard and Straus 2003.
65 Auty 1997; Engerman and Sokoloff 1997; Sokoloff and Engerman 2000; Easterly and Levine 2003.
66 See also Easterly 2007 who tests this relationship more broadly using sugar and grain soil suitability.
68 Hogendorn 1969.
commodity exports destined for Europe, rather than local grain markets, which likely en-
genereed relatively thin domestic production linkages akin to tobacco and cotton in North America. Another important difference was the level of the European settler population. In low settler imperial outposts, extractive institutions prevailed—in which economic production was organized around the extraction of raw commodities for the metropole. Whereas conquistadors were the emissaries of extraction in 16th century Latin America, in 19th century West Africa it was European trading houses.

European traders leveraged the coercive and economic power of their home gov-
ernments to capture markets beyond the coastal ports where they were located during the slave trades. Instrumental in this regard was the use of, what can best be described as, gunboat commerce to eliminate indigenous middlemen, control mar-
kets, and drive down prices. By the Berlin Conference, French and British traders dominated commercial trade on every major river in the region and nearly all ports from Saint-Louis in Senegal to the Oil Rivers in Nigeria.

The spread of imperialism had a number of consequences. For one, as discussed, European capital financed new infrastructure, but railways and paved roads tended to vertically-integrate agricultural zones with export markets. Combined with protective trade policies (e.g., limiting imports to bulk, raw commodities), this further weakened domestic production linkages (as manufacturing was dislocated to Europe)

---

69 Baldwin 1956; North 1959; Price 1974.
70 Engerman and Sokoloff 1997 as well as Acemoglu, Johnson, and Robinson 2001, 2002 emphasize the effects of settlers on institutional development.
71 Acemoglu and Robinson 2012.
72 Young 1994.
73 See Santos 1970; Rodney 1972. Aiding European traders’ penetration of western Africa was the the advent of the steamship and increasing availability of quinine as a prophylaxis for malaria Headrick 2012.
74 With the abolition of the slave trade in 1807, the British Royal Navy started to patrol the coast of West Africa to interdict slave traders. British merchants exploited the presence of the Royal Navy to call on them to bombard villages that resisted their commercial activities on the Niger River and other waterways Headrick 2012. French traders did the same with the French Navy, which started to patrol West Africa after 1820 Daumalin 2004.
75 Rodney 1972.
and stymied the domestic economic differentiation and horizontal integration that otherwise might have resulted.  

A case in point is the protective trade measures France placed on groundnuts in the 1840s. France restricted groundnut imports to bulk unshelled nuts on French vessels, which aided French merchants’ monopolization of the West African commodity trade, while damaging local oil processing in the region. Moreover, buying oligopolies and monopolies, such as George Goldie’s United African Company (later the Royal Niger Company), dominated export markets, squeezing out African and non-European export firms.

The spread of imperial trade institutions paved the way for conquest and colonization, which in turn deepened the organization of African economies around primary commodity exports. However, the 19th century boom in cash crop production that in many ways propelled the imperial conquest led to declining terms of trade in agricultural commodities by the time of the Berlin Conference. Nonetheless, colonial states remained steadfast in promoting adoption of cash crop agriculture. But once a production area was opened to trade, there was a tendency to specialize in these zones. Severe budget constraints imposed by the metropole further forced colonial governments to pick and choose the agricultural schemes it could support. Beyond integrating these prized zones with export markets, investments in transportation infrastructure were scarce. Few horizontal networks were constructed to facilitate intra-African trade.

Another potentially important spatial implication of extractive colonialism was the creation of labor reserves, in which colonial governments used the coercive arm of the state—repression, high taxation rates and deprivation of local economic opportunities—to create a cheap supply of labor to service cash crop and mining enclaves and work

---

77 Brooks 1975.
78 This eventually took the form of marketing boards before and during World War II.
79 Following the Mineral Revolution in South Africa in the 1870s and 1880s, imperial conquest and state building in southern Africa revolved much more around mining—and were led by royally chartered companies, such as the British South African Company and Union Minière du Haut-Katanga.
81 Walle 2009; Gardner 2012.
82 Rodney 1972.
on colonial infrastructural projects.\textsuperscript{83} In effect this led to a policy of dividing their territories into productive and labor extraction zones.\textsuperscript{84}

\textit{***}

The cash crop revolution transformed Africa. How it did so was shaped by the interaction of geography, trade and imperial institutions. Three implications follow from our framework: 1.) we expect agro-climatic suitability for cash crops and historical cash crop production zones to be significant predictors of contemporary patterns of economic development across the continent; 2.) colonial infrastructure investments to reduce trade costs served as a key channel of path dependence; and 3.) Extractive institutions stymied the broader domestic economic gains that likely would have accrued and left non-cash crop areas worse off than expected based on their underlying characteristics, thereby raising motivations and opportunities for armed rebellion.

\textbf{Data and Methods}

\textbf{Measuring the spatial distribution of the cash crop revolution in Africa}

To test this argument, we draw on a novel and comprehensive dataset of historical African economic geography, including detailed geospatial information on cash crop production at the end of the 1950s across 38 countries. We rely on multiple sources of data. As explained above, one remarkable dataset was built by a team led by William Hance.\textsuperscript{85} In the late 1950s, Hance’s team mapped out the source location of more than

\textsuperscript{83}Amin 1972, 1974; Cordell, Gregory, and Piché 1996.

\textsuperscript{84}For example, in West Africa the French identified the Mossi people as “exploitable” due to its high population density in an “inhospitable land.” Quote by French doctor, Charles Crozat, who traveled to the Mossi Kingdom in 1890. Cited in Cordell, Gregory, and Piché 1996. The Mossi homeland falls in the bottom tercile of cash crop suitability. Subsequently, Upper Volta (Burkina Faso), which encompassed the Mossi people, was set up as a labor reserve.

\textsuperscript{85}Hance, Kotschar, and Peterec 1961.
95 percent of exports across 38 states in sub-Saharan Africa (excluding the Union of South Africa and most island colonies) and conveyed them as points at the site of cultivation or extraction. Each point represents a value of $289,270, standardized in 1957 U.S. dollars ($). The dataset covers 9 groups of cash crops, 20 minerals and metals, and forest, animal and manufactured products. In total, it identifies 9,517 geocoded production points (See the upper-right panel in Figure 1.) Published in the Geographical Review in 1961, the map represented a landmark contribution to the study of African economic geography. Yet it was generally overlooked.

Independently, we collected similar data as the Hance team. Drawing on colonial reports, maps and other records documenting the location, volume and value of primary commodity production across 30 countries (28 colonies and the states of Liberia and Ethiopia), the dataset provides administrative data on the subnational distribution of the most important commodity exports produced by each state (i.e., those that make up at least 10 percent of total exports at the end of colonialism) standardized in 1960 US$. We are thus able to validate the Hance dataset at the level of subnational administrative units. Overall and resource-specific correlations are high (see Figure A1).

**Unit of analysis, outcome variables and alternative pathways**

For our empirical analysis we aggregate the historical primary commodity data, the infrastructure and development outcomes as well as all control variables to 28,166 quarter-degree grid cells (the mean land area of cells in our sample is 237 sq km). In the Appendix (see Figures A17 and A18), we rerun many analyses at coarser geographic units aggregating our data to grid cells of 0.5 and 1.0 degree resolution as well as ethnic group polygons from Murdock’s “Tribal Map of Africa.” For each spatial unit, we code binary measures of whether cash crop or mineral production occurred above the threshold export value of $289,270 in 1957, as measured by Hance. These colonial cash crop and mineral dummies serve as the main predictors in our models.

---

86 For details, see Appendix B.
87 Murdock 1959.
and allow a quantitative comparison of the effects of different commodity types. In alternative specifications, we also use continuous operationalizations of our resource variables and take the \( \log (+1) \) of cell-level production value in $100 divided by land area in sq km (See Figures A14 and A15).

Consistent with the expectation that the cash crop revolution transformed the long-run spatial distribution of development within African countries, our main outcomes focus on contemporary infrastructure and economic agglomeration: paved and improved roads; electrification; cities; and household wealth.\(^{88}\) To measure these variables, we use easy-to-interpret dichotomous measures indicating if a cell had: (a) a quality road in 1998; (b) satellite-detectable luminosity at night in 2015; (c) a city with at least 10,000 inhabitants; and (d) the cell mean of asset-based household wealth from all available geocoded Demographic and Health Surveys (DHS) from the years 1990–2017. This measure is based on asset ownership of 747,255 households surveyed in 26 African countries. The household wealth analysis is based on the 24% subsample of our cells that contain DHS enumeration areas. Replacing binary outcomes with logged continuous measures (quality road length per 100 sq km, urban population per sq km, average cell-level luminosity, night lights per 1000 inhabitants) does not substantively alter our results.

To account for the potential confounding effects of biogeographic fundamentals and the legacy of the Neolithic Revolution, we control for calorie-weighted agricultural suitability, elevation and terrain ruggedness, as well as disease prevalence (malaria and tsetse fly). Alternative explanations of spatial inequality focus on: geographic variability and its effects on social fractionalization and ethnic inequality;\(^{89}\) the effects of the slave trades;\(^{90}\) variation in pre-colonial institutions;\(^{91}\) and technological changes that reduced shipping costs, agglomerating economic activities around ports and natural harbors.\(^{92}\) To account for these factors and others, we also

\(^{88}\)For sources of historical and contemporary development outcomes, all other variables and summary statistics, see Tables A1, A2, A3, and A4.

\(^{89}\)Alesina, Michalopoulos, and Papaioannou 2016.

\(^{90}\)Nunn 2008; Nunn and Puga 2010.

\(^{91}\)Boone 2003; Gennaioli and Rainer 2007; Michalopoulos and Papaioannou 2013; Wilfahrt 2018.

\(^{92}\)Henderson et al. 2018; Ricart-Huguet 2017.
control for: distance to the coast and navigable rivers, pre-colonial political institutions, slave affectedness, distance to historic trading and explorer routes, late 19th century urbanization, and proximity to the founding colonial capital.

**Statistical analysis**

We employ four different empirical strategies to estimate the long-term effects of colonial resource extraction on contemporary infrastructure and development outcomes: (i) simple OLS regressions with country fixed effects and a host of geographic and historical pre-treatment controls, (ii) inverse propensity score-weighted regressions, (iii) a spatial randomization inference approach comparing outcomes in historical resource areas to those in 1000 similarly plausible counterfactual spatial equilibria, and (iv) spatial instrumental variable models instrumenting colonial cash crop production with agro-climatic suitability scores while also accounting for spatial patterns in the outcome variables. Each of these methods has its advantages and limitations, yet consistent results across all of them increase the credibility of our findings.

**Baseline models.** Our baseline models take the following form:

\[
Y_{ic} = \alpha_c + \beta \text{ Resource Dummy}_{ic} + \lambda X_{ic} + \epsilon_{ic}
\]

\(Y_{ic}\) is outcome \(Y\) for cell \(i\) nested in former colony \(c\). The coefficient of interest is \(\beta\) identifying the effect of colonial cash crop or mineral production. The fixed effects \(\alpha_c\) control for all unobserved geographic and historical confounds at the level of colonies and enable an intuitive interpretation of coefficients as measures of spatial inequality (i.e. cell-level deviation from the country/colony mean). \(X_{ic}\) contains the geographic and pre-colonial baseline controls described above and \(\epsilon_{ic}\) is the error term. All models are estimated via OLS, which yields linear probability models for the binary outcomes (a)-(c).\(^93\) We cluster standard errors at the level of ethnic re-

\(^93\)We prefer the linear probability model over more traditional binary choice models due to its intuitive interpretation of coefficients as marginal effects as well as its versatility in terms of accom-
gions as defined in Murdock’s map\textsuperscript{94} representing the closest known albeit imperfect approximation of pre-colonial political entities available for the whole continent.\textsuperscript{95} In Appendix Tables A7 and A8, we report detailed regression tables with coefficient estimates and standard errors for all covariates.

**Propensity score weights.** As an alternative strategy to account for observable confounds, we use propensity score weighting. We first estimate a continent-wide propensity score model of receiving the cash crop (mineral) treatment based on all geographic and historic baseline covariates included in the OLS models as well as a cash crop-specific agroclimatic suitability score from the FAO GAEZ database.\textsuperscript{96} We then run OLS regressions with inverse propensity score weights, which under the (admittedly strong) assumption of no unobserved confounding yield the average treatment effect on the treated (ATT).\textsuperscript{97} Reweighting dramatically improves balance on observable pre-treatment characteristics (see Figure A2). For all 19 covariates across two treatments (cash crops and minerals) and two samples (global vs. DHS subsample), the standardized mean differences between treated and control observations remain below the conventional threshold of 0.1.

**Randomization inference: counterfactual spatial equilibria.** We further proceed with a randomization inference-type analysis inspired by Dell and Olken.\textsuperscript{98} The basic intuition is that the observed spatial distribution of colonial resource production is just one realization of many possible spatial equilibria. According to our data, only 12\% of grid cells with above-median cash crop suitability saw actual production in 1957. In addition to raw agro-climatic suitability, a host of other geographic and historical variables may have determined the location of resource extraction.
tion (e.g. proximity to pre-colonial trade routes, navigable rivers, or the coast). To account for this, we construct a set of 1000 similarly plausible counterfactual distributions of both cash crop and mineral extraction. We regard alternative spatial resource distributions as plausible if (i) counterfactual cells score similarly on geographic and historical baseline variables as real cash crop (mining) cells, and (ii) the overall distribution of counterfactual cells exhibit comparable patterns of spatial clustering as the actually observed resource data. The local determinants of colonial resource extraction may vary across different crops and minerals as well as geographic and institutional contexts, especially in the case of decentralized cash crop production by native farmers. However, soil suitability, proximity to natural and pre-existing trade networks, terrain, the disease environment, and pre-colonial development plausibly mattered in most contexts. In addition, the counterfactual analysis enables us to gauge the latent developmental advantage implied by the non-random assignment to resource extraction and, via the imposed spatial patterns, safeguards against mistaking spatial clustering in predictors and outcomes for long-term effects.99

We construct counterfactual resource distributions as follows: We treat the realized distributions of colonial cash crop and mineral production as clustered point processes and estimate two separate models assuming a Thomas cluster process and fitting a spatial trend that depends on the geographic and historical covariates described above. Using a clustered point process model seems appropriate since the distribution of colonial cash crop and mineral extraction exhibits stronger spatial clustering than predicted by covariates alone. We use the estimated parameters from the point process models to simulate 1000 counterfactual spatial equilibria. We then aggregate the points of each simulation to our quarter-degree grid and regress our outcome variables on a counterfactual cash crop (mineral) dummy without adding any additional control variables or fixed effects. We run the same models with the observed resource variables and compare the actual coefficients with the distribution of 1000 counterfactual estimates. We derive point estimates of the treatment effect by subtracting the mean of all counterfactual coefficients from the actual estimate.

99Kelly 2019.
Empirical p-values are calculated by looking at the position of the actual coefficient within (or outside) the absolute value distribution of counterfactual coefficients.\(^{100}\) The top panel of Figure 2 contains all point estimates and empirical p-values, whereas the bottom two panels plot actual coefficients and the distribution of their counterfactual counterparts.

To assess whether our simulation exercise achieves covariate balance, we also regress all geographic and historical baseline covariates on both counterfactual and actual resource variables. Figures A4 and A5 show the coefficient distributions and indicate balance across all observed covariates. Figure A3 maps out the spatial distribution of our 1000 counterfactual draws and visually compares them to the actual distribution of colonial resource extraction.

**Spatial instrumental variable analysis.** All these strategies are based on observables only and may still yield biased estimates if unobserved confounds varying within countries make an area more likely to produce cash crops and simultaneously lead to more infrastructure investment and economic activity. We address this inferential threat by estimating 2SLS-IV models instrumenting observed cash crop production with the mean agro-climatic suitability score across the nine most important African cash crops (cocoa, coffee, cotton, groundnuts, oil palm, tea, sugarcane, tobacco, and bananas). Cash crop suitability is a valid instrument if it predicts actual production and at the same time does not influence infrastructure and subnational development through any other causal pathway. We argue that this exclusion restriction plausibly holds since soil suitability for cash crops is not affected by economic activity and, conditional on our baseline covariates—especially general, non-cash-crop-related agricultural suitability and ethnic group-level agricultural dependence prior to colonization—fertile soils for cash crops are unlikely to affect urbanization rates, state building, infrastructure, and economic activity except through their impact on cash crop production from the colonial era onwards.

\(^{100}\) More formally, the empirical p-values are calculated as \(\frac{r+1}{n+1}\), where \(r\) denotes the number of counterfactual coefficients with greater absolute value than the real coefficient and \(n\) refers to the number of simulations.
In the Supplementary Materials (Figure A8), we conduct placebo tests to probe the plausibility of this assumption. We regress cell-level and ethnic group-level proxies of pre-colonial development on cash crop suitability and the set of geographic baseline variables discussed above. Reassuringly, cash crop suitability is not significantly related to proximity to pre-colonial trade routes and urban centers, town size, social class stratification, political centralization, democratic selection of early elites, and the use of slavery. However, similar spatial patterns in both our instrument and outcomes pose an additional threat to the exclusion restriction. We use a continuous, spatially interpolated and therefore relatively smoothly varying instrument for a sharply varying binary treatment constructed from point data, whereas our outcomes tend to vary less sharply between cells than the colonial resource variables (see Figure 1). Under such conditions, standard IV methods may yield estimates that are even more biased than simple OLS.

To account for spatial dependence in our data, we therefore estimate spatial 2SLS models that include a spatial lag of the dependent variable instrumented with first- and second-order spatial lags of the exogenous baseline controls. We do not have strong theoretical priors as to whether the data generating process in fact follows a spatial lag model. Spatial correlation in our outcome variables may also be due to spatially clustered unobservables or spillover effects of right-hand side variables—a possibility we explore in more detail below. Instead, we choose spatial lag 2SLS models as the most practical method to account for spatial dependence while at the same time instrumenting for an endogenous treatment variable. In our main results, we report estimates from models that use spatial lags based on an inverse-distance weighted nearest neighbor matrix including the 24 most proximate cells. The robustness analysis in the Appendix shows that alternative choices of the spatial weights matrix do not fundamentally change our results (See Figure A7). As for instrument

101 Murdock 1967. There is a statistically significant but substantively small association between cash crop suitability and distance to European explorer routes indicating that, if anything, highly suitable areas are further away from these routes.

102 Betz, Cook, and Hollenbach 2019.

103 Betz, Cook, and Hollenbach 2019.

104 Betz, Cook, and Hollenbach 2019.
strength, cash crop suitability strongly predicts actual production in our first-stage regressions. The relevant first-stage F statistic is 34.2 in the road, city, and night light models and 11.0 in the DHS household wealth analysis. We do not instrument colonial mineral production since we are not aware of a plausibly exogenous predictor of colonial mining activities.

Results

The four maps in Figure 1 visually present the central argument and empirical approach: cash crop suitability (upper-left panel), at least partially shaped the distribution of historical cash crop production (upper-right panel), which in turn structured colonial infrastructure investments and the development of export-oriented cash crop enclaves (lower-left panel) that have wrought severe and persistent subnational variation in economic development (lower-right panel). Figure 2 reports coefficient plots derived from the four different statistical methods that allow us to address concerns about endogeneity and more credibly estimate the effect of historical cash crop cultivation on contemporary development outcomes. Our baseline OLS models indicate that colonial cash crop cells had a 16 percentage points higher probability of having a quality road in 1998; close to a 20 percentage points higher likelihood of emitting nighttime lights in 2015; a 19 percentage points higher likelihood of having a city in 2015; and 14 percent of a standard deviation greater household asset-based wealth.\(^\text{105}\) Historical cash crop production exhibits a comparable effect on contemporary roads, electrification and cities as colonial mineral extraction, despite the more capital-intensive nature of mining. The results from the propensity score-weighted regressions, randomization inference and instrumental variable models align closely with the baseline OLS coefficients, suggesting that our estimates are unlikely to be driven by omitted variables, reverse causation or measurement error. The S2SLS-IV results are robust to alternative definitions of what constitutes the

\(^{105}\) See Tables A7 and A8 for OLS regression results with all controls.
spatial neighborhood of a cell (Figure A7).

**Comparing the effects of cash crop production with plausible counterfactual sites**

In addition to revealing alternative spatial equilibria that remained largely unrealized during colonialism (see Figure A3), the counterfactual analysis allows us to benchmark the latent developmental effects of the non-random siting of cash crop production (e.g., due to underlying agricultural potential or proximity to historical urban centers or the coast) and then more precisely estimate the impact of actual production. As illustrated in the lower panel in Figure 2, locational fundamentals and historical processes appear to have predisposed counterfactual cells to higher levels of development than the average cell, but colonial cash crop production substantially amplified this effect by 136 to 196% depending on the outcome.

**Robustness and heterogeneity**

Overall the effect sizes are fairly consistent across the major five cash crops, with the exception of household wealth, in which coffee, cocoa and oil palm outperform cotton and groundnuts (See Figure A9). Country-by-country regressions reported in Figures A10 and A11 show the results hold for most countries and are not driven by outliers; they also generally persist across different imperial powers but are slightly higher in former British colonies (See Figure A12). Nor have path dependent effects been disrupted by different post-colonial trajectories; results are stable across different levels of democracy, conflict-affectedness, and resource dependence; as well as among landlocked and coastal countries (See Figure A13). The Online Appendix to this paper contains additional robustness checks: estimating cash crop effects at the intensive margin (Figure A14); using continuous instead of dichotomous measures of colonial cash crops for the IV analysis (Figure A15); employing Conley standard errors with varying cutoffs to account for spatially correlated residuals (Figure A16); using different grid cell resolutions and ethnic groups in Murdock as unit of analysis.
Figure 1: Cash crop suitability, colonial cash crop production, infrastructure and subnational development in Africa.

The blue/green shading in the upper panels shows agro-climatic suitability for cash crops. Each orange point represents US $289,270 export value of cash crop production in 1957. Red crosses represent mining sites producing varying export volumes. The lower-left panel illustrates road and railroad infrastructure around 1960. The lower-right panel overlays the colonial resource data with luminosity at night in 2015 as a proxy for economic activity.
Figure 2: The effects of colonial cash crop production on long-run development in Africa.

**Upper Panel.** Estimated effects of colonial cash crop and mineral production on four contemporary infrastructure and development outcomes. In each panel, the first and fifth estimates are based on OLS models with country fixed effects; the second and sixth on propensity-score weighted regressions; the third and seventh on our spatial randomization approach; and the fourth on spatial 2SLS models instrumenting colonial cash crop production with agro-climatic suitability scores. **Lower Panel.** Actual versus counterfactual coefficients from the randomization inference analysis.
(Figures A17 and A18); and comparing the effects of cash crops on long-run development to the slave trades and pre-colonial centralization (Figure A19).

**Mechanisms: serial correlation versus path dependence**

The results raise the question of what accounts for the persistent effects of the cash crop revolution on spatial inequality across Africa? We test for two potential pathways: serial correlation and path dependence.\(^{106}\) Serial correlation links contemporary patterns of economic development to recurring direct effects of locational fundamentals (i.e., the geographic or environmental conditions that spurred the concentration of economic activities in the past exert a similar influence in the present).\(^{107}\) Accordingly, high levels of contemporary development in cash crop zones would be a function of the continuous agricultural production in these highly suitable areas. In contrast, path dependence points to the increasing returns that arise not necessarily from underlying drivers of initial economic activities but from the clustering of capital and infrastructural investments in a given area.\(^{108}\) Prior investments not only lower costs going forward but also serve as a coordination mechanism for subsequent distribution of factors of production.\(^{109}\) Consistent with the pattern visualized in the lower-left panel in Figure 1, we report in Tables A9 and A10 evidence that the historical transition to cash crop agriculture aligned with colonial investments in infrastructure—not just in transportation networks, but also in power generation (based on the digitization of a comprehensive map of power stations produced by the United Nations Economic Commission for Africa from the late 1960s).

We perform a causal mediation analysis to investigate the impact of serial correlation versus path dependence. We operationalize the serial correlation mechanism by using an FAO estimate of the market value of total agricultural production across all crops in the year 2000. To test for path dependence, we use data on early inde-

---


\(^{107}\) Davis and Weinstein 2002.


\(^{109}\) Jedwab, Kerby, and Moradi 2017.
pendence roads, railroads, and electricity generation facilities. For all three of these infrastructure variables, we code dummies indicating whether a cell is crossed by a colonial road, rail, or hosted a power plant in the early post-independence period. As for the railroads, we only use those that were constructed for other than mining or military purposes to make sure that the potential for cash crop exports may have, at least partially, motivated rail construction.

We use sequential g-estimation as a method to test for the relevance of these two mechanisms.110 This two-step estimation procedure allows to estimate average controlled direct effects (ACDE) showing how much of the main effect remains after taking a particular mechanism into account. Valid estimation of the ACDE relies on the assumption of sequential unconfoundedness, i.e. no omitted variables biasing the effect of the treatment and/or mediator on the outcome. To address potential violations of this assumption, we use cash crop suitability instead of actual production in 1957 as the reduced form treatment variable. In addition, we add all geographic covariates from our baseline OLS models as potential pre-treatment confounders and all historical covariates as potential intermediate confounders.

Figure 3 plots the results from our mediation analysis using cash crop suitability as treatment variable. The suitability effects remain larger if we demediate the development outcomes with contemporary crop production than with the infrastructure factor. The difference between the baseline coefficient and the ACDE can be interpreted as the part of the total effect that is mediated by the respective causal pathways: 30-38% of the cash crop suitability effect on contemporary development, depending on outcome, is due to early infrastructure, whereas only 16% – 22% of these effects is mediated by a continued advantage in agricultural productivity. The infrastructure channel is even more important if we use actual cash crop production in the mediation analysis (See Figure A20). Here, we report the more conservative demediated analysis with suitability due to potential confounding. Since most highly suitable cells never saw actual production, the true effect probably lies somewhere in between. Nonetheless, both results suggest that path dependence more than serial

---

correlation accounts for the long-run developmental effects of the cash crop revolution.

Figure 3: Causal Mediation Analysis.

Total effect of cash crop suitability as well as average controlled direct effects demediated with respect to contemporary agricultural productivity and early independence infrastructure (rails, roads, power plants). Throughout the infrastructure channel appears more important than agricultural productivity.

**Benchmarking the effects of cash crops on long-run development versus other geographic variables**

To better grasp the importance of our findings, we compare the effects of agroclimatic cash crop potential on patterns of long-run development to other geographic variables—conditions favorable for food production to sustain dense populations; disease ecology; distance to coast; distance to navigable rivers; elevation; and ruggedness. The analysis is at the cell-level—and enables us to compare relative effects of cash crop suitability to other geographic factors regularly cited in the literature as key determinants of development patterns in Africa.

All variables are standardized to mean 0 and sd 1 to compare effect sizes. Results, reported in the lower panel of Figure 3, are at the continental-level (pooled) and within modern international borders (using fixed effects). Overall the results
confirm the impact of “the daunting nature of Africa’s geography,”\textsuperscript{111} in particular the formidable costs to trade for the region's vast interior. Distance to the coast exhibits the largest effects on contemporary development. Next to coastal distance, the effects of cash crop suitability rival, and in many cases, surpass other important geographic factors, especially if we focus on within-country variation. Results remain substantively the same if we use continuous instead of binary outcome measures (See Figure A21). One takeaway from this analysis is that the prospective yield from the 19th century onward of agricultural crops of high commercial (and limited nutritional) value was at least as consequential in shaping subnational wealth disparities in Africa as the Neolithic potential to feed dense populations (Compare rows 1 and 2 in Figure 4 to rows 3 and 4).

\textsuperscript{111}Herbst 2000.

---

\[\text{Figure 4: Benchmarking of coefficient sizes: Cash Crop Suitability vs. Other Locational Fundamentals.}\]
Spillover effects, country-level inequality, and conflict legacies

In the final part of our analysis, we move beyond purely local effects of colonial cash crop production. So far, our estimates show large effects on patterns of subnational development: ‘treated’ cells continue to be significantly better off than comparable cells within the same country. The broader implications of these findings depend on whether and how colonial resource extraction affected development outcomes beyond the immediate neighborhood. Large-scale agglomeration patterns, as well as forward and backward linkages to other sectors of the economy, may have spurred even more positive effects than our local estimates suggest. As discussed, however, resource-exporting African countries are often characterized as enclave economies with limited potential for positive spillovers. Worse yet, the positive estimates we find may in part be due to adverse effects on resource-poor areas. Such negative spillovers seem particularly likely under colonial institutions where extractive policies, such as excessive taxation, labor coercion, restricted mobility, and the protection of European economic interests, were common practice.

To gauge the effects of colonial resource extraction beyond the very local level, we run models including variables for proximity to colonial cash crop and mineral production. More specifically, we construct distance band dummy variables coding, for each grid cell, whether it falls within 0-25, 25-50 ... 225-250 km of a colonial resource point. We then re-estimate our baseline models replacing cell-level resource indicators with 20 distance band dummies capturing proximity to both cash crop and mineral extraction. More formally, we run the following specification:

$$Y_{ic} = \alpha_c + \sum_{j=0-25}^{225-250} \beta_j \text{Dist. Crops}^j_{ic} + \sum_{j=0-25}^{225-250} \gamma_j \text{Dist. Minerals}^j_{ic} + \lambda X_{ic} + \epsilon_{ic}$$

The vectors $\beta_j$ and $\gamma_j$ contain the coefficient estimates for all distance band dummies to cash crops and minerals. Cells further away than 250 km from colonial re-
source extraction serve as the baseline category. $X_c$ and $\alpha_c$ refer to the same outcomes, control variables and country fixed effects used in our baseline analysis.

Positive coefficients on distance band indicators beyond 25 km would suggest positive spillovers. Negative coefficients would be consistent with adverse spillover effects making cells close to colonial resource extraction worse off than predicted by their baseline covariates.

Figure 5 presents coefficients for distance to cash crops. (See Figure A22 for spillover analysis with continuous outcomes). Across all outcomes, we find positive and significant spillovers up to 50 km. For nighttime luminosity and urbanization, we find negative and mostly significant spillovers between 75 and 250 km from colonial cash crop sites. These results suggest that concentrated investments in colonial cash crop enclaves crowded out development in other areas, which appear worse off today than predicted by geographic conditions and precolonial factors. The effects on household wealth are neutral beyond 50 km, but note that only 24% of our grid cells contain DHS enumeration areas. The more positive findings for road infrastructure are unsurprising, given that paved and improved road segments typically cut through multiple cells and extensions of the road network mostly connect pre-existing segments. Interestingly, the distance to mining coefficients exhibits positive spillovers over a somewhat broader range and never turns negative and significant (See Figure A23).

![Spillover Analysis: Cash Crops](image)

**Figure 5:** Spillover effects of Colonial Cash Crop Production
Limited or even negative spillover effects of geographically concentrated colonial resource enclaves suggest a legacy of severe spatial inequality. A descriptive analysis of country-level aggregate inequality measures confirms this impression. We calculate Gini coefficients of night light inequality in 2015 for each country in our sample (based on the distribution of night light intensity in 2015 across quarter degree grid cells). Figure 6 plots this measure of contemporary spatial inequality across an analogously constructed Gini coefficient of spatial inequality in colonial resource extraction (based on the cell-level sum of production value in 1957 as coded by Hance). Figure 6 shows a strong and positive correlation between the geographic concentration of historical resource extraction and contemporary economic inequality. In combination with the spillover analysis, these associations suggest colonial extractive economies not only shaped local patterns of development but produced macro-level consequences as well that continue to underlie African economic geography.

Figure 6: Colonial Resource and Contemporary Night Light Inequality at the Country Level
Figure 7: Colonial Resource Endowments and Post-Independence Ethnic Rebellion
In a final analysis, we investigate whether the spatial distribution of colonial resources helps to account for patterns of post-colonial ethnic conflict. Unequal economic opportunities create potential for distributive conflict that may, in extreme cases, result in political violence. Where economic inequalities align with politically salient ethnic group identities, widespread grievances among relatively poor groups motivate armed mobilization against the state.\textsuperscript{112} In addition, economic disadvantages and lower levels of historically inherited local state capacity provide armed actors with opportunities for recruitment and mobilization in ethnic settlement areas without colonial resources.\textsuperscript{113}

We use ethnic group-level data on settlement areas and rebellion from the Ethnic Power Relations (EPR) dataset to test this notion.\textsuperscript{114} For all 36 countries in our sample, we identify the ethnic groups that EPR codes as political relevant in the first year of independence. We then aggregate our colonial resource data to the polygons of these 150 groups’ main settlement areas and code, for each group, whether it has been involved in an ethnic civil war between independence and 2017. 35.7\% of ethnic groups that did not see any colonial cash crop production have fought at least one civil war post-independence, whereas the number for groups with colonial cash crops is only 24.5\%. The respective conflict shares for groups with and without colonial mineral extraction are 24.1\% and 28.7\%. Figure 7 presents estimates from three linear probability models with country fixed effects, the geographic and historical control variables from above, and Conley errors with a distance cutoff of 400 km. Colonial cash crop production is associated with a significantly lower likelihood of ethnic rebellion, regardless of whether we operationalize resource endowments as binary indicators or logged 1960 production values per sqkm or capita. The estimates for mineral resource endowments remain close to zero and statistically insignificant.

\textsuperscript{112}Stewart 2008; Cederman, Gleditsch, and Buhaug 2013.
\textsuperscript{113}Collier and Hoeffler 2004; Fearon and Laitin 2003.
\textsuperscript{114}Vogt et al. 2015. EPR provides lists of politically relevant ethnic groups for all independent countries since 1945 and codes information about these groups’ political representation and involvement in intrastate conflict as defined by the UCDP/PRIO dataset Gleditsch et al. 2002. The GeoEPR companion data set provides spatial polygon data on the main settlement areas of relevant ethnic groups. Wucherpfennig et al. 2011.
This null finding may have to do with less widespread economic benefits of mineral extraction or the attractiveness of capital-intensive fixed point resources as a target for armed actors. Overall, however, our conflict analysis suggests that the legacies of colonial resource extraction are not merely economic but, at least in the case of cash crops, extend to political outcomes such as stability and conflict.

Discussion

We argue Africa’s contemporary economic geography was shaped by the cash crop revolution that swept the continent from the late 1700s onward. The end of the slave trades and high global demand for agricultural commodities increased the geographical advantages of areas suitable for cash crop cultivation—leading to the emergence of new spatial equilibria centering around highly-suitable regions. However, the take-off of the cash crop trade with industrializing Europe also contributed to the diffusion of economic imperialism beyond coastal Africa. While European capital fueled the construction of new transportation infrastructure that further intensified the region’s agricultural revolution, extractive and mercantilist institutions structured the distribution of gains arising from this new mode of production and its broader impact on the economy. Consequently, as the colonial cash crop revolution produced powerful and enduring agglomerating effects, these came at the expense of surrounding areas, which ended up significantly worse off than otherwise would be expected based on their underlying characteristics.

An important scope condition of our analysis, however, is, given the extent of colonialism in Africa, it is difficult to know the counterfactual of the long-run effects of commercial export agriculture without colonial institutions and investments. As one of the few exceptions, Ethiopia is potentially instructive. Coffee exports took off in Ethiopia in the first part of the 20th century, spurring the integration of the southern highlands into the Ethiopian Empire through infrastructural and administrative investments. In turn, revenue from the coffee trade contributed to state centraliza-

\[^{115}\text{Lujala 2010; Dube and Vargas 2013; Berman et al. 2017.}\]
Interestingly, despite coffee’s economic importance, we do not see evidence of significant spatial disparities between Ethiopia’s cash crop agricultural zones and the rest of the state (See Figure A10). Whether this is due to the absence of mercantilist institutions and stronger integrative economic processes, a longer history of centralized statehood, or the (downward) leveling effects of the Marxist revolution that brought an end to the Ethiopian Empire in the 1970s requires future research.

Going beyond this single case, another fruitful avenue for research would be to scale up to a global sample to leverage greater variation in the distribution of colonial institutions and commercial export agriculture. Surprisingly, despite cash crop agriculture’s importance to the emergence of early modern globalization from the 16th century onward, its impact on spatial patterns of development around the world has largely been understudied. If Africa is any indication, this represents a significant omission in our study of the subnational wealth and inequality of nations.

---

116McClellan 1980.
References


Achten, Sandra, and Beata Smarzynska Javorcik. 2019. “Spatial inequality, geography and economic activity.”


Online Appendix (not for publication)

Table of Contents

Table A1. Description and Sources of Variables

Table A2. Description and Sources of Variables (continued).

Table A3. Summary Statistics (0.25 Degree Grid Cells)

Table A4. Summary Statistics (Ethnic Homeland-Country Pairs)

Table A5. Night-light Gini Coefficients by World Region

Table A6. Agriculture, Grains and Early Statehood?

Table A7. Baseline OLS Models (Cash Crops)

Table A8. Baseline OLS Models (Minerals)

Table A9. Cash Crop Suitability and Mediators

Table A10. Colonial Cash Crop Production and Mediators

Figure A1. Validation of Colonial Resource Data

Figure A2. Covariate Balance Propensity Score Weights

Figure A3 Distribution of Counterfactual Cells for RI Analysis

Figure A4. Covariate Balance: Counterfactual vs. Actual Cash Crop Cells

Figure A5. Covariate Balance: Counterfactual vs. Actual Mining Cells

Figure A6. RI Treatment Effects with Continuous Outcomes

Figure A7. Robustness of S2SLS-IV results

Figure A8. Placebo Test: Cash Crop Suitability and Pre-Colonial Development
Figure A9. Dissagregating Cash Crops

Figure A10. Country-by-Country Analysis: Cash Crops

Figure A11. Country-by-Country Analysis: Minerals

Figure A12. Heterogeneity across Imperial Powers

Figure A13. Heterogeneity Analysis

Figure A14. Robustness: Continuous Measures & Intensive Margins

Figure A15. Robustness: Continuous Measures & S2SLS Models

Figure A16. Robustness: Conley Errors for Spatial Correlation

Figure A17. Robustness: Changing the Spatial Unit of Analysis (Binary Predictors)

Figure A18. Robustness: Changing the Spatial Unit of Analysis (Continuous Pred.)

Figure A19. Cash Crops, Precolonial Statehood, and the Slave Trades.

Figure A20. Causal Mediation Analysis (Actual Production 1957).

Figure A21. Robustness: Suitability vs. Other Geographic Fundamentals

Figure A22. Spillover Analysis (Cash Crops)

Figure A23. Spillover Analysis (Minerals)

Appendix B. Sources of Historical Primary Commodity Production by Country
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality Road 1998</td>
<td>Any paved or improved roads within the grid cell</td>
<td>Michelin (1998) from Jedwab and Moradi (2016)</td>
</tr>
<tr>
<td>Night Lights 2015</td>
<td>Any detectable night-time lights within the grid cell</td>
<td>Román et al. (2018)</td>
</tr>
<tr>
<td>City 2015</td>
<td>Localities with 10,000 inhabitants or more in 2015</td>
<td>Africapolis.org</td>
</tr>
<tr>
<td>DHS Household Wealth</td>
<td>Composite asset-based measure of a household's living standard from 93 surveys across 26 countries</td>
<td>DHS (1990-2017)</td>
</tr>
<tr>
<td>Road Density 1998 (log)</td>
<td>Natural log of the length of paved or improved roads per square kilometer</td>
<td>Michelin (1998) from Jedwab and Moradi (2016)</td>
</tr>
<tr>
<td>Urban Population Density 2015 (log)</td>
<td>sum of population residing in cities with more than 10'000 inhabitants divided by land area in sqkm</td>
<td>Africapolis.org</td>
</tr>
<tr>
<td>Power Plant 1972</td>
<td>Location of hydropower and thermal plants</td>
<td>Economic Commission for Africa (1972)</td>
</tr>
<tr>
<td>Crop Production Value 2000 (log)</td>
<td>Grid-level estimate of the market value of all agricultural production</td>
<td>FAO GAEZ</td>
</tr>
<tr>
<td>Colonial Cash Crops (Y/N)</td>
<td>Production points greater than $289,270 (1957) across nine different cash crops</td>
<td>Hance, Kotschar, and Peterec (1961)</td>
</tr>
<tr>
<td>Colonial Cash Crop Value (log)</td>
<td>Log of the total value of cash crops produced divided by land area in sqkm</td>
<td>Hance, Kotschar, and Peterec (1961)</td>
</tr>
<tr>
<td>Colonial Mineral Value (log)</td>
<td>Log of the total value of cash crops produced divided by land area in sqkm</td>
<td>Hance, Kotschar, and Peterec (1961)</td>
</tr>
<tr>
<td>Cash Crop Suitability</td>
<td>Mean agro-climatic suitability across nine cash crops (cocoa, coffee, cotton, groundnuts, oil palm, tobacco, tea, sugarcane, and bananas) in a given area</td>
<td>FAO GAEZ</td>
</tr>
<tr>
<td>Caloric Suitability</td>
<td>Potentially attainable caloric yield (post-1500) across all crops suitable for cultivation in a given area</td>
<td>Galor and Özak (2016)</td>
</tr>
<tr>
<td>TseTse Suitability</td>
<td>Temperature-based tsetse fly suitability index calculated using the formula in Alsan (2015) and historical temperature data (1901-2017) from the CRU dataset</td>
<td>Alsan (2014); Harris et al. (2014)</td>
</tr>
<tr>
<td>Malaria Suitability</td>
<td>Temperature-based suitability index for malaria transmission</td>
<td>Weiss et al. (2014)</td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
<td>Source</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>Terrain Ruggedness</td>
<td>Average terrain ruggedness within quarter-degree cell or polygon</td>
<td>Shaver, Carter, and Shawa (2019)</td>
</tr>
<tr>
<td>Elevation</td>
<td>Mean elevation in m above sea level within cell or polygon</td>
<td>srtm.csi.cgiar.org</td>
</tr>
<tr>
<td>Dist. Coast (log)</td>
<td>Minimum geodesic distance to the coast</td>
<td>ngdc.noaa.gov/mgg/shorelines/</td>
</tr>
<tr>
<td>Dist. European Explorer Route (log)</td>
<td>Minimum geodesic distance to European explorer routes</td>
<td>Century Atlas Africa (1911) via Nunn and Wantchekon (2011)</td>
</tr>
<tr>
<td>Dist. 1900 City (log)</td>
<td>Minimum geodesic distance to localities with more than 10,000 inhabitants in 1900</td>
<td>Jedwab and Moradi (2016)</td>
</tr>
<tr>
<td>Dist. Colonial Capital (log)</td>
<td>Minimum geodesic distance to the colonial capital in 1900</td>
<td>Various</td>
</tr>
<tr>
<td>Dist. Precolonial Trade Route (log)</td>
<td>Minimum geodesic distance to precolonial trade routes</td>
<td>Ajayi and Crowder (1985)</td>
</tr>
<tr>
<td>Slaves per Area (int.)</td>
<td>Number of slaves taken from a Murdock ethnic group per sqkm (second tercile)</td>
<td>Nunn and Wantchekon (2011)</td>
</tr>
<tr>
<td>Slaves per Area (high)</td>
<td>Number of slaves taken from a Murdock ethnic group per sqkm (third tercile)</td>
<td>Nunn and Wantchekon (2011)</td>
</tr>
<tr>
<td>Slaves per Area (missing)</td>
<td>Indicates if slave trade data is missing</td>
<td>Nunn and Wantchekon (2011)</td>
</tr>
<tr>
<td>Precolonial Agriculture (int.)</td>
<td>Between 46 and 55% reliance on agriculture</td>
<td>Murdock (1967)</td>
</tr>
<tr>
<td>Precolonial Agriculture (high)</td>
<td>More than 55% reliance on agriculture</td>
<td>Murdock (1967)</td>
</tr>
<tr>
<td>Precolonial Agriculture (missing)</td>
<td>Indicates if precolonial agriculture variable is missing</td>
<td>Murdock (1967)</td>
</tr>
<tr>
<td>Precolonial Chiefdoms (Y/N)</td>
<td>Indicates if ethnic group had some political centralization beyond the village level</td>
<td>Murdock (1967)</td>
</tr>
<tr>
<td>Precolonial Statehood (Y/N)</td>
<td>Indicates if ethnic group had state-like political centralization</td>
<td>Murdock (1967)</td>
</tr>
<tr>
<td>Precolonial Centralization (missing)</td>
<td>Indicates if political centralization variable is missing</td>
<td>Murdock (1967)</td>
</tr>
<tr>
<td>Precolonial Town Size</td>
<td>Average town/settlement size of Murdock ethnic groups</td>
<td>Murdock (1967)</td>
</tr>
<tr>
<td>Social Class</td>
<td>Pre-colonial presence of class stratification</td>
<td>Murdock (1967)</td>
</tr>
<tr>
<td>Headman Elected</td>
<td>Democratic selection rule for local elites</td>
<td>Murdock (1967)</td>
</tr>
<tr>
<td>Slavery</td>
<td>Slavery as a common practice</td>
<td>Murdock (1967)</td>
</tr>
<tr>
<td>Statistic</td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>---------------------------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>Quality Road 1998 (Y/N)</td>
<td>28166</td>
<td>0.334</td>
</tr>
<tr>
<td>Night Lights 2015 (Y/N)</td>
<td>28166</td>
<td>0.267</td>
</tr>
<tr>
<td>City 2015 (Y/N)</td>
<td>28166</td>
<td>0.128</td>
</tr>
<tr>
<td>DHS Household Wealth</td>
<td>6775</td>
<td>−0.391</td>
</tr>
<tr>
<td>Road Density 1998 (log)</td>
<td>28166</td>
<td>−2.676</td>
</tr>
<tr>
<td>Mean Night Lights 2015 (log)</td>
<td>28166</td>
<td>−3.337</td>
</tr>
<tr>
<td>Night Lights per cap. 2015 (log)</td>
<td>28166</td>
<td>−3.490</td>
</tr>
<tr>
<td>Quality Road 1960 (Y/N)</td>
<td>28166</td>
<td>0.107</td>
</tr>
<tr>
<td>Railway 1960 (Y/N)</td>
<td>28166</td>
<td>0.018</td>
</tr>
<tr>
<td>Power Plant 1972 (Y/N)</td>
<td>28166</td>
<td>0.027</td>
</tr>
<tr>
<td>Crop Production Value 2000 (log)</td>
<td>18636</td>
<td>6.841</td>
</tr>
<tr>
<td>Colonial Cash Crop Dummy</td>
<td>28166</td>
<td>0.081</td>
</tr>
<tr>
<td>Colonial Mineral Dummy</td>
<td>28166</td>
<td>0.009</td>
</tr>
<tr>
<td>Colonial Cash Crop Value (log)</td>
<td>28166</td>
<td>0.175</td>
</tr>
<tr>
<td>Colonial Mineral Value (log)</td>
<td>28166</td>
<td>0.022</td>
</tr>
<tr>
<td>Cash Crop Suitability</td>
<td>28166</td>
<td>0.000</td>
</tr>
<tr>
<td>Caloric Suitability</td>
<td>28166</td>
<td>−0.000</td>
</tr>
<tr>
<td>TseTse Suitability</td>
<td>28166</td>
<td>9.763</td>
</tr>
<tr>
<td>Malaria Suitability</td>
<td>28166</td>
<td>0.562</td>
</tr>
<tr>
<td>Terrain Ruggedness</td>
<td>28166</td>
<td>37.836</td>
</tr>
<tr>
<td>Elevation</td>
<td>28166</td>
<td>608.049</td>
</tr>
<tr>
<td>Dist. Navigable River (log)</td>
<td>28166</td>
<td>5.016</td>
</tr>
<tr>
<td>Dist. Coast (log)</td>
<td>28166</td>
<td>6.208</td>
</tr>
<tr>
<td>Dist. European Explorer Route (log)</td>
<td>28166</td>
<td>4.608</td>
</tr>
<tr>
<td>Dist. 1900 City (log)</td>
<td>28166</td>
<td>5.973</td>
</tr>
<tr>
<td>Dist. Colonial Capital (log)</td>
<td>28166</td>
<td>6.222</td>
</tr>
<tr>
<td>Dist. Precolonial Trade Route (log)</td>
<td>28166</td>
<td>4.803</td>
</tr>
<tr>
<td>Slaves per Area (int.)</td>
<td>28166</td>
<td>0.204</td>
</tr>
<tr>
<td>Slaves per Area (high)</td>
<td>28166</td>
<td>0.229</td>
</tr>
<tr>
<td>Slaves per Area (missing)</td>
<td>28166</td>
<td>0.019</td>
</tr>
<tr>
<td>Precolonial Agriculture (int.)</td>
<td>28166</td>
<td>0.462</td>
</tr>
<tr>
<td>Precolonial Agriculture(high)</td>
<td>28166</td>
<td>0.265</td>
</tr>
<tr>
<td>Precolonial Agriculture (missing)</td>
<td>28166</td>
<td>0.040</td>
</tr>
<tr>
<td>Precolonial Chiefdoms (Y/N)</td>
<td>28166</td>
<td>0.619</td>
</tr>
<tr>
<td>Precolonial Statehood (Y/N)</td>
<td>28166</td>
<td>0.126</td>
</tr>
<tr>
<td>Precolonial Centralization (missing)</td>
<td>28166</td>
<td>0.095</td>
</tr>
<tr>
<td>Area on Land (sqkm)</td>
<td>28166</td>
<td>732.293</td>
</tr>
</tbody>
</table>
Table A4: Summary Statistics (Ethnic Homeland-Country Pairs)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>N</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Pctl(25)</th>
<th>Pctl(75)</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality Road 1998 (Y/N)</td>
<td>1,074</td>
<td>0.827</td>
<td>0.379</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Night Lights 2015 (Y/N)</td>
<td>1,074</td>
<td>0.778</td>
<td>0.416</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>City 2015 (Y/N)</td>
<td>1,074</td>
<td>0.709</td>
<td>0.455</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>DHS Household Wealth</td>
<td>771</td>
<td>−0.310</td>
<td>0.470</td>
<td>−1.772</td>
<td>−0.581</td>
<td>−0.086</td>
<td>1.824</td>
</tr>
<tr>
<td>Road Density 1998 (log)</td>
<td>1,074</td>
<td>−0.297</td>
<td>2.145</td>
<td>−4.605</td>
<td>−0.572</td>
<td>1.138</td>
<td>2.865</td>
</tr>
<tr>
<td>Mean Night Lights 2015 (log)</td>
<td>1,074</td>
<td>−1.603</td>
<td>2.496</td>
<td>−4.605</td>
<td>−4.379</td>
<td>0.284</td>
<td>6.564</td>
</tr>
<tr>
<td>Urban Population Density 2015 (log)</td>
<td>1,074</td>
<td>−2.656</td>
<td>1.714</td>
<td>−4.605</td>
<td>−4.605</td>
<td>−1.396</td>
<td>3.290</td>
</tr>
<tr>
<td>Mean Night Lights 2015 (log)</td>
<td>1,074</td>
<td>−1.422</td>
<td>2.374</td>
<td>−4.605</td>
<td>−4.007</td>
<td>0.377</td>
<td>5.755</td>
</tr>
<tr>
<td>Quality Road 1960 (Y/N)</td>
<td>1,074</td>
<td>0.443</td>
<td>0.497</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Railway 1960 (Y/N)</td>
<td>1,074</td>
<td>0.098</td>
<td>0.297</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Power Plant 1972 (Y/N)</td>
<td>1,074</td>
<td>0.303</td>
<td>0.460</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Colonial Cash Crop Dummy</td>
<td>1,074</td>
<td>0.363</td>
<td>0.481</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Colonial Mineral Dummy</td>
<td>1,074</td>
<td>0.067</td>
<td>0.250</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Colonial Cash Crop Value (log)</td>
<td>1,074</td>
<td>0.389</td>
<td>0.744</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>5</td>
</tr>
<tr>
<td>Colonial Mineral Value (log)</td>
<td>1,074</td>
<td>0.067</td>
<td>0.362</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Cash Crop Suitability</td>
<td>1,073</td>
<td>0.003</td>
<td>1.001</td>
<td>−1.588</td>
<td>−0.704</td>
<td>0.572</td>
<td>3.265</td>
</tr>
<tr>
<td>Caloric Suitability</td>
<td>1,073</td>
<td>−0.002</td>
<td>1.004</td>
<td>−2.957</td>
<td>−0.458</td>
<td>0.785</td>
<td>1.312</td>
</tr>
<tr>
<td>Tsetse Suitability</td>
<td>1,074</td>
<td>10.245</td>
<td>2.680</td>
<td>0.000</td>
<td>8.588</td>
<td>12.387</td>
<td>14.464</td>
</tr>
<tr>
<td>Malaria Suitability</td>
<td>1,074</td>
<td>0.626</td>
<td>0.177</td>
<td>0.039</td>
<td>0.488</td>
<td>0.774</td>
<td>0.944</td>
</tr>
<tr>
<td>Terrain Ruggedness</td>
<td>1,074</td>
<td>47.581</td>
<td>52.247</td>
<td>0.603</td>
<td>13.940</td>
<td>60.510</td>
<td>413.404</td>
</tr>
<tr>
<td>Elevation</td>
<td>1,074</td>
<td>608.444</td>
<td>448.116</td>
<td>9.257</td>
<td>292.595</td>
<td>879.243</td>
<td>2,364.656</td>
</tr>
<tr>
<td>Dist. Navigable River (log)</td>
<td>1,074</td>
<td>4.705</td>
<td>1.138</td>
<td>0.196</td>
<td>4.107</td>
<td>5.460</td>
<td>7.128</td>
</tr>
<tr>
<td>Dist. Coast (log)</td>
<td>1,074</td>
<td>6.043</td>
<td>1.221</td>
<td>0.052</td>
<td>5.545</td>
<td>6.896</td>
<td>7.490</td>
</tr>
<tr>
<td>Dist. European Explorer Route (log)</td>
<td>1,074</td>
<td>4.634</td>
<td>1.203</td>
<td>0.119</td>
<td>4.006</td>
<td>5.466</td>
<td>6.865</td>
</tr>
<tr>
<td>Dist. 1900 City (log)</td>
<td>1,074</td>
<td>5.692</td>
<td>0.825</td>
<td>1.813</td>
<td>5.246</td>
<td>6.301</td>
<td>7.290</td>
</tr>
<tr>
<td>Dist. Colonial Capital (log)</td>
<td>1,074</td>
<td>6.078</td>
<td>0.808</td>
<td>2.668</td>
<td>5.628</td>
<td>6.793</td>
<td>7.492</td>
</tr>
<tr>
<td>Dist. Precolonial Trade Route (log)</td>
<td>1,073</td>
<td>4.472</td>
<td>1.274</td>
<td>0.000</td>
<td>3.771</td>
<td>5.398</td>
<td>7.237</td>
</tr>
<tr>
<td>Precolonial Chiefdoms (Y/N)</td>
<td>1,074</td>
<td>0.575</td>
<td>0.495</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Precolonial Statehood (Y/N)</td>
<td>1,074</td>
<td>0.123</td>
<td>0.328</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Slaves per Area (log)</td>
<td>1,070</td>
<td>0.186</td>
<td>0.555</td>
<td>0.000</td>
<td>0.000</td>
<td>0.032</td>
<td>3.774</td>
</tr>
<tr>
<td>Precolonial Agriculture</td>
<td>1,024</td>
<td>6.124</td>
<td>1.812</td>
<td>1.000</td>
<td>5.000</td>
<td>8.000</td>
<td>10.000</td>
</tr>
<tr>
<td>Precolonial Town Size</td>
<td>402</td>
<td>4.206</td>
<td>2.026</td>
<td>1.000</td>
<td>3.000</td>
<td>7.000</td>
<td>8.000</td>
</tr>
<tr>
<td>Precolonial Class Stratification</td>
<td>950</td>
<td>0.575</td>
<td>0.495</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Political Centralization</td>
<td>975</td>
<td>1.317</td>
<td>0.979</td>
<td>0.000</td>
<td>1.000</td>
<td>2.000</td>
<td>4.000</td>
</tr>
<tr>
<td>Headman Elected</td>
<td>1,074</td>
<td>0.068</td>
<td>0.252</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Slavery (Murdock)</td>
<td>1,004</td>
<td>0.493</td>
<td>0.309</td>
<td>0.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Area (sqkm)</td>
<td>1,074</td>
<td>19,709.7</td>
<td>37,242.2</td>
<td>108.8</td>
<td>2,567.2</td>
<td>20,928.8</td>
<td>482,598.9</td>
</tr>
</tbody>
</table>
Table A5: Night-light Gini Coefficients across First and Second Administrative Units by World Regions

<table>
<thead>
<tr>
<th>Region</th>
<th>ADM 1 2000</th>
<th>ADM 1 2010</th>
<th>ADM 2 2000</th>
<th>ADM 2 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>0.693</td>
<td>0.673</td>
<td>0.867</td>
<td>0.847</td>
</tr>
<tr>
<td>Asia</td>
<td>0.604</td>
<td>0.572</td>
<td>0.763</td>
<td>0.747</td>
</tr>
<tr>
<td>South America</td>
<td>0.596</td>
<td>0.578</td>
<td>0.747</td>
<td>0.722</td>
</tr>
<tr>
<td>Oceania</td>
<td>0.468</td>
<td>0.482</td>
<td>0.744</td>
<td>0.739</td>
</tr>
<tr>
<td>North America</td>
<td>0.455</td>
<td>0.452</td>
<td>0.645</td>
<td>0.602</td>
</tr>
<tr>
<td>Europe</td>
<td>0.401</td>
<td>0.402</td>
<td>0.605</td>
<td>0.601</td>
</tr>
</tbody>
</table>

Table A5 draws from Alesina, Michalopoulos, and Papaioannou (2016). It reports Gini coefficients of nighttime lights at first-level and second-level administrative divisions for all countries in the world. Electrification inequality is highest in Africa compared to all world regions.
Table A6: Agriculture, Grains and Early Statehood?

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Precolonial Centralization</th>
<th>Precolonial State (Y/N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caloric Suitability</td>
<td>−0.175***</td>
<td>−0.025*</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Cereal Advantage</td>
<td>−0.182***</td>
<td>−0.027**</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.013)</td>
</tr>
</tbody>
</table>

Observations 986 978 986 978
Adjusted R² 0.180 0.181 0.115 0.112

Notes: OLS regression models with Murdock ethnic group polygons as the unit of analysis. Standard errors are clustered at the country level. Significance codes: *p<0.1; **p<0.05; ***p<0.01

Table A6 reports OLS models regressing pre-colonial centralization as measured by Murdock 1959, 1967 on two agricultural variables prominently discussed in the recent quantitative literature on early state formation. Some studies report an intricate link between agricultural productivity and the emergence of early states (Diamond 1997) whereas others argue that only easy-to-tax cereal grains incentivize state formation (Mayshar et al. 2018; Scott 2017). The first two columns in Table A6 use Murdock’s ordinal variable as coded, whereas the last two columns use a binary indicator of whether a particular ethnic group had state-like political organization. Drawing from (Galor and Özak 2016), the caloric suitability measure is an estimate of a given territory’s potential caloric yield and its ability to feed dense populations in premodern times. Cereal advantage measures the difference in caloric yield between cereal grains on the one hand and roots and tubers on the other as described by Mayshar et al. (2018). The correlation between these two agricultural variables and Murdock’s early statehood indicators is negative suggesting that other mechanisms may be more important in explaining the emergence of precolonial states in the African context.
Table A7: Baseline OLS Models (Cash Crops)

Baseline Analysis (OLS)

<table>
<thead>
<tr>
<th></th>
<th>Quality Road (Y/N)</th>
<th>Lights 2015 (Y/N)</th>
<th>City 2015 (Y/N)</th>
<th>HH Wealth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caloric Suitability</td>
<td>0.070***</td>
<td>0.049***</td>
<td>0.031***</td>
<td>−0.048*</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.009)</td>
<td>(0.007)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>TseTse Suitability</td>
<td>0.001</td>
<td>−0.012***</td>
<td>−0.002</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.002)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Malaria Suitability</td>
<td>−0.096</td>
<td>−0.128**</td>
<td>−0.035</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td>(0.071)</td>
<td>(0.060)</td>
<td>(0.039)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>Ruggedness</td>
<td>−0.0003***</td>
<td>−0.0001</td>
<td>−0.0002***</td>
<td>−0.001***</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0002)</td>
</tr>
<tr>
<td>Elevation</td>
<td>0.0001***</td>
<td>0.0002***</td>
<td>0.0003***</td>
<td>0.0004***</td>
</tr>
<tr>
<td></td>
<td>(0.00003)</td>
<td>(0.00002)</td>
<td>(0.00002)</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>Dist. Navigable River (log)</td>
<td>−0.031***</td>
<td>−0.033***</td>
<td>−0.024***</td>
<td>−0.059***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.004)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Dist. Coast (log)</td>
<td>−0.017*</td>
<td>−0.086***</td>
<td>−0.035***</td>
<td>−0.127***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.008)</td>
<td>(0.006)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>Dist. Trade Route 1900 (log)</td>
<td>−0.020***</td>
<td>−0.012***</td>
<td>0.0004</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Dist. Explorer Route (log)</td>
<td>−0.066***</td>
<td>−0.061***</td>
<td>−0.039***</td>
<td>−0.100***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.009)</td>
<td>(0.007)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Dist. City 1900 (log)</td>
<td>−0.061***</td>
<td>−0.046***</td>
<td>−0.024***</td>
<td>−0.132***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.007)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>Dist. Colonial Capital (log)</td>
<td>−0.009*</td>
<td>−0.005</td>
<td>−0.003</td>
<td>−0.020**</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Slavery (Med.)</td>
<td>−0.016</td>
<td>−0.011</td>
<td>−0.016**</td>
<td>−0.003</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.015)</td>
<td>(0.008)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Slavery (High)</td>
<td>0.008</td>
<td>0.006</td>
<td>−0.013</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.013)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Precolonial Agriculture (Med.)</td>
<td>0.047**</td>
<td>0.056***</td>
<td>0.018*</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.014)</td>
<td>(0.009)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Precolonial Agriculture (High)</td>
<td>−0.004</td>
<td>0.044***</td>
<td>0.026**</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.015)</td>
<td>(0.012)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Precolonial Chiefdom</td>
<td>−0.006</td>
<td>−0.013</td>
<td>0.006</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.013)</td>
<td>(0.011)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Precolonial State</td>
<td>−0.010</td>
<td>−0.006</td>
<td>0.011</td>
<td>−0.054</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.023)</td>
<td>(0.017)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>Precolonial Statehood Missing</td>
<td>0.062**</td>
<td>−0.011</td>
<td>0.029*</td>
<td>−0.069</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.021)</td>
<td>(0.017)</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Colonial Cash Crops (Y/N)</td>
<td>0.164***</td>
<td>0.196***</td>
<td>0.187***</td>
<td>0.144***</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.027)</td>
</tr>
</tbody>
</table>

Notes: OLS regression models. Column 4 restricts the sample to all grid cells with information on asset-based household wealth from the DHS survey. All models include colony fixed effects and standard errors (in parentheses) are clustered at the level of ethnic group polygons from George Peter Murdoch’s “Tribal Map of Africa.” Significance codes: ’p<0.1; **’p<0.05; ***’p<0.01
Table A8: Baseline OLS Models (Minerals)

<table>
<thead>
<tr>
<th>Baseline Analysis (OLS)</th>
<th>Quality Road (Y/N)</th>
<th>Lights 2015 (Y/N)</th>
<th>City 2015 (Y/N)</th>
<th>HH Wealth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calorie Suitability</td>
<td>0.070*** (0.012)</td>
<td>0.049*** (0.009)</td>
<td>0.018*** (0.007)</td>
<td>−0.051* (0.027)</td>
</tr>
<tr>
<td>Tsetse Suitability</td>
<td>0.001 (0.004)</td>
<td>−0.012*** (0.006)</td>
<td>−0.005 (0.002)</td>
<td>0.001 (0.007)</td>
</tr>
<tr>
<td>Malaria Suitability</td>
<td>−0.082 (0.072)</td>
<td>−0.124*** (0.064)</td>
<td>−0.031 (0.043)</td>
<td>0.070 (0.143)</td>
</tr>
<tr>
<td>Raggedness</td>
<td>−0.003*** (0.0003)</td>
<td>−0.0001 (0.0001)</td>
<td>−0.002*** (0.0001)</td>
<td>−0.001*** (0.0001)</td>
</tr>
<tr>
<td>Elevation</td>
<td>0.0002*** (0.0002)</td>
<td>0.0021*** (0.0002)</td>
<td>0.0001*** (0.0001)</td>
<td>0.0004*** (0.0001)</td>
</tr>
<tr>
<td>Dist. Navigable River (log)</td>
<td>−0.017*** (0.009)</td>
<td>−0.031*** (0.009)</td>
<td>−0.025*** (0.007)</td>
<td>−0.058*** (0.007)</td>
</tr>
<tr>
<td>Dist. Coast (log)</td>
<td>−0.012*** (0.006)</td>
<td>−0.099*** (0.006)</td>
<td>−0.039*** (0.006)</td>
<td>−0.131*** (0.006)</td>
</tr>
<tr>
<td>Dist. Trade Route 1900 (log)</td>
<td>−0.020*** (0.006)</td>
<td>−0.012*** (0.004)</td>
<td>0.0001 (0.003)</td>
<td>0.008 (0.007)</td>
</tr>
<tr>
<td>Dist. Explorer Route (log)</td>
<td>−0.070*** (0.011)</td>
<td>−0.066*** (0.010)</td>
<td>−0.044*** (0.009)</td>
<td>−0.105*** (0.010)</td>
</tr>
<tr>
<td>Dist. City 1900 (log)</td>
<td>−0.071*** (0.011)</td>
<td>−0.079*** (0.011)</td>
<td>−0.034*** (0.008)</td>
<td>−0.143*** (0.008)</td>
</tr>
<tr>
<td>Dist. Colonial Capital (log)</td>
<td>−0.009* (0.006)</td>
<td>−0.006 (0.005)</td>
<td>−0.004 (0.003)</td>
<td>−0.022*** (0.008)</td>
</tr>
<tr>
<td>Slavery (Med.)</td>
<td>−0.018 (0.019)</td>
<td>−0.014 (0.015)</td>
<td>−0.018** (0.015)</td>
<td>0.002 (0.012)</td>
</tr>
<tr>
<td>Slavery (High)</td>
<td>0.011 (0.019)</td>
<td>0.010 (0.020)</td>
<td>−0.009 (0.014)</td>
<td>0.021 (0.014)</td>
</tr>
<tr>
<td>Precolonial Agriculture (Med.)</td>
<td>0.048** (0.020)</td>
<td>0.092** (0.015)</td>
<td>0.019** (0.015)</td>
<td>0.015 (0.015)</td>
</tr>
<tr>
<td>Precolonial Agriculture (High)</td>
<td>0.012*** (0.022)</td>
<td>0.051*** (0.015)</td>
<td>0.031** (0.015)</td>
<td>0.017 (0.015)</td>
</tr>
<tr>
<td>Precolonial Chiefdom</td>
<td>−0.001 (0.022)</td>
<td>−0.012 (0.016)</td>
<td>0.007 (0.012)</td>
<td>−0.001 (0.012)</td>
</tr>
<tr>
<td>Precolonial State</td>
<td>−0.012 (0.013)</td>
<td>−0.009 (0.024)</td>
<td>0.008 (0.018)</td>
<td>−0.065 (0.023)</td>
</tr>
<tr>
<td>Precolonial Statehood Missing</td>
<td>0.068*** (0.012)</td>
<td>−0.004 (0.021)</td>
<td>0.016** (0.021)</td>
<td>−0.067 (0.021)</td>
</tr>
<tr>
<td>Colonial Mining (Y/N)</td>
<td>0.154*** (0.015)</td>
<td>0.166*** (0.044)</td>
<td>0.223*** (0.039)</td>
<td>0.391*** (0.067)</td>
</tr>
</tbody>
</table>

Observations: 28,166
Adj. R²: 0.161
Baseline OLS Models (Minerals)

Note: OLS regression models. Column 4 restricts the sample to all grid cells with information on asset-based household wealth from the DHS survey. All models include colony fixed effects and standard errors (in parentheses) are clustered at the level of ethnic group polygons from George Peter Murdock’s “Tribal Map of Africa.” Significance codes: * p<0.1; ** p<0.05; *** p<0.01

A10
Table A9: Cash Crop Suitability, Early Infrastructure, Continued Agricultural Production

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Caloric Suitability</td>
<td>1.862***</td>
<td>0.091***</td>
<td>0.0001</td>
<td>0.004*</td>
</tr>
<tr>
<td>(0.135)</td>
<td>(0.008)</td>
<td>(0.003)</td>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td>TseTse Suitability</td>
<td>−0.229***</td>
<td>−0.098***</td>
<td>−0.001</td>
<td>−0.003***</td>
</tr>
<tr>
<td>(0.056)</td>
<td>(0.003)</td>
<td>(0.001)</td>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Malaria Suitability</td>
<td>0.615</td>
<td>−0.046</td>
<td>−0.040*</td>
<td>−0.020</td>
</tr>
<tr>
<td>(0.654)</td>
<td>(0.044)</td>
<td>(0.024)</td>
<td></td>
<td>(0.014)</td>
</tr>
<tr>
<td>Ruggedness</td>
<td>0.082**</td>
<td>−0.002***</td>
<td>−0.0001</td>
<td>0.00002</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.0002)</td>
<td>(0.0001)</td>
<td></td>
<td>(0.0001)</td>
</tr>
<tr>
<td>Elevation</td>
<td>0.0001</td>
<td>0.0001***</td>
<td>0.00001</td>
<td>0.00002***</td>
</tr>
<tr>
<td>(0.0001)</td>
<td>(0.0002)</td>
<td>(0.0001)</td>
<td></td>
<td>(0.0001)</td>
</tr>
<tr>
<td>Dist. Navigable River</td>
<td>−0.459***</td>
<td>−0.014***</td>
<td>−0.004</td>
<td>−0.006***</td>
</tr>
<tr>
<td>(0.069)</td>
<td>(0.004)</td>
<td>(0.001)</td>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Dist. Coast</td>
<td>−0.325***</td>
<td>−0.095***</td>
<td>−0.012***</td>
<td>−0.019***</td>
</tr>
<tr>
<td>(0.062)</td>
<td>(0.006)</td>
<td>(0.003)</td>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td>Cash Crop Suitability</td>
<td>0.504***</td>
<td>0.048***</td>
<td>0.015***</td>
<td>0.010***</td>
</tr>
<tr>
<td>(0.088)</td>
<td>(0.007)</td>
<td>(0.003)</td>
<td></td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

Observations 28,166 28,166 28,166 28,166
Adjusted R² 0.620 0.218 0.033 0.042
Colony FE Yes Yes Yes Yes

Notes: OLS regression models. All models include colony fixed effects and standard errors are clustered at the level of ethnic group polygons from George Peter Murdock’s “Tribal Map of Africa.” Significance codes: *p<0.1; **p<0.05; ***p<0.01
Table A10: Colonial Cash Crop Production, Early Infrastructure, Continued Agricultural Production

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Caloric Suitability</td>
<td>1.914***</td>
<td>0.086***</td>
<td>0.082</td>
<td>0.006***</td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.007)</td>
<td>(0.001)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Tsetse Suitability</td>
<td>−0.201***</td>
<td>−0.006**</td>
<td>0.0002</td>
<td>−0.002**</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Malaria Suitability</td>
<td>1.390**</td>
<td>0.015</td>
<td>−0.021</td>
<td>−0.009</td>
</tr>
<tr>
<td></td>
<td>(0.630)</td>
<td>(0.041)</td>
<td>(0.022)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Raggedness</td>
<td>0.002</td>
<td>−0.0002***</td>
<td>−0.0008**</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.0001)</td>
<td>(0.0003)</td>
<td>(0.0003)</td>
</tr>
<tr>
<td>Elevation</td>
<td>0.001**</td>
<td>0.0001***</td>
<td>0.00000</td>
<td>0.0002***</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.00002)</td>
<td>(0.00005)</td>
<td>(0.00007)</td>
</tr>
<tr>
<td>Dist. Navigable River</td>
<td>−0.406***</td>
<td>−0.011***</td>
<td>−0.001</td>
<td>−0.005***</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.004)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Dist. Coast</td>
<td>−0.269***</td>
<td>−0.090***</td>
<td>−0.009***</td>
<td>−0.106***</td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td>(0.006)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Colonial Cash Crops (Y/N)</td>
<td>1.168****</td>
<td>0.204***</td>
<td>0.071****</td>
<td>0.490****</td>
</tr>
<tr>
<td></td>
<td>(0.129)</td>
<td>(0.010)</td>
<td>(0.031)</td>
<td>(0.007)</td>
</tr>
</tbody>
</table>

Observations: 28,166 28,166 28,166 28,166
Adjusted R²: 0.620 0.218 0.033 0.042

Note: OLS regression models. All models include colony fixed effects and standard errors are clustered at the level of ethnic group polygons from George Peter Murdock’s “Tribal Map of Africa.” Significance codes: *p<0.1, **p<0.05, ***p<0.01
Figure A1: Correlations between Hance’s primary commodity data (standardized in 1957 US$) with the primary commodity data we collected from colonial reports, maps and other records (standardized in 1960 US$; see additional appendix below). Data have been aggregated to second-level (in some cases first-level) administrative districts (regions) at independence (c. 1960).
Figure A2: Statistical balance of colonial cash crop and mineral production cells before and after using propensity score reweighting (Hirano, Imbens, and Ridder 2003).
Figure A3: Figure S3a maps out counterfactual cash crop (upper panels) and mining (lower panel) cells sampled from 1000 draws. See Methods section in the main text for description of counterfactual sampling procedure. The right-hand panels also show actual colonial production sites from Hance, Kotschar, and Peterec (1961).
Figure A4: Statistical balance on key covariates among actual versus counterfactual cash crop cells.
Figure A5: Statistical balance on key covariates among actual versus counterfactual mining site cells.
Figure A6: Spatial randomization results replacing binary with continuous development outcomes.
Figure A7: Robustness of S2SLS-IV results

Figure A7 reports results from spatial 2SLS-IV regression models but employs alternative definitions of what constitutes the spatial neighborhood of a cell. The first four IV models use spatial lags on inverse-distance weighted nearest-neighbor matrices, whereas the last four models use binary distance cutoffs at 0.35, 0.71, 1.06, and 1.41 lon/lat degrees. The upper panel reports coefficients and confidence intervals for the instrumented colonial cash dummy, whereas the lower panel reports F statistics for the two first stages in each model (i) cash crop dummy instrumented with suitability; (ii) spatial lag of the respective dependent variable instrumented with first and second-order spatial lags of the baseline covariates).
Figure A8: Placebo Test: Cash Crop Suitability and Pre-Colonial Development

Figure A8 conducts placebo tests probing whether our suitability instrument predicts pre-colonial development outcomes. Any developmental advantage of highly suitable areas before the commercial transition to export agriculture could indicate a violation of the exclusion restriction. All placebo regressions include distance to coast, navigable waterways, caloric suitability, Tsetse fly and malaria suitability, ruggedness, and elevation. The first three models are run at the grid cell-level and indicate that cells suitable for cash crops are not significantly closer to pre-colonial trade routes or urban centers and, if anything, further away from European explorer routes. The subsequent models are at the level of Murdock’s ethnic homelands and show no significant associations with variables from Murdock’s Ethnographic Atlas that can be seen as proxies for socioeconomic or institutional development.
Figure A9: Disaggregating effects by five main cash crops: cocoa, coffee, cotton, groundnuts, and oil palm.

Figure A13 reports the impact of colonial cash crops (upper panel) and mining (lower panel) on contemporary development outcomes by various country subgroups: different levels of democracy; conflict-affectedness; agricultural/mineral export dependence; and landlocked vs. coastal. Marginal effect estimates from multiplicative interaction models. Continuous moderators are made categorical based on sample terciles.
Figure A10: Country-by-country regressions of infrastructural and wealth outcomes on colonial cash crop production.
Figure A11: Country-by-country regressions of infrastructural and wealth outcomes on colonial mining.
Figure A12: Empire-by-empire regressions of main outcome variables on colonial cash crop production (upper panel) and mining (lower panel).
Heterogeneity Analysis
OLS Interaction Models with Colonial Cash Crop Dummy as Main Predictor

Heterogeneity Analysis
OLS Interaction Models with Colonial Mineral Dummy as Main Predictor

Figure A13: Heterogeneity by post-independence trajectories.
Figure A14: Continuous outcomes and intensive margin effects

Figure A14 shows estimates from models with continuous instead of dichotomous infrastructure and development outcomes. We also test for intensive-margin effects of our resource variables by subsetting the sample to cells with non-zero production or at least one quality road, city, or some night lights.

Figure A15 shows robustness of our spatial IV results when using continuous versions of the outcome and/or treatment variables.

Figure A15: Spatial IV analysis using continuous outcome variables and/or continuous versions of the cash crop treatment.
| Distance Cutoff for Conley Errors |
|---------------------|---------------------|---------------------|
| Baseline            | Quality Road 1998 (Y/N) | Night Lights 2015 (Y/N) | City 2015 (Y/N) | DHS Household Wealth |
| 100 km              |                      |                      |                |
| 200 km              |                      |                      |                |
| 400 km              |                      |                      |                |
| 800 km              |                      |                      |                |
| 1600 km             |                      |                      |                |
| 3200 km             |                      |                      |                |

**Figure A16**: Conley (1999) standard errors with various distance cutoffs to address concerns about spatial dependence and artificially small uncertainty estimates.
Figure A17: Binary resource variables across different spatial units of analysis.

In Figures A17 and A18, we check whether our results are stable to different definitions of the spatial unit of analysis. For this purpose, we aggregate all variables to grids of 0.5 and 1.0 degree (long/lat) resolution (55.5 and 111 km, respectively, at the equator), resulting in cells that are, on average, 4 and 16 times larger than in our baseline analysis. In addition, we aggregate our data to ethnic homeland-country pairs as in Michalopoulos and Papaioannou 2013.

We use binary and continuous versions of our main outcome and predictor variables. Figure A17 presents results from models employing binary cash crop or mineral dummies as main independent variable. With the exception of the urbanization outcomes, the binary resource coefficients remain stable or decrease in size as the grid gets coarser. This seems consistent with the limited or even negative spillover effects that we report in the main paper as well as in Figures A22 and A23 below. The ethnic group-level coefficients are comparable to our baseline analysis using quarter-
Figure A18: Continuous resource variables across different spatial units of analysis.
degree cells. A similar relationship between grid cell and coefficient size emerges if we replace binary with continuous resource variables (Figure A18).
In Figure A19 we compare the effects of colonial cash crop agriculture to two historical factors often cited as shaping long-run development in Africa—the slave trades (Nunn 2008; Nunn and Puga 2010; Nunn and Wantchekon 2011) and pre-colonial centralization (Gennaioli and Rainer 2007; Michalopoulos and Papaioannou 2013). Both of these variables are available at the level of ethnic homelands as defined by Murdock (1959, 1967). We follow Michalopoulos and Papaioannou (2013) and run models with ethnic homeland-country pairs as the unit of analysis, include country fixed effects and cluster standard errors at the the country level. We find that the historical effects of cash crop production on contemporary infrastructural development at the subnational level are much more robust than those of the other historical variables. The impact of pre-colonial centralization is modest and sensitive to selection of control variables. And, compared to the slave trades’ negative effects on social institutions and trust (Nunn and Wantchekon 2011), areas most affected by the slave trades tend to have slightly higher levels of contemporary urbanization and wealth—due perhaps to their favorable geography and proximity to the coast, the targeting of slave raids to initially densely populated areas, or the persistence of markets organized around trade.
Figure A20: Replication of the mechanism analysis from the main text with actual colonial cash crop production instead of agro-climatic suitability as the main treatment variable.

Figure A21: Effect size comparisons from as in the main text but with continuous instead of binary outcome variables.
Figure A22: Distance band regressions (cash crops) using continuous (lower panel) instead of binary outcome variables (upper panel, for comparison).
Figure A23: Distance band regressions (minerals) using continuous (lower panel) and binary outcome variables.
References


B  Sources of Primary Commodity Production by Colony

The dataset we compiled to validate Hance, Kotschar, and Peterec (1961) in Figure A1 only includes primary commodities that account for 10% or more of total exports as of 1960. Aggregate primary commodity exports data come from the United Nations, Yearbook of International Statistics for as close to 1960 as available. Country-specific subnational production data comes from unique sets of sources for each colony or sets of colonies (French West Africa, French Equatorial Africa), again as close as possible to the year 1960. We use these data sources to validate the Hance map at the level of subnational administrative units (districts or regions). See Figure S1 above.

Angola

- Primary commodity exports as of 1960
  - Coffee (36%)
  - Diamonds (17%)

- Sources of administrative and spatial primary commodity production data
Belgian Congo (DRC)

- Primary commodity exports as of 1960
  - Copper (27%)
  - Coffee (14%)
  - Palm (13%)

- Sources of administrative and spatial primary commodity production data

Cameroon

- Primary commodity exports as of 1960
  - Cocoa (25%)
  - Coffee (20%)
  - Aluminium (20%)

- Sources of administrative and spatial primary commodity production data

A38
Central African Republic

- Primary commodity exports as of 1960
  - Cotton (44.7%)
  - Coffee (24.9%)
  - Diamonds (12.1%)

- Sources of administrative and spatial primary commodity production data

Chad

- Primary commodity exports as of 1960
  - Cotton (80%)

- Sources of administrative and spatial primary commodity production data
Côte d'Ivoire

- Primary commodity exports as of 1960
  - Coffee (45%)
  - Cocoa (20%)

- Sources of administrative and spatial primary commodity production data

Dahomey (Benin)

- Primary commodity exports as of 1960
  - Palm products (66%)

- Sources of administrative and spatial primary commodity production data

Ethiopia

- Primary commodity exports as of 1960
  - Coffee (56.5%)
• Sources of administrative and spatial primary commodity production data

Gambia

• Primary commodity exports as of 1960
  ○ Groundnuts (88.2%)

• Sources of administrative and spatial primary commodity production data

Guinea

• Primary commodity exports as of 1960
  ○ Bauxite/Aluminum (47.4%)
  ○ Diamonds (12.6%)
  ○ Bananas (10.7%)
  ○ Coffee (10.5%)

• Sources of administrative and spatial primary commodity production data
Gold Coast (Ghana)

- Primary commodity exports as of 1960
  - Cocoa (60%)

- Sources of administrative and spatial primary commodity production data

Kenya

- Primary commodity exports as of 1960
  - Coffee (24%)
  - Tea (24%)
  - Sisal (11%)

- Sources of administrative and spatial primary commodity production data


**Liberia**

- Primary commodity exports as of 1960
  - Iron Ore (47%)
  - Rubber (42%)

- Sources of administrative and spatial primary commodity production data
**French Soudan (Mali)**

- Primary commodity exports as of 1960
  - Coffee (56.5%)

- Sources of administrative and spatial primary commodity production data

**Madagascar**

- Primary commodity exports as of 1960
  - Coffee (40%)

- Sources of administrative and spatial primary commodity production data

**Mozambique**

- Primary commodity exports as of 1960
  - Cotton (32%)
  - Sugar (13%)
  - Cashew (13%)
• Sources of administrative and spatial primary commodity production data

Nyasaland (Malawi)

• Primary commodity exports as of 1960
  ◦ Tobacco (45%)
  ◦ Tea (40%)

• Sources of administrative and spatial primary commodity production data

Niger

- Primary commodity exports as of 1960
  - Groundnuts (70–80%)

- Sources of administrative and spatial primary commodity production data
  - République du Niger. Cultures Vivrières; Cultures Industrielles (unsourced map from Library of Congress received in 1962).

Nigeria

- Primary commodity exports as of 1960
  - Palm products (26%)
  - Cocoa (24%)
  - Groundnuts (15%)

- Sources of administrative and spatial primary commodity production data
Northern Rhodesia (Zambia)

- Primary commodity exports as of 1960
  - Copper (90%)

- Sources of administrative and spatial primary commodity production data
  - Northern Rhodesia Department of Geological Survey. 1959. Mineral Map of Northern Rhodesia. Lusaka.
  - Northern Rhodesia Census. 1951. Lusaka: Government Printer.
  - Rhodesia and Nyasaland. 1959. Federation of Rhodesia and Nyasaland: Designed and printed by the Government Printer (Scale [ca. 1:4,000,000]). Salisbury.
Ruanda-Urundi (Rwanda, Burundi)

- Primary commodity exports as of 1960
  - Coffee (75%)

- Sources of administrative and spatial primary commodity production data

Senegal

- Primary commodity exports as of 1960
  - Groundnuts (80%)

- Sources of administrative and spatial primary commodity production data
Sierra Leone

- Primary commodity exports as of 1960
  - Diamonds (55%)
  - Iron ore (18%)

- Sources of administrative and spatial primary commodity production data

Somalia

- Primary commodity exports as of 1960
  - Bananas (70%)

- Sources of administrative and spatial primary commodity production data

Southern Rhodesia (Zimbabwe)

- Primary commodity exports as of 1960
○ Tobacco (30%)

- Sources of administrative and spatial primary commodity production data

Sudan

- Primary commodity exports as of 1960
  ○ Cotton (55%)

- Sources of administrative and spatial primary commodity production data

Tanganyika (Tanzania)

- Primary commodity exports as of 1960
  ○ Sisal (27%)
  ○ Coffee (17%)
  ○ Cotton (16%)

- Sources of administrative and spatial primary commodity production data


**Togo**

- Primary commodity exports as of 1960
  - Cocoa (28%)
  - Coffee (25%)

- Sources of administrative and spatial primary commodity production data

**Uganda**

- Primary commodity exports as of 1960
  - Coffee (40%)
  - Cotton (35%)
• Sources of administrative and spatial primary commodity production data

Zanzibar

• Primary commodity exports as of 1960
  ○ Cloves (75%)

• Sources of administrative and spatial primary commodity production data
  ○ Zanzibar Blue Book. 1946.