Stanford | King Center on Global Development

Concessions, Violence, and Indirect Rule: Evidence from the Congo Free State

Sara Lowes

Eduardo Montero

October, 2020

Working Paper No. 1079

NBER WORKING PAPER SERIES

CONCESSIONS, VIOLENCE, AND INDIRECT RULE: EVIDENCE FROM THE CONGO FREE STATE

Sara Lowes
Eduardo Montero

Working Paper 27893
http://www.nber.org/papers/w27893

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2020

We are grateful for the financial support from: the Danielien Travel and Research Grant, The Eric M. Mindich Research Fund for the Foundations of Human Behavior, Lab for Economics Applications and Policy, Warburg Fund, IQSS, Weatherhead Center Graduate Student Associate Grant, the History Project and the Institute for New Economic Thinking (INET). Sara is grateful for financial support from the NSF Graduate Research Fellowship Program. This project has IRB approval from the Harvard CUHS (IRB15-2086). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Concessions, Violence, and Indirect Rule: Evidence from the Congo Free State
Sara Lowes and Eduardo Montero
NBER Working Paper No. 27893
October 2020
JEL No. D72,N47,O15,O43,Z1,Z13

ABSTRACT

All colonial powers granted concessions to private companies to extract natural resources during the colonial era. Within Africa, these concessions were characterized by indirect rule and violence. We use the arbitrarily defined borders of rubber concessions granted in the north of the Congo Free State to examine the causal effects of this form of economic organization on development. We find that historical exposure to the concessions causes significantly worse education, wealth, and health outcomes. To examine mechanisms, we collect survey and experimental data from individuals near a former concession boundary. We find that village chiefs inside the former concessions provide fewer public goods, are less likely to be elected, and are more likely to be hereditary. However, individuals within the concessions are more trusting, more cohesive, and more supportive of sharing income. The results are relevant for the many places that were designated as concessions to private companies during the colonial era.

Sara Lowes
Department of Economics
University of California at San Diego
9500 Gilman Drive
La Jolla, CA 92093
and NBER
slowes@ucsd.edu

Eduardo Montero
Ford School of Public Policy
University of Michigan
4219 Weill Hall
735 S State St
Ann Arbor, MI 48109
emontero@umich.edu

An online appendix is available at http://www.nber.org/data-appendix/w27893

1. Introduction

The Congo Free State (CFS), what is today the Democratic Republic of Congo (DRC), was the personal colony of King Leopold II of Belgium between 1885 and 1908. Leopold designated large parts of the CFS as concessions to private companies. These private companies extracted natural resources by using extreme violence and by coopting local leaders. Historians have noted that the rubber concessions granted under Leopold II had disastrous consequences for local populations. An estimated 10 million people, approximately half of the population of Congo, died between 1880 and 1920 (Vansina, 2010, Hochschild, 1998).

While the concessions granted under the CFS and the subsequent abuses that occurred are particularly infamous, all major colonial powers granted concessions to private companies. Within Africa, concessions existed in French, British, Belgian, German, and Portuguese colonies (e.g. Angola, Botswana, Central African Republic, Cameroon, Chad, DRC, Gabon, Malawi, Mozambique, Namibia, Nigeria, Republic of Congo, Tanzania, Zambia, Zimbabwe).[1] While the form of concessions varied widely across contexts, a common element is that these companies' primary purpose was to extract natural resources. They were assigned powers that are typically associated with government – such as monopoly over violence and ability to tax. In sub-Saharan Africa, concessions to private companies were characterized by indirect rule, in which the scope and power of local leaders was altered and coopted, and the use of violence to achieve their extraction goals (Mamdani, 1996, Michalopoulos and Papaioannou, 2020).

While concessions to private companies were common across colonial Africa as a form of organizing economic production, causal evidence on their impacts has been difficult to obtain. The lack of evidence is due in part to how ubiquitous these concessions were. Some colonies were completely run as concessions.[2] Additionally, the concessions' boundaries were rarely randomly allocated, and concessions were often granted in resource rich areas.

However, the rubber concessions granted in northern Congo – Abir and Anversoise – offer an opportunity to test how exposure to private concession companies has affected development. Abir and Anversoise were defined using salient geographic features at a time when relatively

[1] Concessions to private companies also existed outside of Africa, for example: the East India Company in India, the United Fruit Company in much of Central America (Méndez-Chacón and Van Patten, 2020), and the Dutch East India company in Indonesia.

[2] For example, all of Rhodesia (present day Zimbabwe) was granted as a concession to the British South Africa company.

little was known of the interior of Congo. Additionally, the former concession boundaries do not align with any present day political boundaries, ethnic group boundaries, or any particular natural resource endowment. The concessions only existed for 14 years, a relatively short amount of time. Finally, the concessions focused exclusively on extraction, rather than on investments in infrastructure or human capital (in contrast to Dell and Olken, 2020 and Méndez-Chacón and Van Patten, 2020). These features allow us to isolate the effects of exposure to extraction-based colonial concessions on our outcomes of interest (Michalopoulos and Papaioannou, 2020). In this study, we examine the effects of the rubber concessions granted in the CFS on development, the performance and accountability of local leaders, and on pro-social norms and behavior.

Exposure to the concessions in the CFS was characterized by violence and the use of village chiefs to enforce rubber collection quotas. The concession companies were given monopoly rights over natural resource extraction within the concession boundaries. European agents had monetary incentives tied to rubber production and were given support from the CFS armed forces (the *Force Publique*), their own militias, and a state mandate to use coercive means to reach their rubber extraction goals. Historical accounts of the rubber concession period highlight how the rubber companies forced village chiefs to support the rubber regime. Those who did not support the rubber regime were killed and replaced by outsiders willing to enforce the rubber quotas (Harms, 1975, Vangroenweghe, 1986). This combination of cooption and coercion is a highly relevant bundle for the colonial context and likely the most relevant for understanding the importance of colonial extraction (Michalopoulos and Papaioannou, 2020).

We examine the implications of the rubber concessions for development outcomes using a geographic regression discontinuity design. Consistent with the idiosyncratic manner with which the historical boundaries were determined, we demonstrate that those areas designated as concessions are geographically and culturally similar to the areas just outside of the concessions. We then use Demographic and Health Survey (DHS) data from 2007 and 2014 to estimate the effects of the historical concessions on present-day education, wealth, and health outcomes. We find that individuals from the former concessions have significantly worse outcomes. The results are robust to a variety of analyses, including: alternative RD specifications, bandwidths, and kernels; a donut hole analysis; dropping observations along the Congo river; standard error adjustments to account for spatial auto-correlation; and an analysis by concession. We address several possible concerns with examining the effects of the historical rubber concessions: the use of river basins to define boundaries, selective migration, and subsequent colonial or missionary investment. We find no evidence that these factors explain our observed results.

We then examine mechanisms. We present a conceptual framework based on three premises. First, the use of indirect rule may have led to less accountable and more despotic leaders (Mamdani, 1996, Vanhee, 2005). Second, weakened institutions and violence may have encouraged more pro-social behavior; as local institutions failed to function, returns to investing in social capital became higher and cooperation became more important (von Dawans, Fischbacher, Kirschbaum, Fehr and Heinrichs, 2012, Bauer, Blattman, Chytilová, Henrich, Miguel and Mitts, 2016). Third, institutions and culture may act as substitutes, leading to a low-development equilibrium (Bisin and Verdier, 2017).

We collected survey and experimental data from 520 individuals in Gemena, DRC, a town on the border of the former Anversoise concession. Gemena was created after the end of the concession era; therefore, those who live there are migrants themselves or descendants of migrants. Our analysis follows the epidemiological approach (Fernández, 2011): we compare data from individuals in Gemena with ancestors from inside the former concessions to those with ancestors from outside the former concessions.

Using data on the villages of origin of our respondents, we examine the selection mechanism for the chief and the extent to which the chief provides various public goods for the village. Generally, village chiefs are chosen either through elections, in which village members or representatives of villagers can vote among several candidates, or are hereditary, coming from one specific lineage within the village. We find that village chiefs within the former concessions are 34 percentage points less likely to be elected to their position and are more likely to be hereditary. This is possibly due to the common historical practice of anointing particular lineages as "ruling" lineages, from which subsequent village chiefs would be chosen. In some cases, medals were given to particular chiefs as proof that their lineage is a ruling lineage (Vanhee, 2005, Omasombo, 2013).

We then examine the responsibilities of the village chief. The village chief serves a particularly important role in this context, as the national government is weak and generally unable to provide basic public goods (Sanchez de la Sierra, 2020). We find that village chiefs inside the former concessions are less likely to provide public goods, such as road maintenance, conflict arbitration, and maintenance of schools. As a consequence, these villages also have fewer public goods.

Building on other work examining the effects of colonial extraction (Dell, 2010, Dell and Olken, 2020) which primarily focuses on institutional and investment channels, we test whether the concession system had consequences for several different cultural traits, including trust, social cohesion, altruism, and support for sharing income. This is important given growing evidence of the feedback between institutions and cultural traits (Bisin and Verdier, 2017).

Using survey questions on trust in a variety of other individuals or groups, we find that individuals from areas exposed to the rubber concessions are more trusting of others than those just outside the former concessions. We also examine measures of social cohesion and support for sharing income. We find that individuals from the former concession areas report feeling closer to a variety of others, are more likely to agree with statements that money should be shared with others, and are more likely to redistribute in an experimental task. Based on the historical narrative and our conceptual framework, we suggest that these changes may due to a need to compensate for the lower quality institutions that arose from the concession era. Additionally, these changes may reinforce each other: chiefs are held less accountable and allowed to stay in power since individuals rely less on their formal institutions and instead rely more on informal sharing norms for support.

We contribute to several literatures. An important set of studies examine the effects of colonialism, pre-colonial characteristics, and exposure to the slave trade on modern outcomes (La Porta, López de Silanes, Shleifer and Vishny, 1998, Acemoglu, Johnson and Robinson, 2001, Gennaioli and Rainer, 2007, Michalopoulous and Papaioannou, 2013, 2014, Alsan, 2015, Nunn, 2008). A related literature has examined the effects of investments made by colonial governments (Huillery, 2009, Cage and Rueda, 2016, 2017, Osafo-Kwaako, 2012, Wantchekon, Klasnja and Novta, 2015, Lowes and Montero, 2020).[3]

However, there is no evidence on the concession system, despite that it was a very common form of economic organization used by all colonial powers to extract natural resources from their colonies (Michalopoulos and Papaioannou, 2020).[4] In the case of these concession companies, they did not make productive investments in these areas. Rubber requires little capital investment to be collected, and it does not require training of the labor force. The primary input is labor, and the concession areas are connected to river networks so that there was no need to invest in road

[3] See Michalopoulos and Papaioannou (2020) for a thorough review of work examining historical legacies in Africa.

[4] Current evidence mostly focuses on cases in which companies made productive investments (Juif and Frankema, 2017, Dell and Olken, 2020)

infrastructure. Thus, the key focus was extraction.[5]

Past work has shown that forms of indirect rule are associated with worse development outcomes (Banerjee and Iyer, 2005, Acemoglu, Reed and Robinson, 2014). The concession era contributed to the proliferation of indirect rule. We find evidence consistent with Mamdani's hypothesis that indirect rule has created less accountable and worse leaders. By coopting local institutions, replacing uncooperative leaders with compliant ones, and creating ruling lineages, the rubber concessions instituted a series of local strong men, who continue to dominate village politics today.

There is growing theoretical evidence that culture responds to institutions and that understanding this process is important for understanding persistence (Bisin and Verdier, 2017, Acemoglu and Robinson, 2019, Nunn, 2020). By combining survey and field work with our historical analysis, we are able to empirically examine the role of both institutions and culture. Additionally, the results speak to the importance of "place" for intergenerational mobility (Alesina, Hohmann, Michalopoulos and Papaioannou, 2020, Chetty and Hendren, 2018, Chetty, Hendren and Katz, 2016). We show that development outcomes for second-generation migrants converge once in Gemena, despite that they maintain different cultural norms.

Finally, we provide the first quantitative evidence on the effects of an historical event of significant magnitude. This itself is an important contribution. After the slave trade, the concessions granted under King Leopold are one of the most important events in modern African history. Joseph Conrad, author of *Heart of Darkness*, describes this era as "the vilest scramble for loot that ever disfigured the history of human conscience and geographical exploration".[6] We find that the rubber concessions granted by Leopold II have large and significant negative effects on economic development. This finding is relevant for the many parts of sub-Saharan Africa, South America, and Asia that were granted as concessions to private companies during the colonial era.

The paper is organized as follows. Section 2 provides historical background on the Congo Free State and the rubber concessions. Section 3 presents the main empirical results from the DHS data. Section 4 describes the fieldwork data and results. Section 5 concludes.

[5] This is also related to the literature on the long-run effects of labor coercion, a common element of colonial extraction in Africa and a common feature of labor relations for much of human history (van Waijenburg, 2015, Dell, 2010, Nunn, 2008, Dippel, Greif and Trefler, 2017, Acemoglu and Wolitzky, 2011).

[6] This is related to literature on the effects of mass killings, such as the Holocaust, the Rwandan genocide, and the expulsion of the Moriscos (Acemoglu, Hassan and Robinson, 2011, Rogall and Yanagizawa-Drott, 2014, Chaney and Hornbeck, 2016).

2. The History of the Rubber Concessions

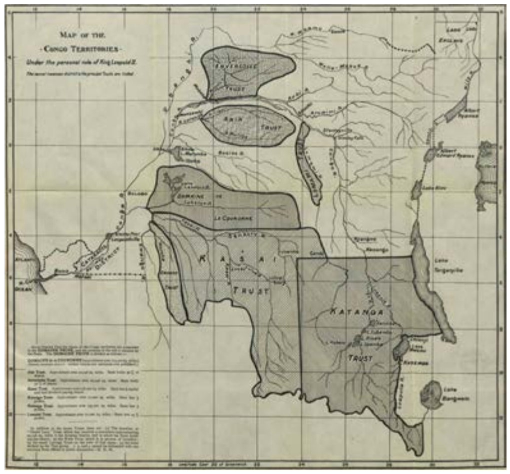
By the mid-1870s, European powers had made claims to most parts of Africa. However, the center of Africa remained largely unexplored. In a bid to make Belgium a colonial power, King Leopold II of Belgium convinced other European colonial powers of his philanthropic goals in Congo, including his mission to end the slave trade. The British, French, and German governments acquiesced to Leopold's interest in Congo to avoid conflict with each other over their own colonial aspirations. Thus, the CFS was created in 1885 as the personal colony of Leopold. According to the Berlin conference in which the borders of the CFS were outlined, Congo was to remain a free trade zone for individuals of all nationalities.

2.1. Concessions in the Congo Free State

Leopold needed to demonstrate continued state presence in the Congo in order to retain his rights over it. This proved a costly endeavor. In 1891 and 1892, in an attempt to increase revenues and contrary to the spirit of the Berlin agreement, he declared all lands and any raw materials found on these lands to be the property of the CFS. This decree divided Congo into three areas. The first area was the *domaine privé*, which was property of the state. Areas of the *domaine privé* were divided into concessions given to private companies. The two largest concessions granted in the *domaine privé* were Anglo-Belgian India Rubber Company (Abir) and Anversoise (Waltz, 1918). An additional part of the *domaine privé* was allocated as private land for the king himself, called the *domaine de la couronne*. A second area, called the "closed area," was to be settled as circumstances allowed. Most of this area was eventually allocated to the Katanga Company in the southwest. The rest of the country was primarily a "free trade zone" where individuals of any nationality could engage in trade. The Kasai region in the South and Southeast remained open to free trade until 1902, when the Kasai trust was established. See Figure 1 for a map of the concessions as of 1904.

The administration of the various areas of the CFS varied depending on whether they were part of a concession, the concession's timing and duration, and the natural resources present in the area. The Abir and Anveroise concessions were the largest focusing on the collection of rubber

Figure 1: 1904 Map of Concessions Granted By Leopold II



Note: The two most northern concessions are Anversoise and Abir.

and existed for 14 years, from 1892 to 1906.[7] The Abir and Anversoise concessions differed from other concessions in that their borders were defined by the extent of river basins, their borders do not coincide with present day political boundaries, they existed for a short period of time, and the concessions focused almost exclusively on the collection of rubber, which required little to no investment. While most of the paper focuses on the Abir and Anversoise concessions, we return to an examination of all of the concessions granted during the CFS era in Section 3.5.

2.2. *Creation of Abir and Anversoise*

Abir and Anversoise were created in the Upper Congo Basin shortly after the invention of the pneumatic tire in 1890, which lead to a dramatic increase in the demand for natural rubber. The Upper Congo Basin had immense natural rubber resources, and Leopold finally saw an opportunity for profits. The state had limited manpower and capacity, so Leopold established concessions to be given to private companies for the exploitation of rubber.

Because most of the interior of the Congo was uncharted at the time, the concession boundaries were defined using salient geographic characteristics, namely major rivers and their basins (Harms, 1975). The contracts establishing the agreements between the CFS and Abir and Anversoise confirm that salient geographic characteristics determined the concession boundaries. Abir was established in 1892 and given rights over the Maringa-Lopori basin. This concession area was

[7] The Kasai area was partially under the free trade regime, then part of a concession company from 1902 to the mid-1950s. The Katanga area was part of a concession, though the extraction focused primarily on copper, rather than rubber.

defined by two rivers and their tributaries: the Maringa river and the Lopori river, plus a 25 km buffer area around them.[8] In the same year Anversoise was created and given extraction rights in the Mongala river basin, defined by the Mongala river and its tributaries.[9] Figure 2a presents the boundaries of the Abir and Anversoise concessions.[10]

To see that the boundaries of the concessions do conform to the definitions as stated in the founding contracts, Figure 2b illustrates the concession boundaries and the associated river basins. The concession borders appear to align almost exactly with the extent of the river basins. Additionally, Figure 2b shows the locations of the rubber collection posts established by the concession companies. The posts all fall within the boundaries of the concessions. In return for the land granted to the concession companies, the state would collect 2% of the companies' profits. Leopold himself was a majority stake holder in Abir and Anversoise (Harms, 1975). Areas just outside of the concessions continued to be free trade zones, but individuals trading in these areas did not have the same rights and resources granted to the concession companies.

2.3. *Rubber Collection*

The concession companies forced individuals within their concessions to collect rubber as a form of paying taxes. The collection of rubber required little capital investment, in contrast to the collection of other natural resources such as diamonds or minerals, and it did not require training of the labor force. The intensity of rubber extraction in concession areas was thus linked to the supply and productivity of labor.

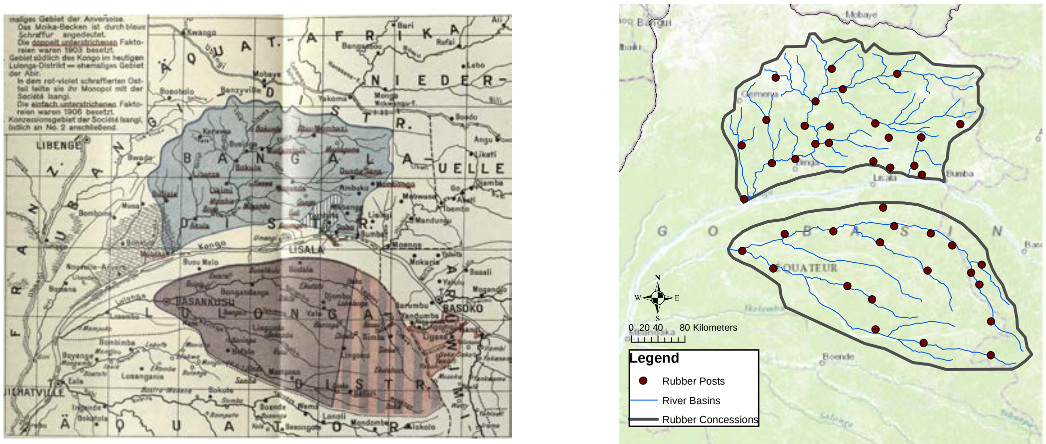
---

[8] The initial contract between the Secretary of the Interior of the CFS, Mr. Eetvelde, and Mr. J.T. North and Alexis Mols, representatives of the Société Anonyme Anglo-Belgian-India-Rubber and Exploration Company defines the boundaries of Abir as follows: "The State of Congo concedes to the undersigned on the other part under the conditions stated in this contract and for a period of 30 years starting today, the right to exploit rubber, gum copal and other products of the forest situated on state lands in the basin of the Lopori and the Maringa, from and including Basakusu and to include the forest situated in an area of 10 kilometers around this post. The state will provide all facilities for such exploitation that will be with the assistance of the District Commissioner and at the sole risk and peril of the concessionary" (Waltz, 1918, p. 372). Article 4 of the document specifies rights to an area of 25 km around each post.

[9] This concession was defined as the area north of part of the Congo River up to the former international border between the CFS and French Equatorial Africa. The initial contract between the Secretary of the Interior of CFS, Mr. Eetvelde, and Mr. Alexander de Browne de Tiège, representative of Anversoise defines the boundaries of the Anversoise concession as follows: "The Congo State accords to the undersigned on the other part, under the conditions indicated in the present contract and for a term of 50 years starting today...the concession of the forests in the state land situated in the basin of the Mongala, with the exclusive right to exploit the rubber, gum copal, and all the other products of the forest" (Waltz, 1918, p. 352).

[10] The maps of the rubber concessions are from Waltz (1918), which describes all of the concessions given by King Leopold II. This includes details on the physical boundaries of the concessions and the year when each concession was granted. Figure 2a is a map of the concessions of interest: Abir and Anversoise. These were the largest concessions in the Upper Congo Basin, and the largest concessions that focused exclusively on rubber (Vangroenweghe, 1985).

Figure 2: Abir and Anversoise Concession Boundaries



(a) Abir and Anversoise Concessions

(b) Concessions, Posts, and River Basins

Notes: The Anversoise rubber concession is the northern concession and the Abir concession is the southern concession (Waltz, 1918). The Mongala river basin is the northern basin and the Maringa-Lopori river basin is the southern basin.

Once the rubber concessions were allocated, the companies set up posts within the concessions to collect rubber. One or two European agents would be assigned to each post within a concession. They would survey surrounding villages and make a census of the number of adult men in the village. Concession companies set quotas for the collection of rubber based on these population censes. Male villagers were required to deliver a quota of about four kilos of dried rubber every two weeks. In addition, villages were required to provide food and supplies to maintain nearby posts (Harms, 1983, 1975).

Most rubber collected during the CFS era was from the vine *landolphia*, which is delicate and easily damaged, rather than from the more hearty rubber trees, *funtumia elastica*, which were more prevalent in the French Congo and West Africa (Harms, 1975). Rubber collection was both time intensive and physically demanding. The rubber collection process could take days, particularly as rubber supplies dwindled and untapped rubber vines became more difficult to find. For example, men in the Baringa area would spend around 10 days of every 14 days in the forest collecting rubber (Harms, 1983).

2.4. *Violence*

The concession companies maintained militias comprised of sentries who were responsible for ensuring compliance with the rubber quotas. Generally, the sentries were outsiders recruited from other areas of Congo; this strategy was purposefully selected to ensure that sentries were willing to use violence against villagers. Approximately 25 to 80 "post sentries" armed with rifles were assigned to each new post established. An additional 65 to 100 "village sentries," armed with muzzle-loading cap guns, were stationed in the villages surrounding the posts. In 1903, one Abir post received 17,600 cartridges for the Albini rifles used by the post sentries (Harms, 1983). To prevent waste, soldiers were required to provide a human hand for every bullet used. The human hands were then smoked for preservation and collected by the European agents.

Individuals were severely punished if they failed to meet their rubber quota. Punishment could take many forms. For example, individuals could be imprisoned and forced to work. Their family members could be held for ransom until the quota was fulfilled. Individuals could also be subjected to various forms of physical violence, including whipping by the *chicotte* (a whip made of hippopotamus hide), burning with gum copal, or death. The chief of the village could also be imprisoned if his village did not meet the quota. In July 1902, records indicate that 44 chiefs were imprisoned in the villages around a single post (Harms, 1983). The sentries from the concession companies' private militias were primarily responsible for carrying out these violent tactics. However, the European agents also engaged in the imprisonment, torture, and killing of villagers.

Testimony collected by Robert Casement, a consul for the British sent to Congo to investigate accusations of atrocities, documents the intensity of the violence:

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

If the sentries faced any resistance, they were able to call on soldiers from the *Force Publique* to provide support.

2.5. *Indirect Rule*

A tactic employed by sentries to ensure rubber production was to undermine and coopt local authority. Non-compliant chiefs were held captive, replaced, or killed. One of the sentries in each village was assigned the position of *kapita*, or head sentry for that village. In fact, *kapita* is a

Lingala word used today to denote "chief". Once in the village, the *kapita* would recruit eight to ten people to serve as bodyguards. He then began the process of asserting his authority over the villagers. To do so, he would attack men in positions of esteem or authority. For example, lineage headman were required to carry soil and rubbish alongside slaves. Anyone who challenged the *kapita* could be whipped or killed.

The *kapitas* severely undermined the prestige, authority, and wealth of lineage headmen and village chiefs. The village headmen were "shamelessly degraded in the eyes of their people, made to fetch and carry for soldiers, cast into chains and flung into prison" (Morel, 1904). Though they were still considered to have important connections to ancestors, the headmen no longer had the authority to make important decisions. They were unable to protect their lineage from the brutality and terror imposed by the sentries. Additionally, since most able-bodied men were required to collect rubber in the forest, there was a power vacuum in the village that was filled by the *kapita*. In fact, some sentries began to take on the responsibilities previously allocated to lineage headmen, such as settling disputes among lineage members. Finally, the sentries would take the wealth from lineage headmen, including marrying their daughters and wives (Harms, 1974).

2.6. *Aftermath*

Though the CFS government objected in principle to the violence, in practice it allowed and encouraged it. The effectiveness of the labor coercion allowed the concession companies to make exorbitant profits (Plas and Pourbaix, 1899). By 1905, the natural rubber supplies were nearly exhausted in the Upper Congo Basin. Due to depleted rubber supplies and increasing condemnation of their labor practices in Europe, Abir and Anversoise left CFS in 1906. In 1908, the CFS became a Belgian colony. After 1910, competitive production of rubber from *hevea* plantations in Southeast Asia and South America, along with the invention of synthetic rubber, led to a large decrease in rubber prices (Harms, 1975).

Historians have noted that the rubber concessions granted under Leopold II had disastrous consequences for local populations. Villages subjected to the rubber regime were unable to tend to their fields, leading to low yields and famine. Sentries raided local livestock. Malnourished individuals became particularly susceptible to disease, including the increasingly rampant sleeping sickness (Harms, 1983). As Hochschild describes, "the world has managed to forget one of the great mass killings of recent history... it was unmistakably clear that the Congo of a century ago had indeed seen a death toll of Holocaust dimensions" (Hochschild, 1998, pp. 3-4). The brutality of the rubber collection tactics resulted in the deaths of an estimated 10 million people and earned the policies the nickname "*Red Rubber*" (Vangroenweghe, 1985).

Since the end of the rubber era, there have been three key periods: Belgian colonial rule (1908 to 1960), the Mobutu era (1965 to 1997), and the present. Across the Equateur region, the Belgian era was characterized by little investment, forced cultivation of cash crops (coffee, palm oil, and cotton), forced labor (particularly during the World Wars), missionary provision of education, and sleeping sickness campaigns (Clement, 2014). Mobutu pursued policies of centralization (reducing the number of provinces); returning to cultural "authenticity" (banning Western influences); and "Zairinization" in which foreign enterprises were expropriated and eventually taken over by the state. Corruption was rampant and the government invested only 1.5% of its budget in the Equateur region (Hesselbein, 2007). Since Mobutu's overthrow, the Equateur region has experienced civil conflict and unrest. Equateur remains one of the poorest regions of Congo. Most individuals are engaged in subsistence agriculture; the non-agricultural sector is close to non-existent (Van Reybrouck, 2014).

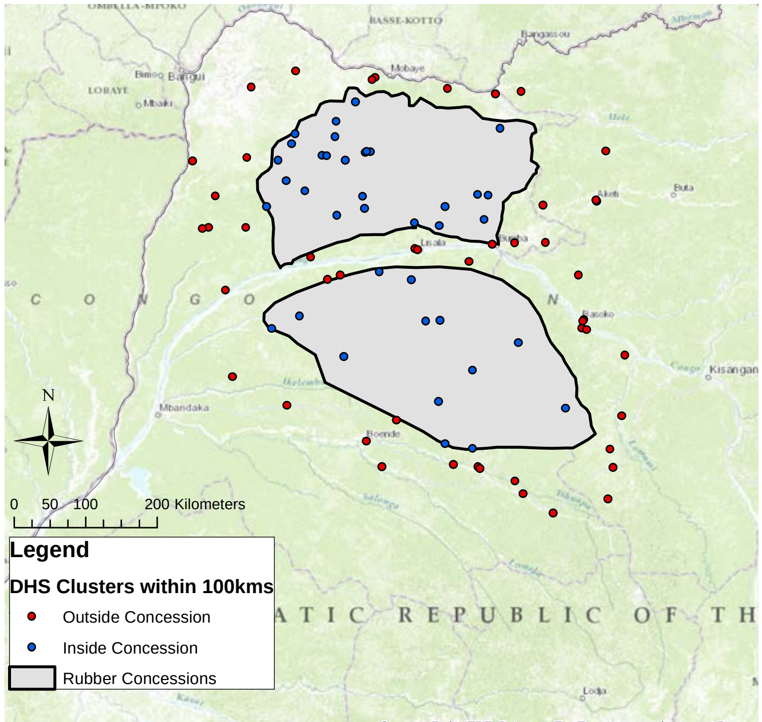
3. The Effects of the Rubber Concessions on Development

3.1. Data

To examine the long run impact of the rubber concessions on development outcomes we combine Demographic and Health Survey (DHS) data from 2007 and 2014 with maps of the boundaries of Abir and Anversoise. The DHS surveys from the DRC provide detailed information on education, assets, and health outcomes. These data sources and the variables used in our analyses are described in detail in Appendix A.[11] Figure 3 provides a map with the rubber concession borders and the DHS clusters from 2007 and 2014 that are within 100 kms of the borders of the rubber concessions. Appendix C.1 provides summary statistics for clusters in and outside of the former concessions.

[11] In Appendix I we also explore other datasets including data on population density (Landscan and African Population Database); nightlights; and conflict (PRIO).

Figure 3: Maps of Clusters from the DHS 2007 and 2014 for DRC - Within 100 kms



3.2. *Empirical Strategy*

We can estimate the causal effect of exposure to the rubber concessions on the outcomes of interest with the following regression discontinuity (RD) specification:

$$y_{i,v} = \alpha + \gamma RubberConcession_{i,v} + f(location_v) + \mathbf{X}_i\beta + \phi + \varepsilon_{i,v} \quad for \quad v \in bw \tag{1}$$

where $y_{i,v}$ is our outcome of interest for individual i from village v; $RubberConcession_{i,v}$ is an indicator equal to 1 if v is inside a rubber concession area and equal to 0 otherwise; $\mathbf{X}_i$ is a vector of covariates for individual i such as gender, age, and age squared; ϕ is a nearest concession fixed effect; and $f(location_v)$ is the RD polynomial, which controls for smooth functions of geographic location for village v. Following Calonico, Cattaneo and Titiunik (2014), Cattaneo, Idrobo and Titiunik (2020), and Gelman and Imbens (2016), our baseline specification is a local linear polynomial in distance to the concession border estimated separately on each side of the concession. We use a triangular weighting kernel and calculate the optimal bandwidth bw using the MSE-minimizing procedure suggested by Cattaneo et al. (2020). We also present results with a wider fixed bandwidth of 100 kms from the border.[12] We check robustness to using various other forms of the RD polynomial, kernel, and bandwidths (see Appendix C).

Our coefficient of interest is γ: the effect of being just inside the concession area on our outcome of interest. The intuition behind this specification is that concession borders arbitrarily allocated some villages to be part of the concessions and others to be just outside the concessions. These villages should have similar geography, culture, history, and institutions prior to the concession era, allowing us to identify the effect of rubber extraction on contemporary outcomes.[13]

The RD approach presented in equation (1) requires two identifying assumptions. The first assumption is that all relevant factors before the concessions were granted varied smoothly at the concession boundaries. This assumption is needed to ensure that individuals located just outside the concessions are an appropriate counterfactual for those located just inside them.

A key concern for identification is that the rubber concessions were chosen strategically for certain characteristics that could also affect our outcomes of interest. For example, these areas might have been more suitable for certain crops or populated by ethnic groups with different cultures. However, whether an area was exposed to rubber extraction is a deterministic and discontinuous function of whether or not a village fell inside the concession boundaries. As described in Section 2, these concessions were granted at a time when much of the Congo had not been explored. The concession boundaries were defined by salient geographic characteristics - in this case, rivers and river basins. Thus, the concession boundaries are unlikely to have been selected based on local characteristics that also vary discontinuously at the concession border.

To assess the plausibility of this first assumption, Table 1 presents summary statistics and estimates using specification (1) for important geographic and pre-concession characteristics. This analysis is at the 20km by 20km grid cell level. These results are presented with standard errors clustered at the territory level for geographic characteristics and the ethnicity level for cultural characteristics; we also present results using Conley standard errors with a cut-off window of 50 kms to account for spatial auto-correlation (Conley, 1999) in Appendix C.2 and with wild

[12] We calculated the Imbens and Kalyanaraman optimal bandwidth for several of our outcomes of interest with distance to the border as the running variable. The optimal bandwidth was generally between 75 and 125 kilometers depending on the outcome. Thus, we chose our fixed bandwidth to be 100 kms.

[13] This RD approach has been used in multiple settings to examine the effects of historical events, such as in: Dell (2010), Miguel and Roland (2011), Grosfeld, Rodnyansky and Zhuravskaya (2013), Michalopoulous and Papaioannou (2014), Becker, Boeckh, Hainz and Woessmann (2015), Fontana, Nannicini and Tabellini (2016).

bootstrap p-values (Cameron, Gelbach and Miller, 2008) in Appendix C.3.[14] In addition to showing balance at the grid cell level, we also show balance on pre-concession characteristics at the DHS cluster level in Appendix C.4.

Consistent with the first identification assumption, we find balance on a number of important geographic and pre-concession characteristics. The geographic characteristics include: elevation, rainfall, rainfall variability, land suitability, ruggedness, river characteristics, disease suitability measures, and distance to Kinshasa and the coast.[15] We find balance on these characteristics, suggesting that the areas inside and outside the concession are comparable along the border.

Finally, we also show balance on a number of pre-concession characteristics using data from the Ethnographic Atlas (EA) (Murdock, 1967) and Nunn and Wantchekon (2011). Ideally we would present balance on pre-concession characteristics for all ethnic groups near the concession. However, the EA only contains information for a sub-sample of groups in our area of interest. Using the sub-sample with information in the EA, we find balance on a number of pre-concession characteristics such as exposure to the Atlantic slave trade, population density prior to colonization, levels of centralization, and polygamy. Additionally, the concession borders do not align with Murdock ethnic group borders (see Appendix Figure B4) nor do they align with present day political borders.

The second important assumption for the RD approach is that there was no selective sorting across the RD threshold when the concession borders were established. Selective sorting would require certain villages be able to select out of being allocated to a concession. This is unlikely to have happened given that villages were unable to negotiate the boundaries of the concessions.

An important related concern is selective migration either during the rubber era or subsequently, which would be considered an outcome of the rubber concessions.[16] It is likely that some migration took place during the rubber era, as individuals tried to avoid the rubber demands and the associated violence. Unfortunately, there are no data available to quantify the magnitude of

[14] For the clustered standard errors, we cluster at the territory level, the lowest administrative level for which there is spatial data. For RD bandwidths below 50 kms, the number of clusters is slightly below thirty, potentially leading to overly optimistic standard errors (Cameron et al., 2008). Overall, the clustered standard errors tend to be quite consistent with (and more conservative than) both the wild bootstrap and Conley standard errors.

[15] Note, rubber in the CFS era was from a vine called *landolphia owariensis*, rather than from rubber trees. We do not have data on the suitability for the rubber vine, which was found in forested areas. Reassuringly, the entire Congo basin region is heavily forested, and this does not align with the concession boundaries. Appendix Figure B2 visually presents the extent of river networks.

[16] By selective migration, we mean migration of only the "highest ability" individuals from just inside the concessions to just outside the concessions.

Table 1: Balance on Geographic and Pre-Concession Characteristics

	Within 100 km			*Within 50 km*			*RD Estimates*	
	Inside	Outside	SE	Inside	Outside	SE	RD Coefficient	SE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Geographic Characteristics:								
Elevation	437.26	430.24	(8.10)	440.06	432.46	(5.87)	2.76	(17.39)
Rainfall (Avg.)	72.49	76.42	(4.08)	70.43	74.19	(3.22)	0.01	(7.59)
Rainfall (St. Dev.)	1.31	1.43	(0.07)	1.32	1.41	(0.07)	-0.02	(0.16)
Land Suitability	5.78	7.62	(1.64)	5.01	7.84	(1.94)	-2.26	(2.42)
Ruggedness	4.94	6.25	(1.02)	5.31	5.74	(0.87)	0.26	(2.24)
Obs.	349	504		232	272		894	
River Characteristics:								
Navigable River Density	10.07	10.51	(2.58)	12.20	9.19	(3.41)	1.64	(4.48)
>0 Navigable Rivers	0.21	0.22	(0.05)	0.24	0.19	(0.07)	0.04	(0.08)
>0 Rivers	0.44	0.51	(0.05)	0.38	0.46	(0.07)	-0.11	(0.09)
Obs.	349	504		232	272		894	
Disease Characteristics:								
Malaria Suitability	18.76	18.75	(0.24)	18.71	18.84	(0.19)	-0.23	(0.29)
TseTse Fly Suitability	1.33	1.34	(0.01)	1.32	1.33	(0.01)	0	(0.02)
Obs.	349	504		232	272		894	
Location Characteristics:								
Distance: Kinshasa	767.30	792.81	(37.53)	764.17	776.68	(25.17)	21.82	(58.51)
Distance: Coast	1093.71	1047.54	(19.33)**	1082.01	1064.67	(11.35)*	-18.23	(26.47)
Obs.	349	504		232	272		894	
Pre-Concession Characteristics:								
Num. Enslaved (Atlantic Trade, 1000s)	0.65	2.40	(1.66)	0.91	4.24	(3.22)	-3.92	(6.41)
Obs.	236	314		159	170		573	
Num Ethnic Groups	11	23		10	17		24	
Population Density	1.41	1.19	(0.19)	1.36	1.18	(0.12)	0.09	(0.19)
Obs.	121	187		74	89		329	
Num Ethnic Groups	3	7		3	6		7	
Centralization	0	0.21	(0.10)	0	0.14	(0.07)	-0.04	(0.08)
Obs.	124	280		103	148		426	
Num Ethnic Groups	5	11		5	10		11	
Polygynous	0.46	0.47	(0.11)	0.49	0.45	(0.11)	0.06	(0.21)
Obs.	247	322		173	176		593	
Num Ethnic Groups	7	13		7	12		13	
Hereditary Local Headman Selection	1	0.90	(0.10)	1	0.96	(0.04)	-0.04	(0.03)
Obs.	204	240		142	139		462	
Num Ethnic Groups	5	11		5	10		11	

Notes: The unit of observation is a 20 by 20 km grid cell. Columns 1, 2, 4, and 5 present the mean of the corresponding variable. Columns 3 and 6 present clustered standard errors for the difference in means clustered at either the territory level for geographic variables or ethnic group level for pre-colonial variables. Inside and Outside indicate whether a grid cell is inside or outside the former rubber concession area respectively. Columns 7 and 8 give the estimated RD coefficient and standard error that uses the corresponding variable as its outcome using a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. The RD MSE optimal bandwidth is determined using the procedure suggested by Cattaneo et al. (2020). Column 7 uses the average of all optimal bandwidths (39.30 kms). Regressions include a nearest concession fixed effect. Variable definitions and data sources used in this analysis are described in detail in Appendix A. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

migration during the rubber era.[17] While we are unable to analyze migration during the rubber era, we conduct a number of tests using present-day DHS data and fieldwork data to show that selective migration is unlikely to fully explain the effects. In Appendix F, we examine the sensitivity of the results to selective migration and to heterogeneity by ease of migrating from inside the concession to outside the concession boundaries. Rates of selective migration would have to be quite high to fully explain our results, and there is no evidence of differential effects based on ease of migration.

[17] We can only highlight the difficulties associated with migration. Anecdotal evidence from Harms (1975) suggests that the rubber companies greatly controlled migration (using village censuses they collected themselves) and forced people to remain in their villages. Harms (1975) notes that local chiefs were held accountable when individuals that migrated did not meet their quotas, incentivizing chiefs to prevent migration. Finally, since the concessions were defined by the extent of river basins, and rivers were used for transport, migration outside of the concessions would likely have been difficult.

3.3. *Regression Discontinuity Results*

To examine the long-run effects of exposure to the rubber concessions, we analyze 2007 and 2014 DHS data on education, wealth, and health. All variables are defined in the table notes. Table 2 reports estimates for specification (1) with education outcomes in Panel A, asset wealth outcomes in Panel B, and health outcomes in Panel C. Section 3.4 discusses additional RD polynomials and other robustness checks. The results in Table 2 Panel A show that areas inside the concession have significantly lower levels of education across all specifications and bandwidths. Individuals just inside the former rubber concessions are estimated to have approximately 1.3 fewer years of education than individuals just outside the concessions.

Table 2 Panel B reports estimates for the asset wealth measures available in the DHS survey.[18] Individuals in villages inside the former rubber concessions are approximately 25% less wealthy than similar individuals outside the rubber concessions. In standard deviation terms, areas inside the former concessions are about 0.5 standard deviations less wealthy.

Table 2 Panel C reports estimates for different health outcomes and finds evidence that individuals from inside the former concessions have worse health outcomes. Children and women inside the former concessions have approximately 9 percentage points lower height-to-age percentile. Overall, we find evidence that individuals residing in villages inside the former rubber concessions are less educated, less wealthy, and have worse health outcomes today than individuals in villages outside the former rubber concessions.

Figure 4 presents standard RD plots for our main outcomes of interest, with distance to the border as the running variable and a local linear trend to each side of the discontinuity. For these outcomes, we observe there is a clear discontinuity at the concession border. See Appendix B for the spatial RD plots.

3.4. *Robustness of DHS Results*

There are three main empirical concerns for the DHS results presented in Table 2: robustness to alternative RD specifications, random displacement of DHS clusters, and the use of basins to define borders. The first concern is whether the results are robust to alternative specifications of the RD-polynomial. In Appendix C.5, we show that the results are robust to alternative bandwidths,

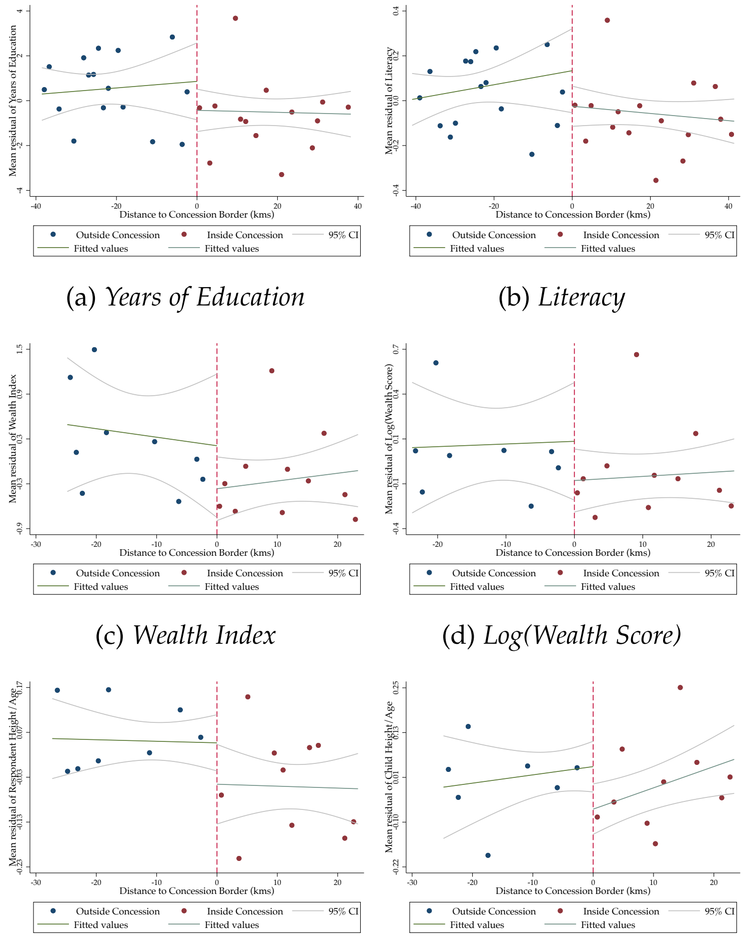
[18] To reduce the influence of outliers and to ease interpretation, we show wealth results using the log of the wealth factor score.

Table 2: Rubber Concessions and Economic Development

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	−2.061**	−1.396**	−0.205**	−0.150**
	(0.842)	(0.635)	(0.089)	(0.069)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	2,661	42,157	2,847	4,266
Clusters	44	85	47	85
Bandwidth	40.34	100.00	41.85	100.00
Mean Dep. Var.	5.060	5.109	0.448	0.465
SD Dep. Var.	3.754	3.821	0.497	0.499
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.913*	−0.686**	−0.412**	−0.259**
	(0.478)	(0.306)	(0.191)	(0.126)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	1,816	4,281	1,816	4,281
Clusters	27	85	26	85
Bandwidth	25.04	100.00	24.31	100.00
Mean Dep. Var.	2.150	2.034	10.948	10.912
SD Dep. Var.	1.119	1.060	0.453	0.443
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.147***	−0.093**	−0.148***	−0.085**
	(0.054)	(0.038)	(0.038)	(0.033)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	745	1,422	754	1,314
Clusters	32	85	32	85
Bandwidth	29.20	100.00	28.96	100.00
Mean Dep. Var.	0.278	0.259	0.267	0.247
SD Dep. Var.	0.262	0.252	0.331	0.316

Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 4: RD Plots for Development Outcomes



(a) *Years of Education* (b) *Literacy*

(c) *Wealth Index* (d) *Log(Wealth Score)*

(e) *Respondent Ht./Age* (f) *Child Ht./Age*

Notes: The figure presents RD plots for our main outcomes and the mean value of each outcome variable at each 2.5 km bin along the running variable (distance to concession border) as well as with a local linear trend estimated separately on each side of the discontinuity. Each regression is estimated using the optimal bandwidth chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020). Regressions controls for age, age squared, gender, survey year, and include a nearest concession fixed effect. Standard errors are clustered at the DHS cluster level and the figures show 95% confidence intervals.

bandwidth selection procedures, and kernels (e.g uniform, triangular, and Epanechnikov). In Appendix C.6 we show that the results are robust to using local quadratic polynomials. We also test robustness to latitude and longitude specifications in Appendix C.7, where we modify $f(.)$ in equation (1) to be a function of latitude and longitude.

A second potential issue is that the DHS randomly displaces the coordinates of the clusters in order to maintain the confidentiality of the respondents. The GPS coordinates are displaced by up to 5 km for all urban clusters and 99% of rural clusters, and up to 10 km for 1% of rural clusters. Importantly, this displacement is random and induces classical measurement error. This would bias our coefficient towards zero. However, with the regression discontinuity approach, one might be concerned that the results are being driven by clusters right along the border that might be incorrectly assigned to inside or outside the concession because of the random displacement. Thus, we estimate our regression discontinuity results with a "donut-hole" of 10 kms (the maximum possible displacement) in Appendix C.8 and find that the results are robust to excluding observations close to the border.

A third possible concern is that because the concession borders were drawn using major river basins as the salient geographic feature for the borders, the results may reflect some inherent characteristic of river basins, rather than exposure to the concessions. To assess this claim, in Appendix D we conduct a falsification exercise where we run our main specification across all major river basins in DRC and show that the estimated effects for the former concessions are larger and more negative than the estimated effects for most of the other major river basins in DRC. The falsification exercise presents important evidence that the results are not a consequence of the concessions being drawn using river basins, but instead suggests that our estimates represent the impact of exposure to rubber concessions. Additionally, to address concerns that the concession borders may have been manipulated, Appendix D shows the estimated effects are similar when we use the relevant basin borders and a 25 km buffer as an instrument for being inside the former concessions.

In Appendix C, we show our results are robust to the following additional robustness tests. We analyze the results: looking at each concession individually to ensure that the results are not being driven by one particular concession (C.9); without covariates (e.g. gender, age, and age squared) (C.10); with Conley standard errors (C.11); at the DHS cluster level, rather than the individual level since assignment to treatment occurs at the village level (C.12); dropping observations along the Congo river to address concerns that villages along the Congo river are different than those farther away from the river (C.13); and dropping one DHS cluster at a time (C.14). In Appendix E we find no evidence of differential missionary presence or subsequent colonial investment, and in Appendix I, we examine the relationship between the concessions and road network density, population density, nighttime lights, and conflict.

Finally, in Appendix I we present two additional complementary analyses. First, in I.4 we examine the correlation between intensity of exposure to the rubber concessions – as measured by year a post is constructed and the number of tons of rubber collected at that post – and our key outcomes of interest. We find that greater intensity of exposure is negatively correlated with education, wealth, and health. We also test whether there appears to be convergence in outcomes overtime across age cohorts in I.5. We find no evidence for convergence.

3.5. *Analysis for All Concessions*

We have found that those individuals from the former Abir and Anversoise concession areas have lower levels of education, wealth, and health today. Given that other concessions were granted during the CFS era, it is natural to examine the broader implications of the concession system for the development of DRC.

In Appendix J, we present RD results for education, wealth, and health examining all concessions granted in DRC as of 1904 (see Figure 1 for a map of all of the concession boundaries). We present results pooling all of the concessions as well as results excluding Abir and Anversoise. We find that across all concessions in DRC, individuals experience worse education, wealth, and health outcomes. The coefficients are always negative, though sometimes not significant when Abir and Anversoise are excluded. Given that the 60% of DRC's landmass that was formally part of a concession, wealth would be about 30% higher had they not been part of a concession. While these estimates are unlikely to be causal, given that these other concession boundaries correspond with present day political boundaries and have different histories than Abir and Anversoise, they are suggestive of the detrimental long-run legacy of the concession system.

4. Testing Mechanisms with Fieldwork

4.1. *Conceptual Framework*

Our analysis so far has documented lasting negative effects of the rubber concessions granted in Northern Congo on development outcomes. These concessions were characterized by indirect rule, violence, and no productive investment – characteristics that were common across the many concessions granted during the colonial era (Coquery-Vidrovitch, 1972, Vangroenweghe, 2006). However, the concessions only lasted 14 years; thus, it is important to understand the mechanisms through which the negative effects of the rubber concessions persist.

Our conceptual framework consists of three main premises and is informed by the recent theoretical literature on how institutions and culture matter for development (e.g. Bisin and Verdier, 2017, Acemoglu and Robinson, 2019, Nunn, 2020, Persson and Tabellini, 2020). This literature takes as a starting point that institutions are an important factor for explaining economic development (e.g. North, 1990, Acemoglu et al., 2001). However, a key insight from this emerging literature is that one cannot understand historical persistence by focusing only on institutions; rather, it is important to also understand how culture evolves and endogenously responds to institutions. Informed by work from historians and oral histories, we apply these insights to our setting and discuss conditions when the framework is most likely to be relevant.

The first premise is that the rubber era undermined local institutions via the use of indirect rule. A key aspect of the way the rubber concessions were able to function was through the cooption of local leaders (Marchal, 1996, Harms, 1974, 1975, 1983, Stengers and Vansina, 1985). Prior to the concession era, lineage headman held positions of authority and leadership within communities. These leaders either cooperated with the rubber regime or were replaced by individuals who would. As Stengers and Vansina (1985, p. 336) write, "Chiefs were auxiliary agents of the state. They did not need to have any traditional legitimacy".

Compliance with the rule of coopted leaders was achieved through extreme violence. This violence was often perpetrated by armed sentries recruited from outside the area with the support of European agents. Local institutions suffered, and leaders could no longer be relied upon to protect the community or fulfill important leadership functions (Vangroenweghe, 1986, Harms, 1974).

Thus, we will examine how local institutions differ inside the former concessions. Specifically, we will examine how chiefs are selected – either hereditary or through elections – and whether chiefs are responsible for public goods provision.

The second premise of our conceptual framework is that cultural traits likely responded to these changes in local institutions.[19] Bisin and Verdier (2017) highlight that certain cultural traits will respond endogenously to changes in institutions and that ex-ante, it is unclear how culture may respond (see also Tabellini, 2008). Under certain conditions, changes in culture will reinforce changes in local institutions, leading to persistence, while in others, cultural change will counteract institutional change.[20]

Historians have highlighted how villagers had to develop alternative coping mechanisms as they faced a brutal rubber regime and local leaders who were unable to protect them (Harms,

[19] Culture is generally defined as values and preferences that are internal to individuals in a society (Tabellini, 2008).

[20] This insight is related to a growing literature on how institutions can affect culture (e.g. Guiso, Sapienza and Zingales, 2004, Tabellini, 2010, Guiso, Sapienza and Zingales, 2016, Lowes, Nunn, Robinson and Weigel, 2017).

1974, Nelson, 1994, Marchal, 1996). The rubber period required communities to adapt and cooperate in order to survive. In particular, there was an increased reliance on horizontal ties (e.g. age sets) and mutual insurance. This allowed individuals to bypass corrupt and ineffective elders who had been targeted by the rubber regime.[21]

Changes in the political power structure may have led to important changes in cultural norms related to pro-social behavior, cooperation, and trust. Therefore, we examine how pro-social norms and behavior differ inside the former concessions.

The final premise of the conceptual framework is that the interaction between institutions and culture is important for understanding the long-run effects of the rubber era. As modeled in Bisin and Verdier (2017), institutional quality and culture may act as substitutes, enforcing an equilibrium of low institutional quality but more pro-social norms (see also Jha, 2013, Platteau, 2000).[22] If culture and institutions are substitutes, stronger pro-social norms may undermine long-run development if they allow less effective and less accountable local leaders to remain in power. We explore correlational evidence consistent with this premise using fieldwork data.[23]

We discuss when these mechanisms are likely to be present. Four main conditions appear central. First, the conceptual framework applies to concessions characterized by indirect rule, violence, and a focus on extraction. Second, violence was generally perpetrated by outsiders – either European agents or sentries recruited from outside of the region. Third, there was minimal productive investment (unlike in Dell and Olken, 2020). The concessions granted across large parts of sub-Saharan Africa generally met these three conditions (Coquery-Vidrovitch, 1972, Vangroenweghe, 2006). Finally, initial state capacity is likely important. The Northern concessions had little state centralization prior to colonization; anecdotally, ethnic groups with more centralization (e.g. in the South of Congo) were better able to resist concession companies and their violence (Stengers and Vansina, 1985). Additionally, the Congolese state since independence has been weak and ineffective at projecting power outside of urban areas.

[21] For example, the rubber period was associated with an increase in cooperation among individuals of same age grade. These "pacts of friendship and mutual aid between age-mates facilitated the social mobility required in the search for rubber" (Nelson, 1994, p.110), as people were often forced to collect rubber in groups far away from their village.

[22] As in Bisin and Verdier (2017), we view institutional change as a "myopic" short-run process. In contrast, culture evolves more slowly over time, through socialization and cultural transmission.

[23] Bisin and Verdier (2017, p.38) write that their theoretical work "underlines the fact that the search for a unique origin for long-term development can be quite an arduous and even sterile undertaking. Focusing more systematically on the positive or negative interactions between culture and institutions along the development process might be more fruitful in terms of historical understanding".

4.2. Data Collection

Existing data from DRC does not allow us to examine institutional or cultural mechanisms. Therefore, we conducted surveys and collected experimental data in Gemena, DRC. Gemena is the capital of Sud-Ubangi province and is situated near the border of the former Anversoise concession. Gemena is inside the former concession boundary, but less than 10 km away from the border. Gemena was created by colonial administrators in the mid-1920s, after the CFS period, and therefore consists primarily of migrants from surrounding areas. Nearly all individuals in our sample identify their "village of origin" as a village outside the town of Gemena. A "village of origin" is the village where an individual's family or ancestors are from.[24]

The data were collected between July and August 2015. We randomly sampled individuals in Gemena for a total sample size of 520 individuals for the first visit and 484 for the second visit. We divided our survey into two visits per household to avoid survey fatigue; the first visit consisted of the main survey module and second visit consisted of lab experiments and a short survey. Of those sampled, 49.71% percent identified their village of origin as being from inside the boundaries of one of the former concessions and a total of 511 originate from villages within 200 kms of the former concession boundaries. Figure 5 presents a map of the locations of villages of origin for our sample, the location of Gemena, and the borders of the former rubber concessions. For more details on sampling and survey methods, see Appendix G.1.

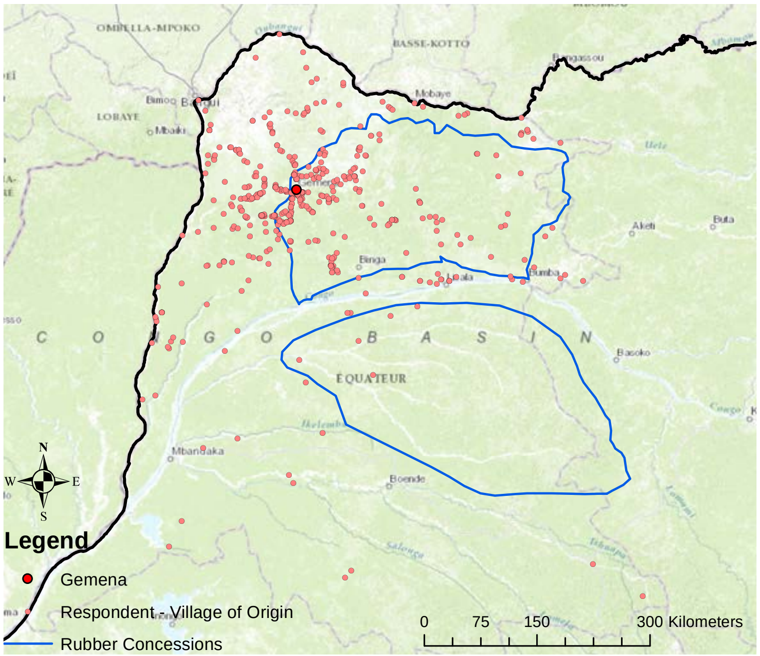
We use the data from the Gemena-based sample to compare individuals with ancestors from inside the former concessions to those with ancestors from outside the former concessions. This approach has two main advantages. First, logistically, it is considerably easier to work in one main town rather than numerous villages in the area, as transportation infrastructure is of very poor quality. Second, it allows us to more precisely identify cultural differences: by examining individuals removed from their original institutional environments and who now share the same institutional environment, any differences in behavior in experimental measures or responses to survey questions are capturing differences in internalized cultural norms.[25]

Individuals answered a series of questions on demographics, migration history, income, trust, and beliefs. In addition to collecting individual level data, we ask respondents detailed questions

[24] This is a commonly understood concept in this area, and all respondents knew their village of origin. A village of origin is not necessarily where an individual is born.

[25] This is the epidemiological approach as described by Fernández (2011). This approach was employed in e.g. Lowes et al. (2017) and Alesina, Giuliano and Nunn (2013).

Figure 5: Gemena, the Rubber Concessions, and Location of Origin Villages within the Sample



about the institutions in their villages of origin. Individuals who were familiar with their village of origin were asked questions on the selection mechanism for the village chief, public goods available in their villages of origin, and the responsibilities of the chief. Finally, individuals completed two behavioral experiments and an Implicit Association Test (IAT), which are described in detail below.

The survey data has multiple questions that could be used to test the hypotheses of interest. We present all of our survey-based results using thematic indices that group related questions. We follow Anderson (2008) and compute the inverse covariance weighted (ICW) index across outcomes within an index. In Appendix G.7 we include coefficient plots of each of the individual components of the index in addition to the estimated ICW index coefficient. As with the previous RD analyses, we estimate equation (1).

Summary statistics are presented in Appendix G.5 by whether an individual originates from inside the former concession. On average, individuals from inside the concession have fewer years of education and lower income than those from outside the concession, but these differences are not statistically significant. Additionally, there is no relationship between being from inside the former concessions and being knowledgeable about one's village of origin, which mitigates concerns about differential knowledge of villages of origin. A possible concern with data collected in Gemena is differential selective migration based on whether an individual is from the former concession area. In Appendix Table G3 we present mean differences on key migration characteristics for individuals from inside and outside the former concessions. We find very little evidence of differences in reasons for migration for individuals from inside and outside the concession.[26]

4.3. *Economic Development in Villages of Origin*

We first examine whether villages of origin within the former concessions are less developed than those just outside the former borders using our survey data from Gemena. This analysis is similar to Section 3.3 where we use DHS data to test for differences in development outcomes. We asked individuals about the public goods available in their villages of origin and their perception of the relative wealth of their village of origin. Panel A of Table 3 presents the coefficients for two ICW indices: an index of public goods available in the village of origin and an index of a respondent's subjective measures of the development of their village of origin. All questions included in the index and their response options are reported in the notes of the table. Villages within the former concession are described as having fewer public goods and are rated as less developed.

4.4. *Differences in Village Chief Selection and Performance*

We turn to the first premise in the conceptual framework: the rubber era may have undermined local institutional quality. The historical accounts of the rubber period and the oral histories from individuals in Gemena suggest that the position of chief may have been affected by the rubber period (Young, 1965, Vanhee, 2005). Today, village chiefs are tasked with organizing public good maintenance and construction, resolving conflict, and welcoming outsiders. If the rubber regime altered the accountability and quality of village chiefs, this could explain the worse development outcomes we observe inside the former concessions.

Selection and Performance of Village Chiefs

Panel B of Table 3 presents the results on the chief selection mechanism. Chiefs in villages inside the former rubber concessions are 35 percentage points less likely to be chosen by election.

[26] Because it is difficult to rule out selective migration with the Gemena sample we also corroborate our main fieldwork findings using baseline data for another project that was collected from 300 villages in the region (Lowes, Montero, Nunn and Robinson, 2016) in Appendix H.

Table 3: Rubber Concessions and Village Institutions

	Panel A: Village Development			
	Village Public Goods Index		*Village Subjective Ratings Index*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.127^{*}	-0.115^{**}	-0.880^{***}	-0.434^{**}
	(0.074)	(0.055)	(0.252)	(0.176)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	267	304	147	212
Clusters	136	221	87	162
Bandwidth	34.40	100.00	22.55	100.00
Mean Dep. Var.	-0.003	0.003	0.062	0.068
SD Dep. Var.	0.285	0.291	0.740	0.690
	Panel B: Chief Quality and Accountability			
	Chief Elected		*Chief Public Good Index*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.423^{**}	-0.345^{***}	-0.277^{**}	-0.194^{*}
	(0.182)	(0.129)	(0.136)	(0.106)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	189	255	250	296
Clusters	104	189	119	210
Bandwidth	26.85	100.00	28.26	100.00
Mean Dep. Var.	0.503	0.506	-0.083	0.016
SD Dep. Var.	0.502	0.501	0.519	0.494
	Panel C: Respect for Authority			
	Survey Questions Index		*Chief IAT Score*	
	(1)	(2)	(3)	(4)
Inside Concession	0.136	0.197	0.165	0.026
	(0.258)	(0.149)	(0.176)	(0.091)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	274	465	259	417
Clusters	142	313	127	285
Bandwidth	21.66	100.00	21.17	100.00
Mean Dep. Var.	0.068	0.003	-0.091	-0.088
SD Dep. Var.	0.739	0.750	0.566	0.549

Notes: Standard errors are clustered at the origin village level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions include nearest concession fixed effects and control for age, age squared and sex. *Village Public Goods Index* is a summary index for the following questions (with the number of components for each question in brackets): (1) What material is the road in your village of origin made of? [2: 0=Sand, 1=Gravel or Pavement] (2) Is your village of origin on a main road? (3) Does your village of origin have a secondary school? [2] (4) Does your village of origin have a Health Dispensary? [2] (5) Does your village of origin have a Hospital? [2] (6) Does the water in your village of origin come from a well? [2: 0=Spring water, 1=Well]. *Village Subjective Ratings Index* is a summary index for the following questions (with the number of components for each question in brackets): (1) How would you rate the quality of the primary school in your village of origin? [5] (2) How would you rate the quality of the secondary school in your village of origin? [5] (3) How would you rate the quality of the road in your village of origin relative to other roads in the area? [5] (4) Relative to other villages in the area you have visited, how would your rate your village of origin overall? [5] *Chief Public Good Index* is a summary index for the following questions: Is the chief in your village of origin responsible for providing (1) road maintenance, (2) new roads, (3) school maintenance, (4) land allocation, (5) protection of property rights, (6) tax collection, (7) jobs, (8) conflict arbitration, and (9) road brushing; all questions answered as a 0 to 2 categorical variable where 0 is No, 1 is Partially, and 2 is Yes. *Chief Elected* is an indicator variable equal to 1 if the village chief of a respondent's origin village is selected by elections. *Respect for Local Authority Index* is a summary index for the following questions: (1) How much do you trust your village of origin chief? [4], (2) How much do you trust your sub-tribe chief? [4], (3) How satisfied are you with your village of origin chief? [4], (4) Would you vote for your village of origin chief if there were an election held tomorrow? [2], (5) How much confidence do you have in local chiefs? [4]. *Chief IAT Score* is the D-Score for the Implicit Association Test that asked respondents to sort sounds of words related to local chief authority, where more positive values indicate a more positive implicit association with local chiefs. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Instead, they are more likely to be hereditary, i.e. chosen from a particular lineage within the community. This lineage is known as the "ruling" lineage, and chiefs then tend to come exclusively from this lineage.

To examine whether there are differences in the quality of chiefs, we construct an index that combines all questions on whether chiefs are responsible for providing specific public goods (and their maintenance) in the villages of origin; a lower value on this index suggests chiefs are of lower quality in the sense that they are not considered responsible for providing key public goods at the village level. We find that chiefs inside the former concessions are responsible for providing fewer public goods.

Respect for Chief Authority

An important consideration when examining differences in village institutions is to account for differences in respect for authority. If respect for chief authority is lower inside the concessions due to the rubber concession period, then local chiefs may be less able to organize productive activities, resolve conflicts, and provide order, even if the chiefs themselves are of the same quality.

To examine respect for village chief authority, we first construct an index of subjective survey questions on confidence and trust in chiefs. We scale all variables so that more positive values indicate greater respect for local chiefs. Because respondents may be unwilling to answer potentially sensitive questions about local political figures truthfully, we also conducted a Single-Target Implicit Association Test (ST-IAT) to measure implicit attitudes towards chiefs.[27]

Panel C of Table 3 reports the estimates from these two different measures of respect for authority. Individuals from inside the former concessions report that they respect chiefs more in the subjective index, though the results are not statistically significant. Likewise, the IAT results do not providence evidence of different levels of respect for authority. Overall, the measures of respect for authority in Panel C suggest that the result of lower public goods provision by chiefs inside the former concessions is not driven by lack of respect for authority.

4.5. Differences in Trust, Social Cohesion, Altruism, and Support for Sharing

The second premise of the conceptual framework is that the rubber regime led to changes in cultural traits. The historical accounts suggest that exposure to the rubber regime affected a

[27] See Appendix: IATs for more information on the ST-IAT and its implementation.

series of important outcomes related to cooperation and sharing by increasing the importance of and reliance on mutual insurance. We examine differences in trust, feeling of closeness with others, and survey and experimental measures of support for sharing.

Trust in Others

We examine whether trust is different across the former concession borders in Panel A of Table 4 by constructing an index of questions on how much individuals trust various people.[28] The coefficients on trust inside the former concessions are positive and statistically significant, suggesting that individuals from the former concessions are in fact more trusting than those outside the former concessions. The coefficient plot for each question individually is presented in Appendix G.7. We also ask respondents how close they feel to people to various groups of people. We present the results on differences in closeness in Panel A of Table 4. We find that individuals from the former concessions report feeling closer to a wide variety of other individuals.

Strength of Beliefs in Importance of Sharing

To test whether there are differences in beliefs in the importance of sharing, we first construct an index of survey questions asking individuals whether they think it is appropriate to share income in a variety of different situations. The index includes questions on whether you and others should share income when it is earned by luck and when it is earned by work. We also ask the respondent how they think people in their village of origin would respond to the same series of questions to understand their expectations regarding the beliefs of others.

Panel B of Table 4 present the estimates for each of these measures. Individuals from the former concessions are more likely to agree that income should be shared with others. They are also more likely to report that individuals in their villages of origin would also agree that income should be shared. Individuals support sharing income regardless of whether it is earned by work or luck and regardless of whether they are speaking about sharing their own income with others or others sharing with them.

We also collected experimental measures of support for sharing. Individuals in our sample participated in a dictator game (DG) to measure altruism and in a reverse dictator game, to

[28] We chose these survey questions following work by Johnson and Mislin (2011) and Johnson and Mislin (2012) who demonstrate that trust survey questions have a positive, robust correlation with experimental measures of trust (i.e. amount sent in the trust game).

Table 4: Survey and Experimental Measures of Trust and Sharing Beliefs

	Panel A: Trust and Closeness			
	Trust Index		*Closeness Index*	
	(1)	(2)	(3)	(4)
Inside Concession	0.299	0.313***	0.521**	0.551***
	(0.236)	(0.120)	(0.211)	(0.135)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	304	465	299	465
Clusters	151	313	158	313
Bandwidth	23.61	100.00	24.95	100.00
Mean Dep. Var.	0.142	0.041	0.156	0.045
SD Dep. Var.	0.698	0.707	0.717	0.717
	Panel B: Survey Measures of Sharing Norms			
	Respondent		*Village of Origin*	
	(1)	(2)	(3)	(4)
Inside Concession	0.521**	0.339*	0.489**	0.268*
	(0.255)	(0.184)	(0.229)	(0.143)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	305	453	285	363
Clusters	156	304	144	255
Bandwidth	26.14	100.00	28.09	100.00
Mean Dep. Var.	0.068	0.031	0.012	0.006
SD Dep. Var.	0.835	0.825	0.785	0.783
	Panel C: Experimental Measures of Sharing Norms			
	Dictator Game: Share Sent		*Effort Task: Share Redistributed*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.018	0.001	0.067**	0.053**
	(0.025)	(0.020)	(0.030)	(0.023)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	258	438	368	437
Clusters	135	300	193	300
Bandwidth	21.48	100.00	37.12	100.00
Mean Dep. Var.	0.449	0.445	0.410	0.405
SD Dep. Var.	0.122	0.123	0.134	0.134

Notes: Standard errors clustered at the origin village level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions include nearest concession fixed effects and control for age, age squared and sex. *Trust Index* is a summary index for the following questions: How much do you trust (1) people from your village of origin, (2) people of another tribe, (3) people of your own tribe, (4) people you meet for the first time, (5) your family, (6) your neighbors, (7) people of another nationality, and (8) people of your sub-tribe; all questions answered on a 0 (Not at All) to 4 (Completely) scale. *Closeness to Others Index* is a summary index for the following questions: (1) How close to you feel to people from your village of origin?, (2) How close to do you feel to people of Gemena?, (3) How close do you feel to people of your own tribe?, (4) How close do you feel to people of your age set from your origin village?, and (5) How close do you feel to people of your age set in Gemena?; all questions answered in a scale from 0 (Not Close at All) to 5 (Very Close). *Sharing Norms Index* is a summary index for the following questions: (1) If you get money from luck you should share it, (2) If you earn money from hard work you should share it, (3) If someone else earns money from luck they should share it, (4) If someone else earns money from hard work they should share it; all questions answered in a scale from 1 (Strongly Disagree) to 5 (Strongly Agree). *Sharing Norms Index Village of Origin* is a summary index for the following questions, where all questions start with "How much would someone from your village of origin agree with the following statements", for the same statements listed above. *Dictator Game: Amount Shared* measures the amount sent to an anonymous player 2 in the standard Dictator Game. *Effort Task: Share Redistributed* is the total share taken (weighted by the maximum budget amount possible to take) in the effort task from the anonymous player 1's earned income. It represents an experimental measure of respect for earned income property rights. Two individuals declined participating in the Dictator Game, and one additional individual declined participating in the Reverse Dictator Game. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

measure support for redistribution. In the standard DG, a player 1 is given an endowment and is asked to allocate it between themselves and a player 2. The reverse DG differs in two key ways from the standard DG. First, the player 1 earns an endowment by completing a task.[29] Second, the player 2 is told how much the player 1 earned and is asked what share of the player 1's earned income they would like to keep for themselves. The amount player 2 decides to take from Player 1's earned income therefore represents a measure of willingness to redistribute.[30] See Appendix G.2 for more details on the reverse DG with earned income, the protocols used, and a description of the earnings task.

Panel C of Table 4 presents the estimates for the experimental measures of altruism and willingness to redistribute. We find no significant differences in amount sent in the dictator game, thought the coefficient on inside concession is positive. For the reverse dictator game, we find that individuals from the former concessions redistribute a larger share of the other player's earned endowment to themselves. We interpret this as having greater support for redistribution, consistent with the survey measures on sharing that suggest individuals think income should be shared.[31]

One implication of greater support for sharing income is that we would expect villages inside the former concessions to have less income inequality. Consistent with greater support for sharing income, in Appendix I.3, we find lower levels of income inequality, as measured by both the standard deviation of the wealth score and the inter-quartile range of the wealth score, within DHS clusters inside the former concessions.

The results in Table 4 provide evidence that individuals from inside the former concessions are more trusting, feel closer to others, believe it is important to share income, and redistribute more in a reverse dictator game. The results all point to more pro-social beliefs and values within

[29] For the earnings task, we selected a task that could be easily understood by all respondents and for which more effort was rewarded by more income. Subjects played a "clicking-game" on touch screen tablets. In this "clicking-game," a small blue dot appears in a random location on the screen every three seconds and the respondent has one second to push the dot before it disappears. Importantly, this effort task did not rely on physical strength or skill but instead relied on concentration and perseverance. The game lasted five minutes and respondents were paid based on the number of successful "clicks," earning 100 Congolese Francs (approximately $0.10) per 10 successful clicks. Respondent were very engaged in the task and earned on average 700 CF in this task.

[30] Variation (i) of the DG has been used before by Hoffman, McCabe, Shachat and Smith (1994) and Cherry, Frykblom and Shogren (2002); subjects tend to be much less generous when they earned their own income, which Farh and Irlenbusch (2000) refer to as *earned property rights.* Variation (ii) on its own changes the standard DG to what is known as a Reverse DG, which has been used many times before (List, 2007). Jakiela (2011) combines these two variations to get a measure of respect for earned property rights and finds that subjects in the US tend to others' respect earned income much more than subjects in Kenya.

[31] Work by Platteau (2000) discusses how there may be a dark side to sharing and redistribution norms. For example, they can dampen incentives for investment.

the former concessions.

Note, our priors before data collection had been that the rubber concessions may have led to both worse institutions and less pro-social norms. These priors were informed by work by Nunn and Wantchekon (2011), who find that the slave trade undermined trust. Thus, while potentially surprising, our results are in line with recent work by Bauer et al. (2016), who review findings from 16 post-conflict settings and have found that individuals exposed to conflict are more pro-social because they are more dependent on local informal systems of risk-sharing and insurance.[32] Our results suggest that in response to weakened institutions and violence, cultural traits related to cooperation and mutual aid changed.

4.6. Discussion of Results

An important question is: how do the changes in institutions and culture lead to a low-development equilibrium? Of course, this cannot be directly tested in our data, but a compelling explanation underlying our conceptual framework is that these changes in institutional quality and culture reinforce each other: chiefs are held less accountable and allowed to stay in power since individuals do not rely on their formal institutions as much and instead rely on informal sharing norms for support. Indeed, consistent with the conceptual framework, in Appendix Figure I1, we find a negative correlation between pro-social norms and a chief's provision of public goods; a positive correlation between chief's provision of public goods and wealth; and a negative correlation between pro-social norms and wealth. We interpret this as suggestive evidence that more pro-social norms serve as a substitute for better quality institutions.

Our fieldwork data also lets us explore whether the cultural changes on their own are sufficient to lead to low development outcomes, or whether they only lead to less development when paired with low quality local institutions. We do this by comparing first and second generation migrants from in and outside the concessions. These individuals are removed from their village institutions and share a common institutional environment presently.

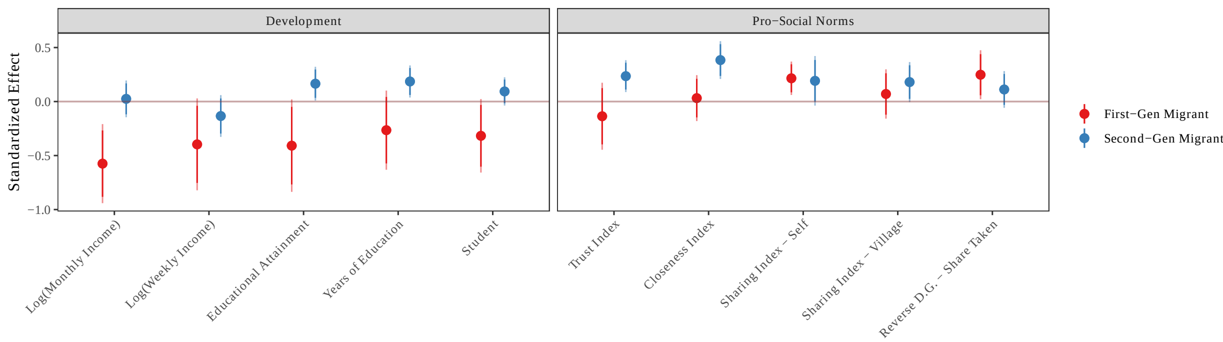
We first explore whether migrants converge in terms of development outcomes when they share a common institutional environment in Figure 6. We find differences in education and

[32] This finding is also related to the "tend-and-befriend" hypothesis in psychology, which suggests people build alliances during times of stress (see von Dawans et al., 2012).

income that are quite stark if we look only at first generation migrants.[33] However, if we examine second generation or higher migrants separately, we find convergence in these outcomes.

We then turn to the cultural outcomes. The results in Figure 6 suggest that both first and second generation migrants exhibit more pro-social norms. This has important implications for understanding if our observed effects are "place" or "person" specific. It suggests that removing individuals (and their cultural norms) from the former concession areas actually leads to relatively quick convergence in education and wealth outcomes.

Figure 6: Development and Pro-Social Norms By First and Second Generation Migrants



Notes: Figure presents estimates from estimating specification (1) for the sub-sample for first-generation migrants (in red) and second-generation or higher migrants (in blue) to Gemena. Data is from the survey conducted in Gemena. Pro-social norm indexes are defined as in Table 4. *Educational Attainment* is a 0 to 3 categorical variable, where 0=No schooling, 1=Primary Schooling, 2=Secondary Schooling, and 3=Post-Secondary Schooling. Standard errors clustered at the origin village level. Figure presents 95% and 90% confidence intervals (with and without transparency, respectively). Regressions include nearest concession fixed effects and control for age, age squared, sex, latitude and longitude. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel using a 100 km bandwidth from the former concession borders.

5. Conclusion

We examine how exposure to the concession system, which was characterized by violence and indirect rule, has shaped the development of areas of Congo. The rubber concession period was characterized by its extreme brutality and violence and the use of local leaders to achieve rubber production quotas. We provide micro-level evidence on how these common colonial "treatments" – concessions, violence, and indirect rule – matter for understanding the comparative development of sub-Saharan Africa.

This study documents that former rubber concession areas have lower levels of education, wealth, and health than areas outside of the concessions. A likely mechanism is that inside the former concessions, chiefs are less likely to be elected, are more likely to be hereditary, and

[33] Interestingly, the differences are of very similar magnitude to the DHS results we present in Section 3.3.

provide fewer public goods. Despite the negative effects on development and local institutions, we also find evidence that individuals inside the former concessions are more pro-social.

The results highlight that even a short exposure to extractive institutions can have a meaningful impact on the development of an area, particularly when local institutions are integrated into supporting the extraction. It is important to keep in mind our counterfactual; most of Congo was exposed to various forms of violence and colonial extraction. Thus we are capturing the effect of relatively greater exposure, suggesting that we are providing a lower bound of an effect relative to no extraction.

We present the first quantitative evidence on the effects of this common form of economic organization and on an important historical event. These results suggest that concessions, where private companies are given state-like powers for the purpose of generating profit, can have significant and negative long term consequences. These results have important implications for the development of the DRC and other parts of Africa, much of which was granted as concessions during the colonial era.

References

Académie Royale des Sciences d'Outre-Mer, *Atlas Général du Congo et du Ruanda-Urundi*, Brussels: Académie Royale des Sciences d'Outre-Mer, 1954.

Acemoglu, Daron and Alexander Wolitzky, "The Economics of Labor Coercion," *Econometrica*, March 2011, *79* (2), 555–600.

_ and James A. Robinson, *The Narrow Corridor: States, Societies, and the Fate of Liberty*, New York, USA: Penguin Press, 2019.

_ , Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 2001, *91* (5), 1369–1401.

_ , Tarek A. Hassan, and James A. Robinson, "Social Structure and Development: A Legacy of the Holocaust in Russia," *Quarterly Journal of Economics*, 2011, *126*, 895–946.

_ , Tristan Reed, and James A. Robinson, "Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone," *Journal of Political Economy*, 2014, *122* (2), 319–368.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn, "On the Origins of Gender Roles: Women and the Plough," *Quarterly Journal of Economics*, 2013, *128* (2), 469–530.

_ , Sebastian Hohmann, Stelios Michalopoulos, and Elias Papaioannou, "Intergenerational Mobility in Africa," 2020. Forthcoming, *Econometrica*.

Alsan, Marcella, "The Effect of the TseTse Fly on African Development," *American Economic Review*, January 2015, *105* (1), 382–410.

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

Army Map Services, Corps of Engineers, U.S. Army, "Series 2201, Kisangani Map, Sheet 23," 1968.

Banerjee, Abhijit and Lakshmi Iyer, "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India," *American Economic Review*, 2005, *95* (4), 1190–1213.

Bauer, Michal, Christopher Blattman, Julie Chytilová, Joseph Henrich, Edward Miguel, and Tamar Mitts, "Can War Foster Cooperation?," *Journal of Economic Perspectives*, 2016, *30* (3), 249–274.

Becker, Sascha O., Katrin Boeckh, Christa Hainz, and Ludger Woessmann, "The Empire is Dead, Long Live the Empire! Long Run Persistance of Trust and Corruption in the Bureaucracy," *The Economic Journal*, February 2015, *126* (590), 40–74.

Bisin, Alberto and Thierry Verdier, "On the Joint Evolution of Culture and Institutions," 2017. NBER Working Paper No. 23375.

Cage, Julia and Valeria Rueda, "The Long-Term Effects of the Printing Press in sub-Saharan Africa," *American Economic Journal: Applied Economics*, 2016, *8* (3), 69–99.

_ and _ , "Sex and the Mission: The Conflicting Effects of Early Christian Investments on sub-Saharan Africa's HIV Epidemic," July 2017.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik, "Robust Data-Driven Inference in the Regression-Discontinuity Design," *Stata Journal*, 2014, *14* (4), 909–946.

Cameron, Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik, *A Practical Introduction to Regression Discontinuity Designs: Foundations*, Cambridge, U.K.: Cambridge University Press, 2020.

Chaney, Eric and Richard Hornbeck, "Economic Dynamics in the Malthusian Era: Evidence from the 1609 Spanish Expulsion of the Moriscos," *Economic Journal*, 2016, *126* (594), 1404–1440.

Cherry, Todd L., Peter Frykblom, and Jason F. Shogren, "Hardnose the Dictator," *American Economic Review*, 2002, *94* (4), 1218–1221.

Chetty, Raj and Nathaniel Hendren, "The Impact of Neighborhoods on Intergenerational Mobility II: County-Level Estimates," *Quarterly Journal of Economics*, August 2018, *133* (3), 1163–1228.

_ , _ , and Lawrence F. Katz, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, April 2016, *106* (4), 855–902.

Clement, Piet, "Rural Development in the Belgian Congo. The Late-colonial "Indigenous Peasantry" Programme and its Implementation in the Equateur District (1950s)," *Bulletin des Séances. Académie Royale des Sciences d'Outre Mer*, 2014, *60* (2), 251–286.

Conley, Timothy, "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, *92*, 1–45.

Coquery-Vidrovitch, Catherine, *Le Congo au temps des grandes compagnies concessionnaires 1898-1930*, La Haye, Mouton, 1972.

de Rouck, René, *Atlas Géographique et Historique du Congo Belge et Des Territoires sous Mandat du Ruanda-Urundi* 1945.

Dell, Melissa, "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 2010, *78* (6), 1863–1903.

_ and Benjamin A. Olken, "The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java," *The Review of Economic Studies*, 2020, *87* (1), 164–203.

Dippel, Christian, Avner Greif, and Dan Trefler, "Outside Options, Coercion, and Wages: Removing the Sugar Coating," 2017.

Farh, René and Bernd Irlenbusch, "Fairness as a Constraint on Trust in Reciprocity: Earned Property Rights in a Reciprocal Exchange Experiment," *Economic Letters*, 2000, *66* (3), 275–282.

Fernández, Raquel, "Does Culture Matter?," in Matthew O. Jackson Jess Benhabib and Alberto Bisin, eds., *Handbook of Social Economics*, Vol. 1A, Elsevier, 2011, chapter 11, pp. 481–510.

Fontana, Nicola, Tommaso Nannicini, and Guido Tabellini, "Historical Roots of Political Extremism: Effects of Nazi Occupation of Italy," March 2016.

Gelman, Andrew and Guido Imbens, "Why High-Order Polynomials Should not be Used in Regression Discontinuity Designs," 2016. NBER Working Paper No. 20405.

Gennaioli, Nicola and Ilia Rainer, "The Modern Impact of Precolonial Centralization in Africa," *Journal of Economic Growth*, September 2007, *12* (3), 185–234.

Grosfeld, Irena, Alexander Rodnyansky, and Ekaterina Zhuravskaya, "Persistent Antimarket Culture: A Legacy of the Pale Settlement after the Holocaust," *American Economic Journal: Economic Policy*, August 2013, *5* (3), 189–226.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "The Role of Social Capital in Financial Development," *American Economic Review*, 2004, *94* (3), 526–556.

_ , _ , and _ , "Long-Term Persistence," *Journal of the European Economic Association*, 2016, *14* (6), 1401–1436.

Harms, Robert, "Abir: The Rise and Fall of a Rubber Empire," March 1974. Masters Thesis University of Wisconsin.

_ , "The End of Red Rubber: A Reassessment," *Journal of African History*, 1975, *16* (1), 73–88.

_ , "The World Abir Made: The Maringa-Lopori Basin, 1885-1903," *African Economic History*, 1983, *12*, 125–139.

Hesselbein, Gabi, "The Rise and Decline of the Congolese State: An Analytical Narrative on State-Making," 2007. Crisis States Research Centre, Working Paper No. 21.

Hochschild, Adam, *King Leopold's Ghost: A Story of Greed, Terror, and Heroism in Colonial Africa*, Houghton Mifflin, 1998.

Hoffman, Elizabeth, Kevin McCabe, Kevin Shachat, and Vernon Smith, "Preferences, Property Rights, and Anonymity in Bargaining Games," *Games and Economic Behavior*, 1994, *7* (3), 346–380.

Huillery, Elise, "History Matters: The Long-Term Impact of Colonial Public Investments in French West Africa," *American Economic Journal: Applied Economics*, 2009, *1* (2), 176–215.

Imbens, Guido and Karthik Kalyanaraman, "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *The Review of Economic Studies*, July 2012, *79* (3), 933–959.

Jakiela, Pamela, "Social Preferences and Fairness Norms as Informal Institutions: Experimental Evidence," *American Economic Review: Papers & Proceedings*, 2011, *101* (3), 509–513.

Janssens, E., "Africa archives, papers D1366 11/22/1904," 1904. Federale Overheidsdienst Buitenlandse Zaken.

Jha, Saumitra, "Trade, Institutions and Ethnic Tolerance: Evidence from South Asia," *American Political Science Review*, November 2013, *107* (4).

Johnson, Noel D. and Alexandra Mislin, "Trust Games: A Meta-Analysis," *Journal of Economic Psychology*, 2011, *32*, 865–889.

_ **and** _ , "How Much Should We Trust the World Values Survey Trust Question?," *Economic Letters*, 2012, *116* (2), 11–44.

Juif, Dácil and Ewout Frankema, "From Coercion to Compensation: Institutional responses to labour scarcity in the Central African Copperbelt," *Journal of Institutional Economics*, 2017.

La Porta, Raphael, Florencio López de Silanes, Andrei Shleifer, and Robert Vishny, "Law and Finance," *Journal of Political Economy*, 1998, *106* (6), 1113–1155.

List, John, "On the Interpretation of Giving in Dictator Games," *Journal of Political Economy*, 2007, *115* (3), 482–493.

Lowes, Sara and Eduardo Montero, "The Legacy of Colonial Medicine in Central Africa," 2020.

_ , _ , Nathan Nunn, and James A. Robinson, "Age Organization & Accountability: Evidence from the Democratic Republic of Congo," 2016. AEA RCT registration number: *AEARCTR-0005370*. J-PAL GI Funding Request. See https://www.povertyactionlab.org/evaluation/age-organization-accountability-evidence-democratic-republic-congo.

_ , Nathan Nunn, James A. Robinson, and Jonathan Weigel, "The Evolution of Culture and Institutions: Evidence from the Kuba Kingdom," *Econometrica*, 2017, *85* (4), 1065–1091.

Mamdani, Mahmood, *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*, Princeton: Princeton University Press, 1996.

Marchal, Jules, *L'Etat Libre du Congo: Paradis Perdu. L'Histoire du Congo 1876-1900*, Borgloon: Bellings, 1996.

Méndez-Chacón, Esteban and Diana Van Patten, "Multinationals, Monopsony, and Local Development," July 2020.

Michalopoulos, Stelios and Elias Papaioannou, "Historical Legacies and African Development," *Journal of Economic Literature*, 2020, *58* (1), 53–128.

Michalopoulous, Stelios and Elias Papaioannou, "Pre-colonial Ethnic Institutions and Contemporary African Development," *Econometrica*, 2013, *81* (1), 113–152.

_ and _ , "National Institutions and Subnational Development in Africa," *Quarterly Journal of Economics*, 2014, *129* (1), 151–213.

Miguel, Edward and Gérard Roland, "The Long-Run Impact of Bombing Vietnam," *Journal of Development Economics*, September 2011, *96* (1), 1–15.

Morel, Edmund Dene, *King Leopold's Rule in Africa*, London: W. Heinemann, 1904.

Murdock, George Peter, *Ethnographic Atlas*, Pittsburgh, PA: University of Pittsburgh Press, 1967.

Nelson, Samuel H., *Colonialism in the Congo Basin, 1880-1940*, Columbus: Ohio University Press, 1994.

North, Douglass, *Institutions, Institutional Change and Economic Performance* Political Economy of Institutions and Decisions, New York: Cambridge University Press, 1990.

Nunn, Nathan, "The Long Term Effects of Africa's Slave Trades," *Quarterly Journal of Economics*, 2008, *123* (1), 139–176.

_ , "Religious Conversion in Colonial Africa," *American Economic Review Papers and Proceedings*, 2010, *100* (2), 147–152.

_ , "History as Evolution," August 2020. Forthcoming, Handbook of Historical Economics.

_ and Leonard Wantchekon, "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, 2011, *101* (7), 3221–3252.

Omasombo, Jean Tshonda, *Sud-Ubangi: bassins d'eau et espace agricole* Monographies des provinces de la Républicque Démocratique du Congo, Tervuren: Musée royal de l'Afrique centrale, 2013.

Osafo-Kwaako, Philip, "Disease and Development: The Effect of Life Expectancy on Economic Growth," 2012. Harvard University Thesis, mimeo.

Persson, Torsten and Guido Tabellini, "Culture, Institutions, and Policy," August 2020.

Plas, J. and Victor Pourbaix, *Les Sociétés Commerciales Belges et le Régime Économique et Fiscal de L'état Indépendant du Congo*, Brussels: Van Assche, 1899.

Platteau, Jean-Philippe, *Institutions, Social Norms, and Economic Development*, Reading, U.K.: Harwood Academic, 2000.

Rogall, Thorsten and David Yanagizawa-Drott, "The Legacy of Political Mass Killings: Evidence from the Rwandan Genocide Evidence from the Rwandan Genocide," March 2014.

Sanchez de la Sierra, Raul, "On the Origin of States: Stationary Bandits and Taxation in Eastern Congo," *Journal of Political Economy*, 2020, *128* (1), 32–74.

Stengers, Jean and Jan Vansina, *The Cambridge History of Africa*, Vol. 6 of *The Cambridge History of Africa*, Cambridge: Cambridge University Press,

Tabellini, Guido, "The Scope of Cooperation: Values and Incentives," *Quarterly Journal of Economics*, 2008, *123* (3), 905–950.

_ , "Culture and Institutions: Economic Development in the Regions of Europe," *Journal of the European Economic Association*, 2010, *8* (4), 677–716.

Van Reybrouck, David, *Congo: The Epic History of a People*, New York: Ecco, 2014.

van Waijenburg, Marlous, "Financing the Colonial State: The Revenue Imperative and Forced Labor," 2015. Working Paper: African Economic History Working Paper 20.

Vangroenweghe, Daniel, "Le Red Rubber de L'Anversoise 1899-1900 Documents Inédits," *Annales Aequatoria*, 1985, *6*, 39–65.

_ , *Du sang sur les lianes: Léopold II et son Congo*, Didier Hatier, 1986.

_ , "The 'Leopold II' Concession System Exported to French Congo with as Example the Mpoko Company," *Belgisch Tijdschrift voor Nieuwste Geschiedenis*, 2006, *36* (3-4), 323–372.

Vanhee, Hein, "Over Heren en Knechten: Chefs Médaillés in het Koloniale Congo," in "Het Geueugen van Congo: de Koloniale Tijd," Tervuren: Koninklijk Museum voor Midden Africa, 2005, pp. 79–82.

Vansina, Jan, *Being Colonized: The Kuba Experience in Rural Congo, 1880-1960*, Madison: University of Wisconsin Press, 2010.

von Dawans, Bernadette, Urs Fischbacher, Clemens Kirschbaum, Ernst Fehr, and Markus Heinrichs, "The Social Dimension of Stress Reactivity: Acute Stress Increases Prosocial Behavior in Humans," *Psychological Science*, 2012, *23* (6), 651–660.

Waltz, Heinrich Jakob, *Das Konzessionswesen im Belgischen Kongo* Veroffentlichungen des Reichskolonialmts, No 9, G Fischer Jena, 1918.

Wantchekon, Leonard, Marko Klasnja, and Natalija Novta, "Education and Human Capital Externalities: Evidence from Colonial Benin," *The Quarterly Journal of Economics*, 2015, *130* (2), 703–757.

Young, Crawford, *Politics in the Congo: Decolonization and Independence*, Princeton: Princeton University Press, 1965.

Web Appendix for

Concessions, Violence, and Indirect Rule: Evidence from the Congo Free State

Sara Lowes

UC San Diego, NBER, and CIFAR

Eduardo Montero

University of Michigan

23 September 2020

Contents

List of Tables

List of Figures

Appendix A. Data Sources and Variable Definitions

A.1. Geographic Data and Variables

- **Elevation:** The elevation data is provided by the Global Climate Database created by Hijmans, Cameron, Parra, Jones and Jarvis (2005) and available at http://www.worldclim.org/. This data provides elevation information in meters at the 30 arc-second resolution (approximately at the 1 km^2 level near the equator). The elevation measure is constructed using NASAs SRTM satellite images (http://www2.jpl.nasa.gov/srtm/). Our paper's elevation variable calculates the mean elevation for each 20 km by 20 km grid cell in meters.[34]

- **Precipitation:** Precipitation data is provided by the Global Climate Database created by Hijmans et al. (2005) and available at http://www.worldclim.org/. This data provides monthly average rainfall in millimeters. We calculate the average rainfall for each month for each 20 km by 20 km grid cell and average this over the twelve months to obtain our yearly precipitation measure in millimeters of rainfall per year. We also construct the standard deviation of rainfall over the twelve months to obtain our measure for the variation in rainfall in millimeters of rainfall per year.

- **Soil Suitability:** Soil suitability is the soil component of the land quality index created by the Atlas of the Biosphere available at http://www.sage.wisc.edu/iamdata/ used in Michalopoulos (2012) and Ramankutty, Foley, Norman and McSweeney (2002). This data uses soil characteristics (namely soil carbon density and the acidity or alkalinity of soil) and combines them using the best functional form to match known actual cropland area and interpolates this measure to be available for most of the world at the 0.5 degree in latitude by longitude level. (The online appendix in Michalopoulos (2012) provides a detailed description of the functional forms used to create this dataset.) This measure is normalized to be between 0 and 1, where higher values indicate higher soil suitability for agriculture. Our Soil Suitability variable measures the average soil suitability in each 20km by 20km grid cell to provide a measure of soil suitability that also ranges between 0 and 1, with higher values indicate higher soil suitability for agriculture.

- **Ruggedness:** We use the measure of terrain ruggedness created by Nunn and Puga (2012). This data uses elevation and constructs a terrain ruggedness index as the square root of the sum of the squared differences in elevation between a central point and the eight adjacent points on a grid of 30 arc-seconds. Our ruggedness variable measures the average terrain ruggedness (normalized by 1000) in each 20km by 20km grid cell to provide a measure of ruggedness, with higher values indicate higher terrain ruggedness.

- **Rivers and Navigable Rivers:** The *Referentiel Geographique Commun*, an online repository for GIS maps for DRC, provides shape files for the DRC on all rivers and navigable rivers in DRC as of 2010. Our variables *Access to Navigable Rivers* and *Access to any River* are indicator variables equal to one if the 20 km by 20 km grid cell contains a navigable river or any river, respectively *Navigable River Density* is defined as total length in meters of navigable rivers in each grid divided by the grid's surface area in kilometers squared.

- **Malaria Suitability:** Malaria data uses the Malaria Ecology index created by Kiszewski, Mellinger, Spielman, Malaney, Sachs and Sachs (2004). The index was created by Kiszewski et al. (2004) to approximate the prevalence of severe forms of malaria, and adjusts for the

[34] See Figure B7 for a Map of the Grid Cells.

mosquito type that is most prevalent in a region. We calculate the average of this measure for each 20km by 20km grid cell to provide a malaria suitability.

- **Tsetse Fly Suitability:** The tsetse suitability index (TSI) is from Alsan (2015). The TSI is measured as the Z-score of the potential steady-state tsetse fly population constructed using global climate data. We calculate the average of this measure for each 20km by 20km grid cell to provide a measure of Tsetse Fly Suitability.

- **Distance to Kinshasa:** We calculate the euclidean distance (in kms) from the centroid of each 20km by 20km grid cell to Kinshasa, the capital city of DRC.

- **Distance to the Coast:** We calculate the euclidean distance (in kms) from the centroid of each 20km by 20km grid cell to the ocean.

A.2. DHS Survey Data and Variables

Survey data on development outcomes for individuals is provided by the 2007 DHS survey on the DRC implemented by The DHS Program with the help of the DRC Ministry of Planning. The fieldwork was carried out from January 2007-August 2007 and sampled 9995 women between the ages of 15-49 and 4757 men between the ages of 15-59. The survey provides detailed information on education, assets, and health outcomes for individuals in multiple villages.As well, the DHS 2007 DRC survey provides GPS coordinates for each village (i.e. *clusters* in the survey); these coordinates are displaced by up to 5km for all urban clusters, and 99% of rural clusters and up to 10 km for 1% of rural clusters. Importantly, this displacement is random, and simply induces classical measurement error. The survey data and detailed information on the sampling procedure and variable definitions is available at http://dhsprogram.com/data/Data-Variables-and-Definitions.cfm. Below we explain the variable definitions for the variables used in this paper from the DHS 2007 DRC survey:

- **Years of Education:** For each individual surveyed, the DHS survey asks the individual the total number of years of education in single years.

- **Literacy:** Literacy is a 0 to 1 indicator variable for each individual where 0 is "cannot read at all", and 1 is "able to read only parts of a sentence" or "able to read a whole sentence".

- **Wealth Factor Score:** Wealth Factor Score is an index generated by the DHS using principle component analysis on asset ownership for each individual. For the log of the wealth factor score, we define it as *Log(Wealth Score) = Log(Wealth Factor Score + min(Wealth Factor Score) + 1)*.

- **Wealth Index:** Wealth Index is a is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile (in the entire DRC 2007 sample) from the Wealth Factor Score.

The DHS survey runs a survey instrument on health behavior to a subsample of the sampled female population (about a third of the entire sample). The following variables are only defined for this subsample:

- **Respondent Ht/Age Percentile:** The aforementioned subsample of the female population measures respondent's height (cms) and weight (kgs). Respondent Ht/Age Percentile divides each respondent's height by their age and determines the percentile for this measure relative to the entire subsample. This index is then normalized by the DHS to be within 0 and 10000.

A.3. Colonial Data and Variables

- **Missionary Stations in 1897**: Missionary post locations in 1897 is from a map in Rouck (1945). This map contains missionary post locations for 1897 and was digitized in ArcGIS. Our variable *Number of Missionary Stations in 1897* is defined as the total number of missionary stations in 1897 located in each 20 km by 20 km grid cell.

- **Missionary Stations in 1924**: Missionary post location in 1924 is from Nunn (2010) and is available at http://scholar.harvard.edu/nunn/pages/data-0 in the form of a GIS shapefile. This shapefile was created by Nathan Nunn by digitizing maps from *"Ethnographic Survey of Africa: Showing the Tribes and Languages; also the Stations of Missionary Societies"* published by Roome (1924). Our variable *Number of Missionary Stations in 1924* is defined as the total number of missionary stations in 1924 located in each 20 km by 20 km grid cell.

- **Missionary Stations in 1953**: Missionary post location in 1953 is from a map in the Académie Royale des Sciences d'Outre-Mer (1954). This map contains missionary post locations for 1953 and was digitized in ArcGIS. Our variable *Number of Missionary Stations in 1953* is defined as the total number of missionary stations in 1953 located in each 20 km by 20 km grid cell.

- **Telecommunication Stations in 1953**: Telecommunication Station locations in 1953 are from a map in the Académie Royale des Sciences d'Outre-Mer (1954). This map contains telecommunication post locations for 1953 and was digitized in ArcGIS. Our variable *Number of Telecommunication Stations in 1953* is defined as the total number of Telecommunication stations in 1953 located in each 20 km by 20 km grid cell.

- **Health Centers in 1953**: Health center location in 1953 for the DRC is from the Académie Royale des Sciences d'Outre-Mer (1954). The Académie Royale des Sciences d'Outre-Mer (1954) includes a map with missionary post locations for 1953 that was digitized in ArcGIS. Our variable *Number of Health Centers in 1953* is defined as the total number of health centers in 1953 located in each 20 km by 20 km grid cell.

- **Road Network Density in 1968**: Maps outlining the road network in 1968 for the DRC are available at the UT Map Library (Perry-Castaneda Map Collection), specifically the Africa Map Series made by the Army Map Service, Corps of Engineers, U.S. Army. This series was made in 1968 using the best available sources at the time, and is available at http://www.lib.utexas.edu/maps/ams/africa/africa_index.html. The DRC maps and roads were digitized in ArcGIS. Our *Road Network Density in 1968* variable is defined as total length in meters of roads in 1968 in each 20 km by 20 km grid divided by the grid's total surface area in kilometers squared.

A.4. Precolonial Data and Variables

- **Number Enslaved (Atlantic Trade)**: We use data from Nunn and Wantchekon (2011) on the number of individuals enslaved (in 1000s) from each ethnic group – where ethnic groups are defined using maps from Murdock (1959) – during the Atlantic slave trade.

- **Precolonial Data**: Precolonial data are from the Ethnographic Atlas created by Murdock (1967). Note that not all ethnic groups have data for pre-colonial information (and many groups are missing information for some variables), so the sample size for these variables is small for our area of interest. The variables we use are the following:

- **Population Density**: Population data are estimated by Murdock (1959) for African ethnic groups. Population density is presented as the mean size of the local community (v30) provided in bins equal to: <50, 50-99, 100-199, 200-399, 400-1,000, 1,000-5,000, 5,000-50,000, and more than 50,000. We follow Alsan (2015) as use her measure of population density for each group, defined as logarithm (inhabitants per square kilometer).
- **Centralization**: Centralization is defined as in Alsan (2015): an indicator variable equal to one if there are > 1 level of hierarchy above the local authority (v33) in Murdock (1967).
- **Hereditary Local Headman Selection**: Hereditary Local Headman Selection is defined as an indicator variable equal to one if the succession to the office of local headman (v72) is through either the patrilineal heir or the matrilineal heir. (However, note that most of the groups in our region of interest were not centralized.)
- **Polygynous**: Polygynous is defined as an indicator variable equal to one if the Domestic Organization (v8) is said to be "Polygynous: unusual co-wives pattern" or "Polygynous: usual co-wives pattern" in Murdock (1967).

A.5. Road Network Data and Variables

- **Road Density** The *Referentiel Geographique Commun* also provides a GIS shapefile on the road network in the DRC as of 2010. Our *Road Density* variable is defined as total length in meters of roads in each 20 km by 20 km grid divided by the grid's total surface area in kilometers squared.

- **Bridges** The *Referentiel Geographique Commun* provides a GIS shapefile on the location of all bridges in the DRC as of 2010. Our *Number of Bridges* variable is defined as the total number of bridges located in each 20 km by 20 km grid cell.

Appendix B. Additional Maps and Figures

B.1. Additional Maps

Figure B1: Navigable Rivers and Concessions

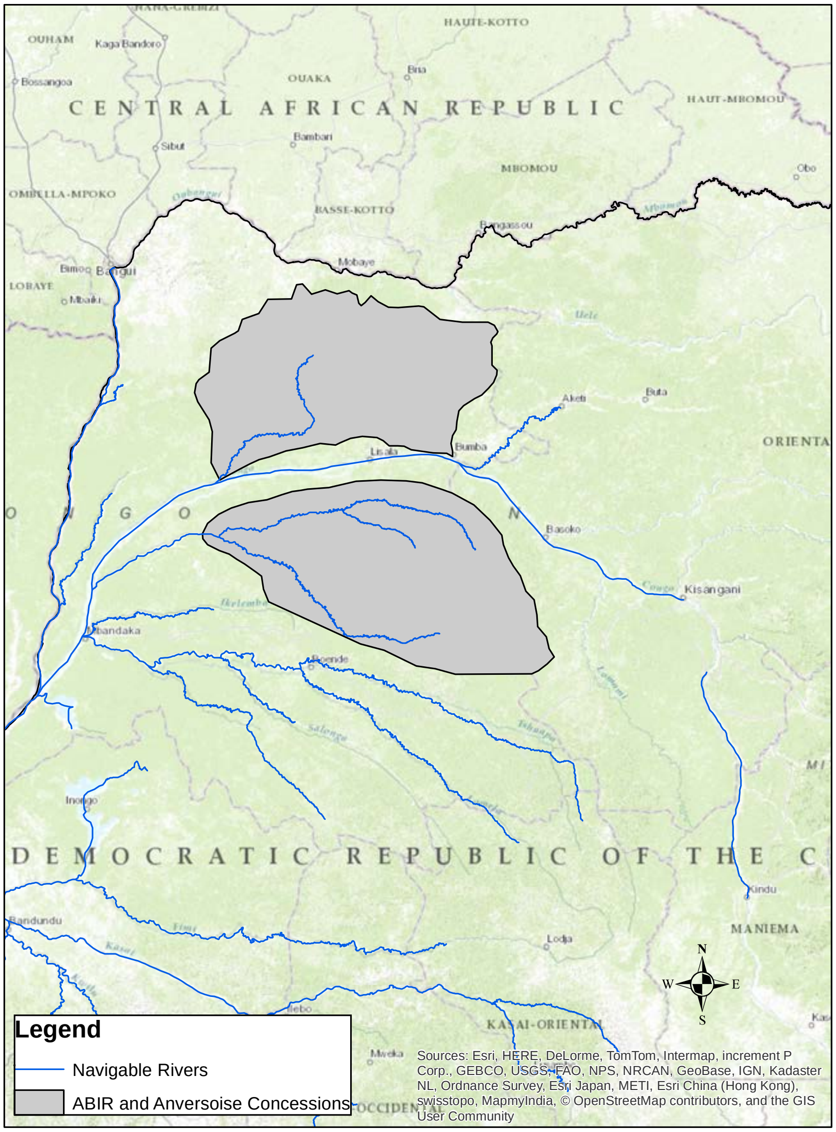


Figure B2: Rivers and Concessions

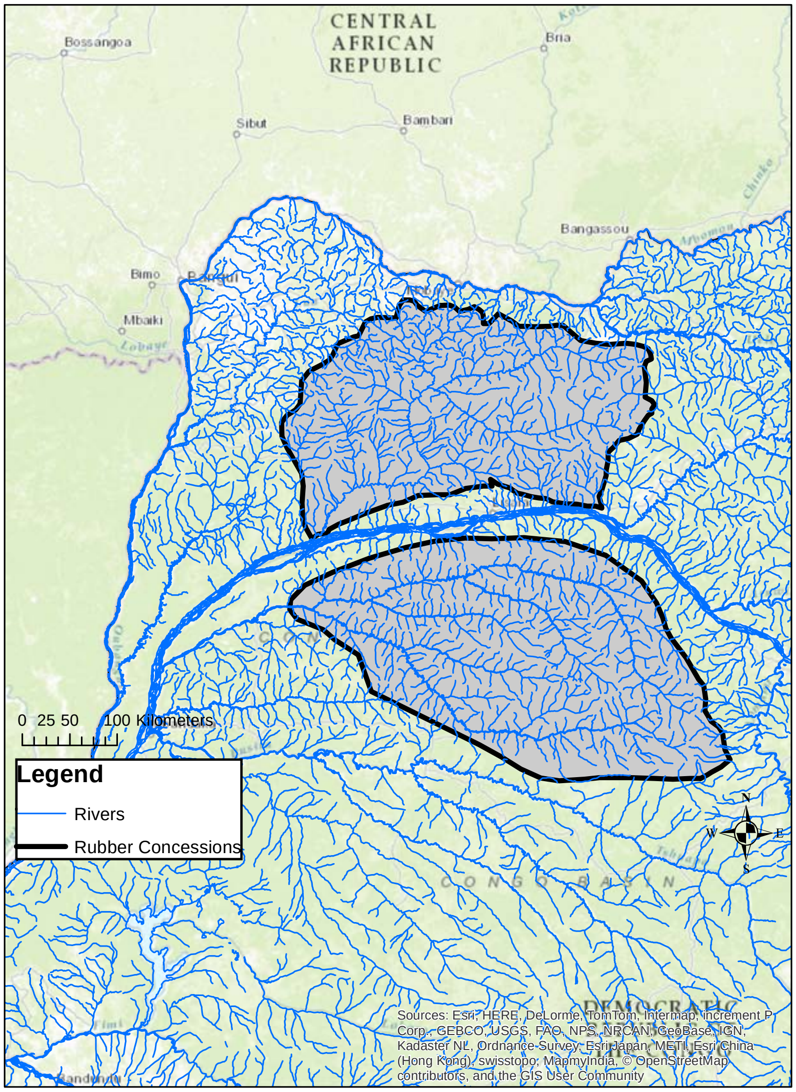


Figure B3: Road Networks and Rubber Concessions

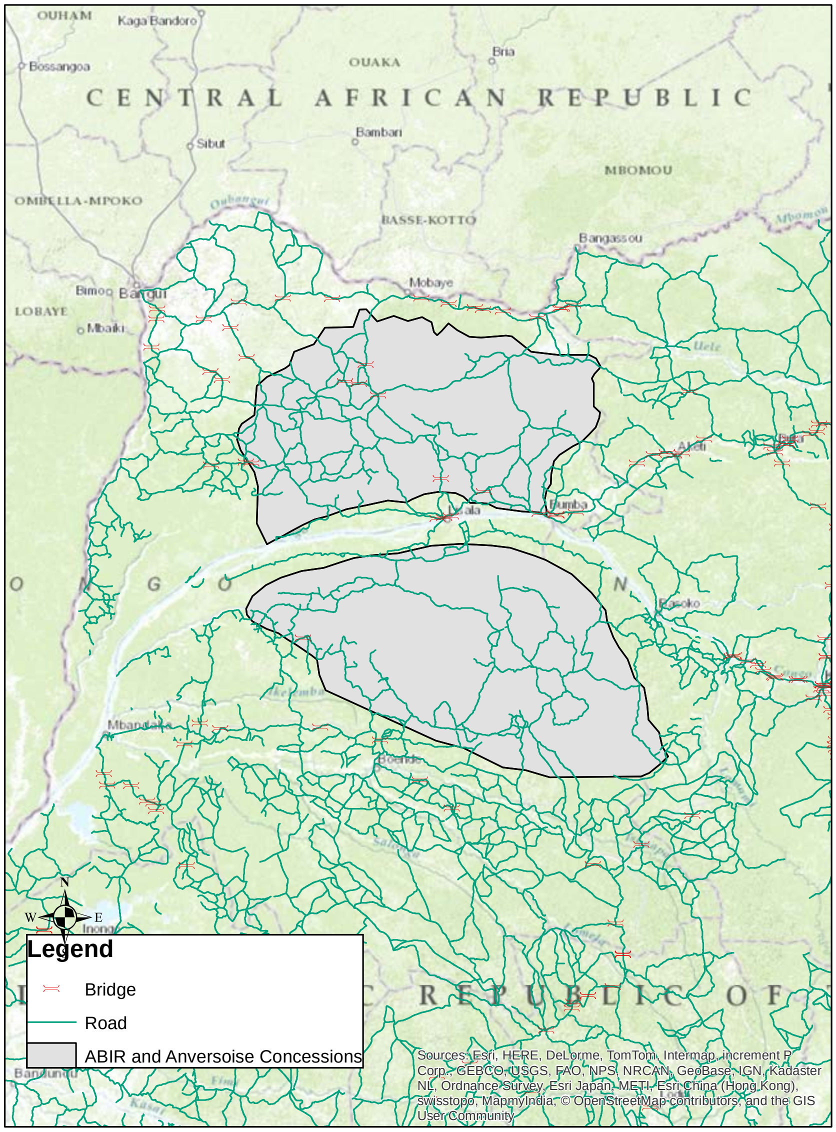


Figure B4: Murdock Ethnic Group Borders and Rubber Concessions

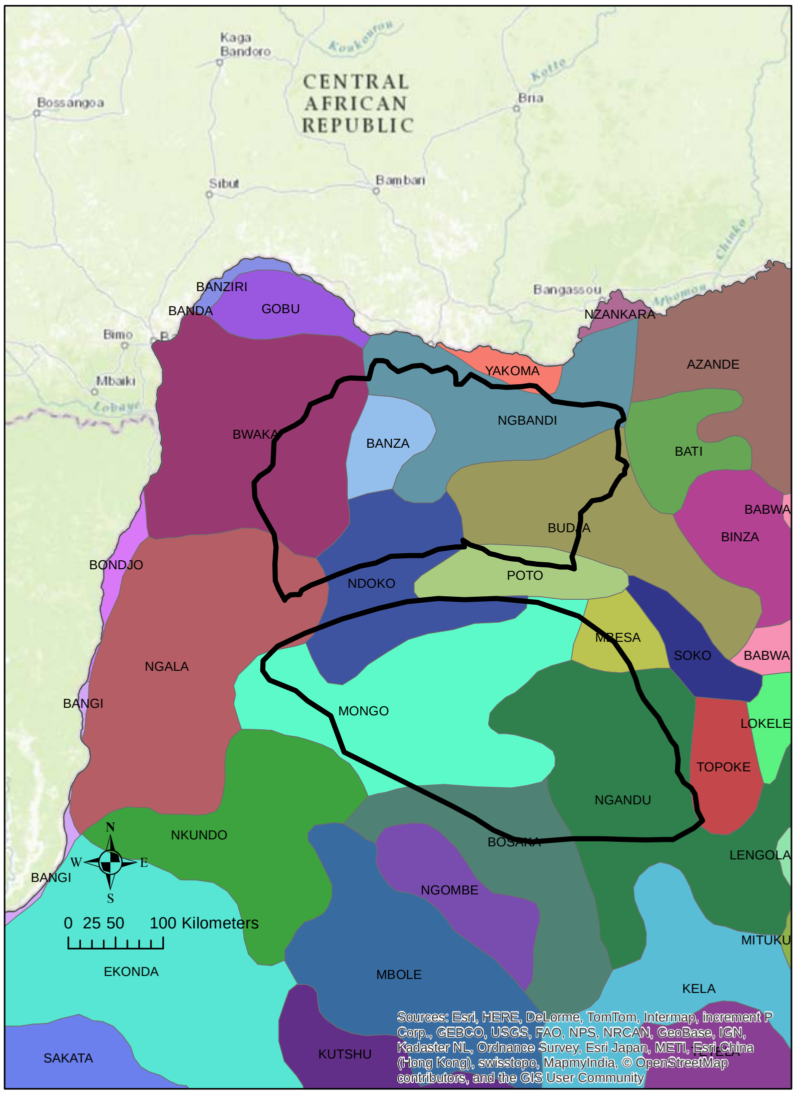


Figure B5: Missions in 1924 and Rubber Concessions

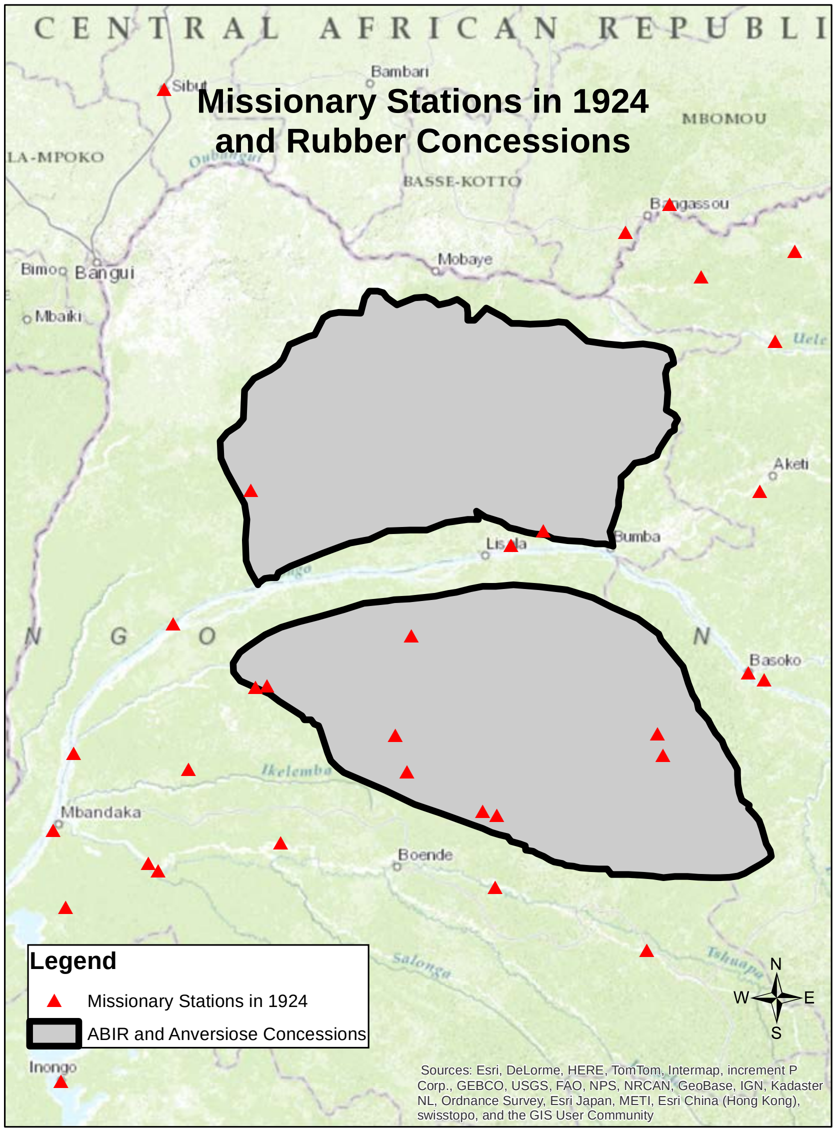


Figure B6: Rivers in 1908 and Rubber Concessions

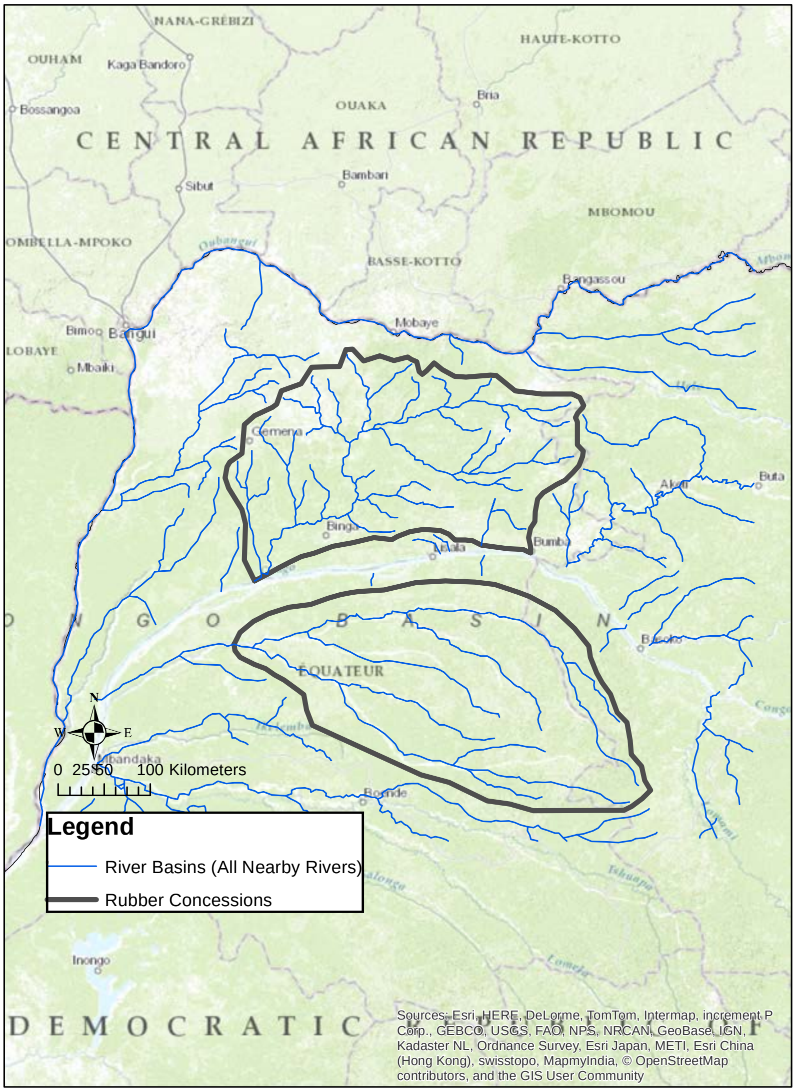
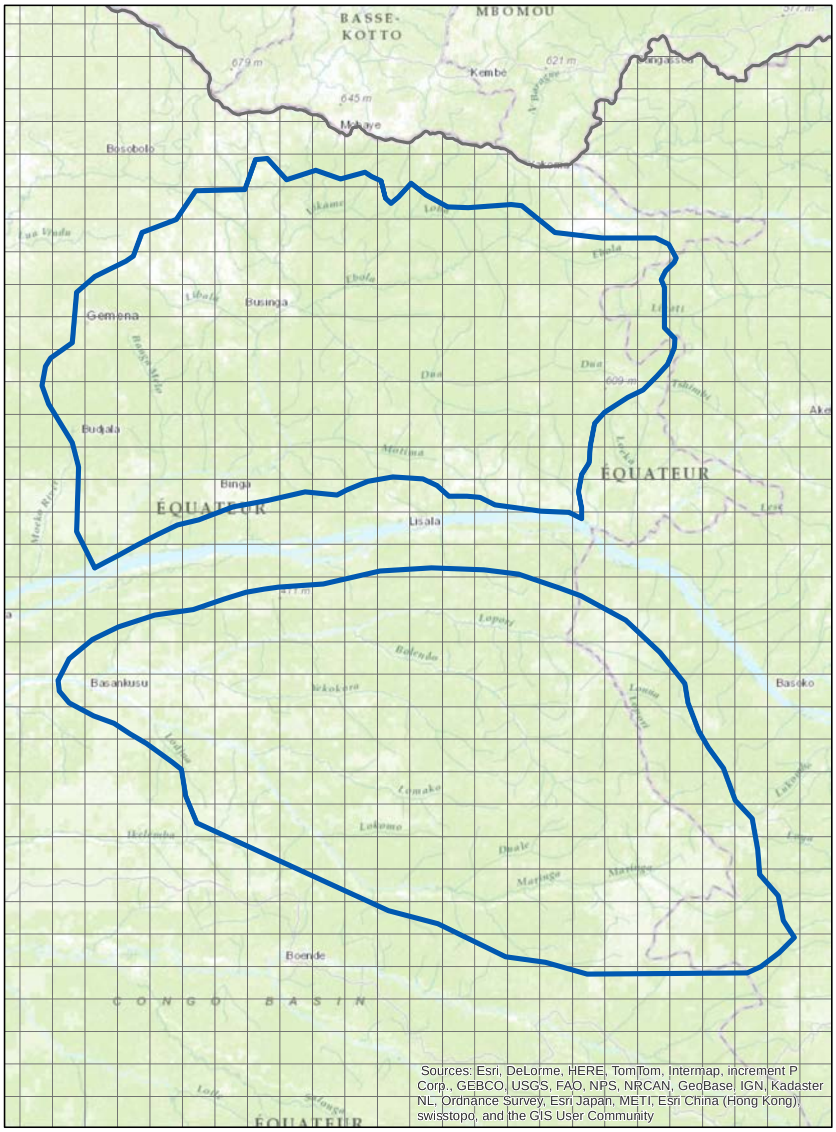


Figure B7: 20 km by 20 km Grid Cell Example



B.2. 3D RD Plots

Note: The figures below are spatial RD plots. Each figure plots the geographic scatterplot of the DHS clusters, each shaded with the mean value in each cluster of the outcome variable of interest. The background shows predicted values for a finely spaced grid of longitude-latitude coordinates from a regression using a linear polynomial in latitude and longitude and the $RubberConcession$ indicator variable. Darker values represent worse development outcomes and vice-versa.[35]

[35] We present results using the linear polynomial in latitude and longitude for simplicity, comparability to other work, and as a complement the RD plots presented using the local linear in distance to the border specifications.

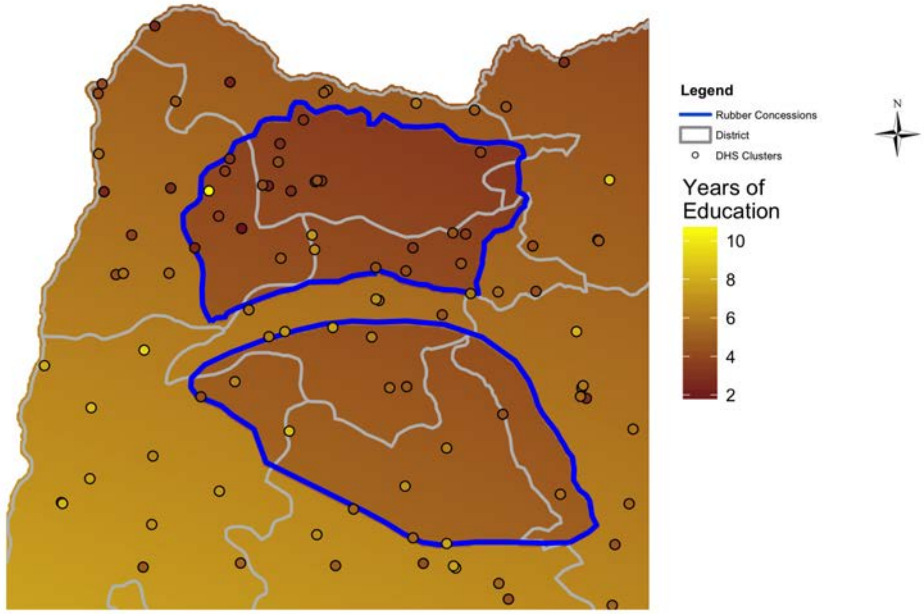


Figure B8: RD Plots - Years of Education

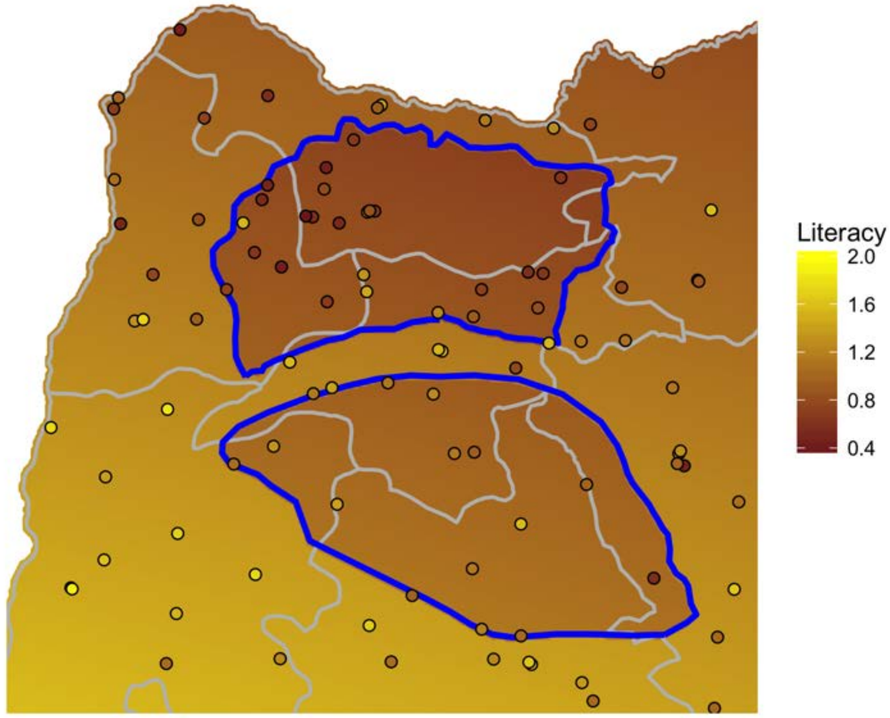


Figure B9: RD Plots - Literacy

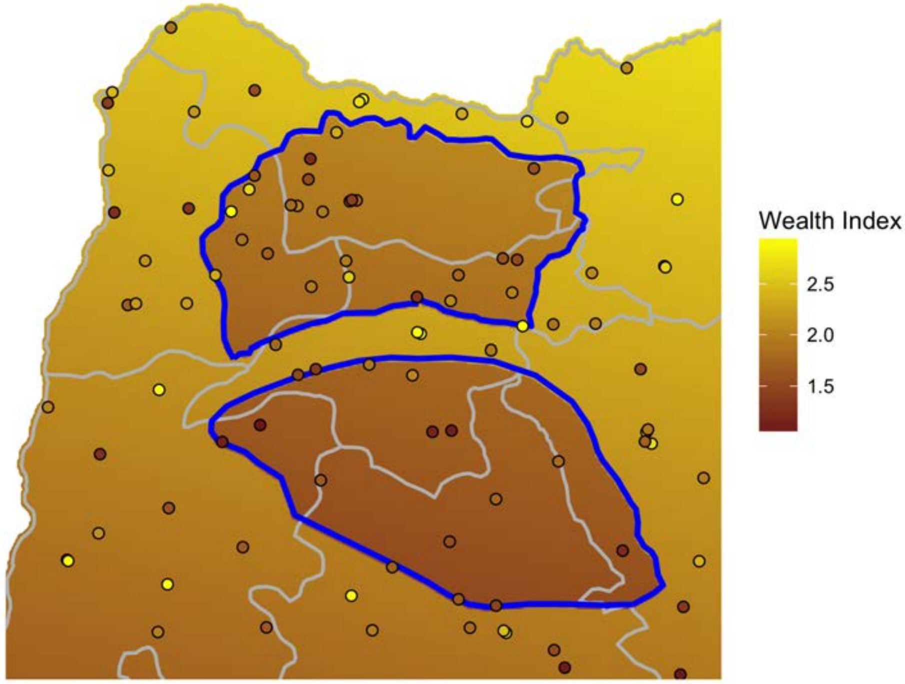


Figure B10: RD Plots - Wealth Index

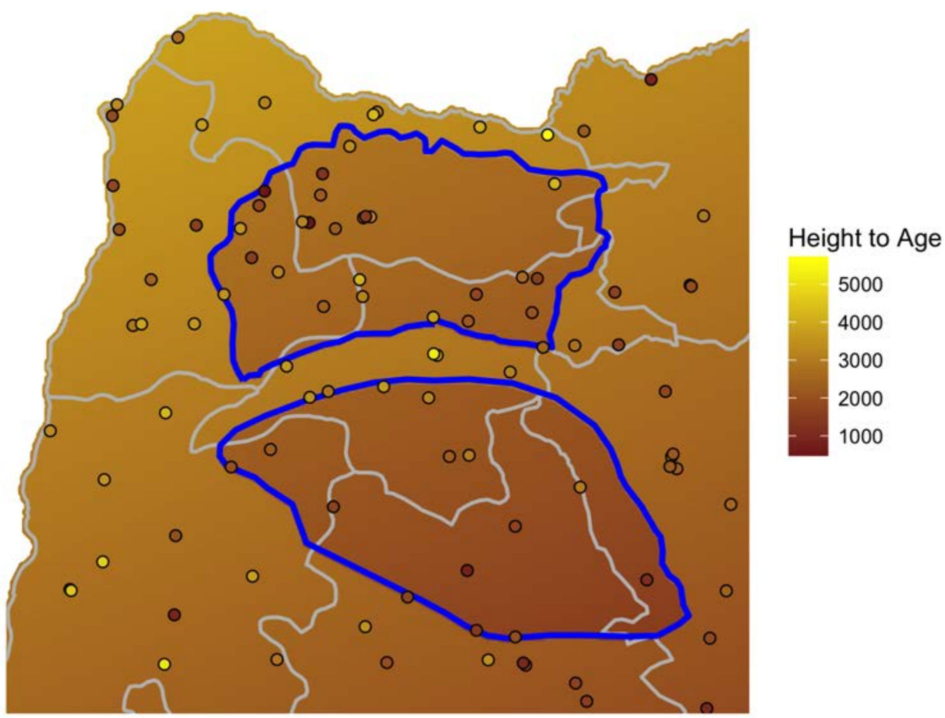


Figure B11: RD Plots - Height-to-Age

Appendix C. Robustness Tables

C.1. DHS Summary Statistics

Table C1 presents simple differences in means inside and outside the concession areas for variables from the DHS. We restrict our analysis for these differences in means to observations that are within 200 kms of the rubber concession borders in order to compare relatively similar areas. Simply comparing differences in means, it appears that the concession areas are less educated, less wealthy, and have worse health outcomes than the areas just outside the concession borders. We have also examined these differences in means between areas inside the former concessions and areas outside the concessions for bandwidths of 100 kms and 50 kms and for all DHS clusters in the DRC. The summary statistics are generally consistent with Table C1.

Table C1: Summary Statistics

	Individuals Within 200 kms of Concession Borders			
	Mean Inside	Mean Outside	Clustered S.E.	(p-value)
Educational Attainment	1.013	1.409	(0.069)	0.0003
Obs	1,843	3,894	–	–
Years of Education	4.228	6.289	(0.368)	0.0006
Obs.	1,837	3,891	–	–
Literacy	0.884	1.283	(0.071)	0.0002
Obs.	1,836	3,870	–	–
Wealth Index	1.824	2.505	(0.156)	0.0009
Obs.	1,843	3,894	–	–
Wealth Score	-54,511	-18,419	(9,494)	0.0008
Obs.	1,843	3,894	–	–
Women Ht/Age Percentile	2,469	2,994	(204.8)	0.012
Obs.	545	1080	–	–
Child Ever Vaccinated	107.0	264.2	0.037	0.033
Obs.	599	1070	–	–
Child Ht/Age Percentile	2,314	2,633	182.1	0.082
Obs.	557	1055	–	–

Notes: The data are from the DHS 2007 and 2014 DRC surveys. Standard errors are clustered at the DHS cluster level. There are 109 clusters within 200 kms of the historical rubber borders. Educational Attainment is a 0 to 3 categorical variable where 0 is no education and 3 is higher education. Literacy is a 0 to 2 categorical variable where 0 is cannot read at all and 2 is able to read a whole sentence. Wealth Factor is an index generated by the DHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Factor Score. Ht/Age Percentile divides each respondent's height by their age and finds their percentile in the sample and normalizes this percentile to be within 0 and 10000. The DHS only records respondent's height and weight for a subsample of the female population. Child Ever Vaccinated is an indicator variable equal to one if the child has ever received a vaccination. Child Ht/Age Percentile divides each children's height by their age and finds their percentile in the sample and normalizes this percentile to be within 0 and 10000. See Data Appendix for more details.

C.2. Balance Table with Conley Standard Errors

Table C2: Balance on Geographic and Pre-Concession Characteristics: Conley Standard Errors

	Within 100 km			*Within 50 km*			*RD Estimates*	
	Inside	Outside	SE	Inside	Outside	SE	RD Coefficient	SE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Geographic Characteristics:								
Elevation	437.26	430.24	(6.84)	440.06	432.46	(8.21)	5.05	(6.85)
Rainfall (Avg.)	72.49	76.42	(2.80)	70.43	74.19	(3.39)	-3.62	(2.79)
Rainfall (St. Dev.)	1.31	1.43	(0.07)	1.32	1.41	(0.09)	-0.11	(0.07)
Land Suitability	5.78	7.62	(1.33)	5.01	7.84	(1.67)	-0.85	(1.33)
Ruggedness	4.94	6.25	(0.85)*	5.31	5.74	(1.10)	-1.48	(0.83)*
Obs.	349	504		232	272		894	
River Characteristics:								
Navigable River Density	10.07	10.51	(2.28)	12.20	9.19	(2.84)	-0.93	(2.33)
>0 Navigable Rivers	0.21	0.22	(0.04)	0.24	0.19	(0.05)	-0.02	(0.04)
>0 Rivers	0.44	0.51	(0.05)	0.38	0.46	(0.06)	-0.07	(0.04)
Obs.	349	504		232	272		894	
Disease Characteristics:								
Malaria Suitability	18.76	18.75	(0.16)	18.71	18.84	(0.18)	0.06	(0.16)
TseTse Fly Suitability	1.33	1.34	(0.01)	1.32	1.33	(0.01)	0	(0.01)
Obs.	349	504		232	272		894	
Location Characteristics:								
Distance: Kinshasa	767.30	792.81	(23.69)	764.17	776.68	(26.69)	-22.71	(23.45)
Distance: Coast	1093.71	1047.54	(12.18)***	1082.01	1064.67	(14.13)	49.66	(12.05)***
Obs.	349	504		232	272		894	
Pre-Concession Characteristics:								
Num. Enslaved (Atlantic Trade, 1000s)	0.65	2.40	(1.20)	0.91	4.24	(2.13)	-1.53	(1.13)
Obs.	236	314		159	170		573	
Num Ethnic Groups	11	23		10	17		24	
Population Density	1.41	1.19	(0.07)**	1.36	1.18	(0.07)	0.15	(0.07)**
Obs.	121	187		74	89		329	
Num Ethnic Groups	3	7		3	6		7	
Centralization	0	0.21	(0.04)***	0	0.14	(0.04)**	-0.15	(0.04)***
Obs.	124	280		103	148		426	
Num Ethnic Groups	5	11		5	10		11	
Polygynous	0.46	0.47	(0.06)	0.49	0.45	(0.08)	-0.05	(0.06)
Obs.	247	322		173	176		593	
Num Ethnic Groups	7	13		7	12		13	
Hereditary Local Headman Selection	1	0.90	(0.04)**	1	0.96	(0.02)	0.11	(0.04)***
Obs.	204	240		142	139		462	
Num Ethnic Groups	5	11		5	10		11	

Notes: The unit of observation is a 20 by 20 km grid cell. Columns 1, 2, 4, and 5 give the mean of the corresponding variable. Columns 3 and 6 give the Conley (1999) spatial standard error assuming a cut-off window of 50 kms. Inside and Outside indicate whether a grid cell is inside or outside the former rubber concession area respectively. Columns 7 and 8 give the estimated RD coefficient and Conley (1999) spatial standard error for the corresponding variable as its outcome using a local linear specification estimated separately on each side of the concession boundary. Regressions include a nearest concession fixed effect. The RD MSE optimal bandwidth is determined using the procedure suggested by Cattaneo et al. (2020). Column 7 uses the average of all optimal bandwidths (39.30 kms). Variable definitions and data sources used in this analysis are described in detail in Appendix A. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.3. Balance Table with Wild Bootstrap P-Values

Table C3: Balance on Geographic and Pre-Concession Characteristics: Wild Bootstrap P-Values

	Within 100 km			*Within 50 km*			*RD Estimates*	
	Inside	Outside	p-value	Inside	Outside	p-value	RD Coefficient	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Geographic Characteristics:								
Elevation	437.26	430.24	0.71	440.06	432.46	0.72	5.05	0.69
Rainfall (Avg.)	72.49	76.42	0.49	70.43	74.19	0.47	-3.62	0.43
Rainfall (St. Dev.)	1.31	1.43	0.04**	1.32	1.41	0.09*	-0.11	0.06*
Land Suitability	5.78	7.62	0.52	5.01	7.84	0.41	-0.85	0.60
Ruggedness	4.94	6.25	0.05*	5.31	5.74	0.43	-1.48	0.05**
Obs.	349	504		232	272		894	
River Characteristics:								
Navigable River Density	10.07	10.51	0.93	12.20	9.19	0.67	-0.93	0.76
>0 Navigable Rivers	0.21	0.22	0.83	0.24	0.19	0.65	-0.02	0.73
>0 Rivers	0.44	0.51	0.42	0.38	0.46	0.44	-0.07	0.45
Obs.	349	504		232	272		894	
Disease Characteristics:								
Malaria Suitability	18.76	18.75	0.92	18.71	18.84	0.60	0.06	0.83
TseTse Fly Suitability	1.33	1.34	0.78	1.32	1.33	0.86	0	0.69
Obs.	349	504		232	272		894	
Location Characteristics:								
Distance: Kinshasa	767.30	792.81	0.63	764.17	776.68	0.77	-22.71	0.63
Distance: Coast	1093.71	1047.54	0.00***	1082.01	1064.67	0.05*	49.66	0.00***
Obs.	349	504		232	272		894	
Pre-Concession Characteristics:								
Num. Enslaved (Atlantic Trade, 1000s)	0.65	2.40	0.33	0.91	4.24	0.41	-1.53	0.33
Obs.	236	314		159	170		573	
Num Ethnic Groups	11	23		10	17		24	
Population Density	1.41	1.19	0.62	1.36	1.18	0.56	0.15	0.62
Obs.	121	187		74	89		329	
Num Ethnic Groups	3	7		3	6		7	
Centralization	0	0.21	0.22	0	0.14	0.28	-0.15	0.24
Obs.	124	280		103	148		426	
Num Ethnic Groups	5	11		5	10		11	
Polygynous	0.46	0.47	0.72	0.49	0.45	0.91	-0.05	0.70
Obs.	247	322		173	176		593	
Num Ethnic Groups	7	13		7	12		13	
Hereditary Local Headman Selection	1	0.90	0.47	1	0.96	0.44	0.11	0.46
Obs.	204	240		142	139		462	
Num Ethnic Groups	5	11		5	10		11	

Notes: The unit of observation is a 20 by 20 km grid cell. Columns 1, 2, 4, and 5 give the mean of the corresponding variable. Columns 3 and 6 give the wild bootstrap p-values clustered at the territory level with 500 repetitions for the differences in means. Inside and Outside indicate whether a grid cell is inside or outside the former rubber concession area respectively. Columns 7 and 8 give the estimated RD coefficient and wild bootstrap p-value, respectively, usingthe corresponding variable as its outcome using a local linear specification estimated separately on each side of the concession boundary. Regressions include a nearest concession fixed effect. The RD MSE optimal bandwidth is determined using the procedure suggested by Cattaneo et al. (2020). Column 7 uses the average of all optimal bandwidths (39.30 kms). Variable definitions and data sources used in this analysis are described in detail in Appendix A. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.4. Balance Table at DHS Cluster Level

Table C4 presents the estimates for differences in geographic characteristics for the DHS clusters in the sample instead of at the grid cell level as in Section 3.2. As in Table 1, there are no significant differences in the main geographic variables of interest for our baseline specification.

Table C4: Balance on Geographic and Pre-Concession Characteristics: DHS Cluster Level

	Within 100 km			*Within 50 km*			*RD Estimates*	
	Inside	Outside	SE	Inside	Outside	SE	RD Coefficient	SE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Geographic Characteristics:								
Elevation	394.03	387.58	(14.45)	396.67	389.12	(19.69)	-7.83	(39.43)
Rainfall (Avg.)	149.66	152.07	(1.51)	151.62	154.19	(1.97)	-1.24	(3.17)
Rainfall (St. Dev.)	1.13	1.49	(0.11)*	1.20	1.47	(0.15)	-0.34	(0.23)
Land Suitability	0.03	0.06	(0.01)	0.04	0.05	(0.02)	0	(0.02)
Ruggedness	6.25	5.79	(1.71)	6.58	5.66	(2.41)	5.22	(3.76)
Obs.	37	48		27	25		85	
River Characteristics:								
Navigable River Density	10.72	15.82	(7.96)	11.69	16.14	(8.41)	-5.21	(11.82)
>0 Navigable Rivers	0.19	0.27	(0.09)	0.22	0.28	(0.13)	-0.02	(0.19)
>0 Rivers	0.51	0.42	(0.11)	0.48	0.52	(0.15)	-0.17	(0.22)
Obs.	37	48		27	25		85	
Disease Characteristics:								
Malaria Suitability	18.54	18.82	(0.19)	18.42	18.87	(0.22)	-0.64	(0.38)*
TseTse Fly Suitability	1.30	1.34	(0.01)	1.31	1.33	(0.02)	0.03	(0.03)
Obs.	37	48		27	25		85	
Location Characteristics:								
Distance: Kinshasa	697.88	810.43	(34.95)*	715.33	807.84	(39.33)	-17.17	(57.92)
Distance: Coast	1086.03	1057.30	(16.20)***	1085.27	1072.70	(21.16)*	9.62	(26.66)
Obs.	37	48		27	25		85	
Pre-Concession Characteristics:								
Num. Enslaved (Atlantic Trade, 1000s)	0.05	6.98	(6.22)	0.06	12.88	(11.08)	-21.89	(18.12)
Obs.	37	48		27	25		85	
Num Ethnic Groups	8	16		8	12		18	
Population Density	1.32	1.03	(0.22)	1.32	0.88	(0.18)	0.33	(0.12)***
Obs.	15	13		10	3		28	
Num Ethnic Groups	2	4		2	2		5	
Centralization	0	0.11	(0.04)	0	0.06	(0.04)	0.06	(0.05)
Obs.	17	28		13	16		45	
Num Ethnic Groups	3	9		3	7		9	
Polygynous	0.57	0.43	(0.19)	0.59	0.41	(0.21)	0.33	(0.23)
Obs.	28	30		22	17		58	
Num Ethnic Groups	5	10		5	8		11	
Hereditary Local Headman Selection	1	0.91	(0.08)	1	0.93	(0.07)	-0.06	(0.04)
Obs.	23	22		17	14		45	
Num Ethnic Groups	4	8		4	6		9	

Notes: The unit of observation is a DHS cluster for the 2007 and 2014 DRC DHS Surveys. Columns 1, 2, 4, and 5 give the mean of the corresponding variable. Columns 3 and 6 present robust standard errors. Inside and Outside indicate whether a grid cell is inside or outside the former rubber concession area respectively. Columns 7 and 8 give the estimated RD coefficient and standard error that uses the corresponding variable as its outcome using a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Regressions include a nearest concession fixed effect. The RD MSE optimal bandwidth is determined using the procedure suggested by Cattaneo et al. (2020). Column 7 uses the average of all optimal bandwidths (99.05 kms). Variable definitions and data sources used in this analysis are described in detail in Appendix A. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

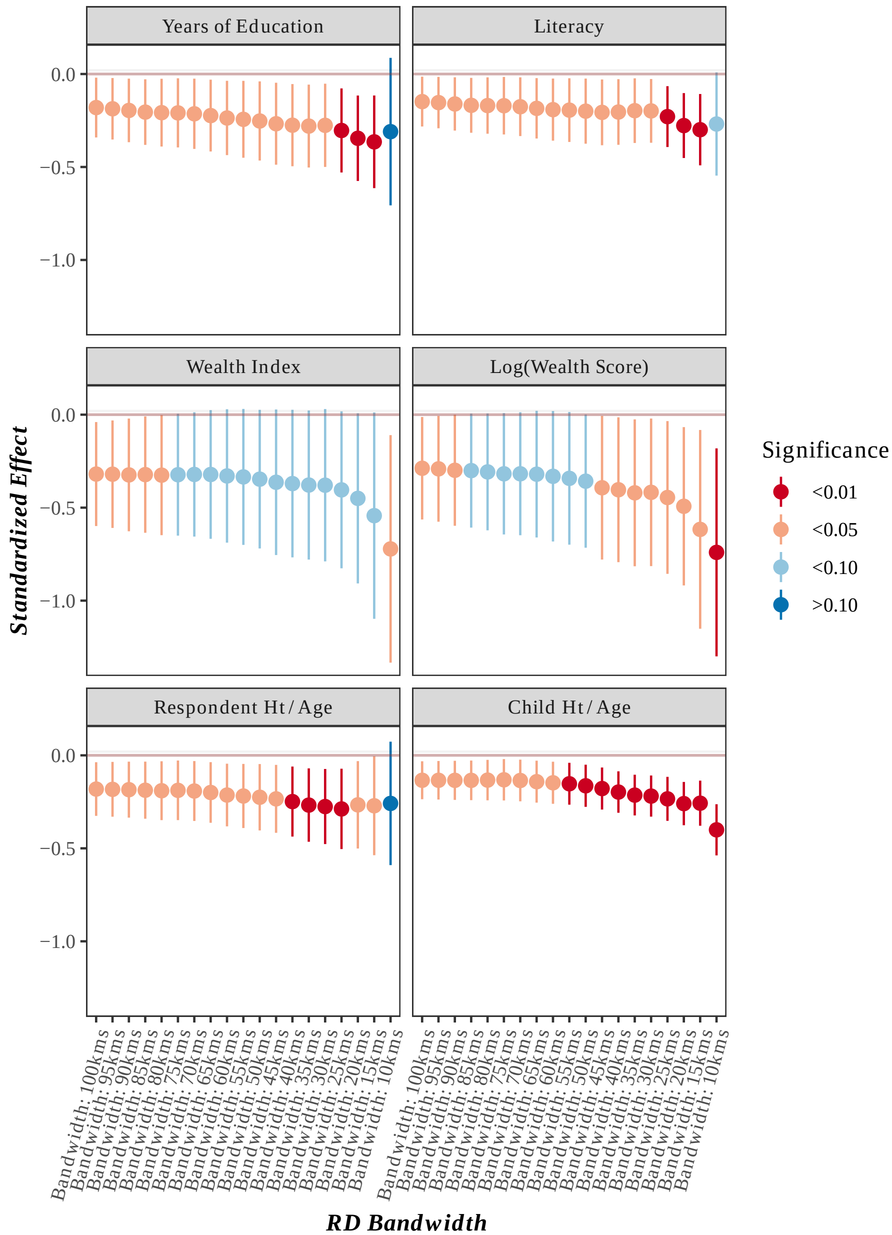
Table C5: Balance on Geographic and Pre-Concession Characteristics: DHS Cluster Level with Conley Standard Errors

	Within 100 km			*Within 50 km*			*RD Estimates*	
	Inside	Outside	SE	Inside	Outside	SE	RD Coefficient	SE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Geographic Characteristics:								
Elevation	394.03	387.58	(14.74)	396.67	389.12	(19.52)	6.39	(14.74)
Rainfall (Avg.)	149.66	152.07	(1.78)	151.62	154.19	(2.15)	2.06	(1.78)
Rainfall (St. Dev.)	1.13	1.49	(0.12)*	1.20	1.47	(0.17)	-0.21	(0.12)*
Land Suitability	0.03	0.06	(0.01)	0.04	0.05	(0.02)	-0.01	(0.01)
Ruggedness	6.25	5.79	(2.05)	6.58	5.66	(2.80)	-0.96	(2.05)
Obs.	37	48		27	25		85	
River Characteristics:								
Navigable River Density	10.72	15.82	(7.76)	11.69	16.14	(8.01)	-3.01	(7.76)
>0 Navigable Rivers	0.19	0.27	(0.10)	0.22	0.28	(0.12)	-0.02	(0.10)
>0 Rivers	0.51	0.42	(0.12)	0.48	0.52	(0.15)	0.13	(0.12)
Obs.	37	48		27	25		85	
Disease Characteristics:								
Malaria Suitability	18.54	18.82	(0.22)	18.42	18.87	(0.25)	-0.07	(0.22)
TseTse Fly Suitability	1.30	1.34	(0.01)	1.31	1.33	(0.02)	0	(0.01)
Obs.	37	48		27	25		85	
Location Characteristics:								
Distance: Kinshasa	697.88	810.43	(39.87)*	715.33	807.84	(43.10)	-65.93	(39.87)*
Distance: Coast	1086.03	1057.30	(18.45)***	1085.27	1072.70	(22.90)*	61.39	(18.45)***
Obs.	37	48		27	25		85	
Pre-Concession Characteristics:								
Num. Enslaved (Atlantic Trade, 1000s)	0.05	6.98	(3.44)**	0.06	12.88	(6.16)**	-6.75	(3.44)**
Obs.	37	48		27	25		85	
Num Ethnic Groups	8	16		8	12		18	
Population Density	1.32	1.03	(0.09)***	1.32	0.88	(0.12)**	0.26	(0.09)***
Obs.	15	13		10	3		28	
Num Ethnic Groups	2	4		2	2		5	
Centralization	0	0.11	(0.03)	0	0.06	(0.04)	-0.04	(0.03)
Obs.	17	28		13	16		45	
Num Ethnic Groups	3	9		3	7		9	
Polygynous	0.57	0.43	(0.12)	0.59	0.41	(0.14)	0.05	(0.12)
Obs.	28	30		22	17		58	
Num Ethnic Groups	5	10		5	8		11	
Hereditary Local Headman Selection	1	0.91	(0.06)	1	0.93	(0.06)	0.08	(0.06)
Obs.	23	22		17	14		45	
Num Ethnic Groups	4	8		4	6		9	

Notes: The unit of observation is a DHS cluster for the 2007 and 2014 DRC DHS Surveys. Columns 1, 2, 4, and 5 give the mean of the corresponding variable. Columns 3 and 6 present Conley (1999) spatial standard error assuming a cut-off window of 50 kms. Inside and Outside indicate whether a grid cell is inside or outside the former rubber concession area respectively. Columns 7 and 8 give the estimated RD coefficient and and Conley (1999) spatial standard error for the corresponding variable as its outcome using a local linear specification estimated separately on each side of the concession boundary. Regressions include a nearest concession fixed effect. The RD MSE optimal bandwidth is determined using the procedure suggested by Cattaneo et al. (2020). Column 7 uses the average of all optimal bandwidths (99.05 kms). Variable definitions and data sources used in this analysis are described in detail in Appendix A. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

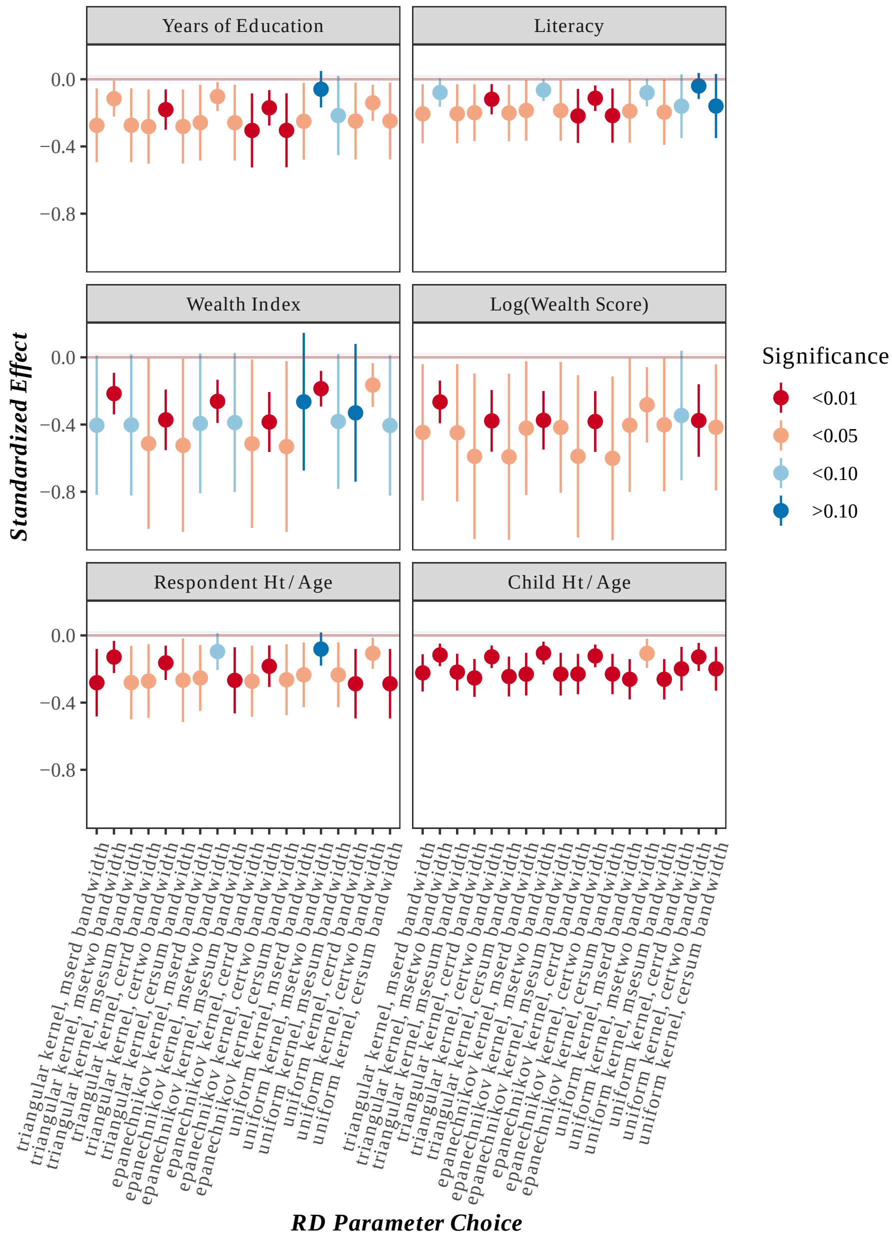
C.5. DHS Results - Varying Local Linear RD Parameters

Figure C1: DHS Results – Robustness to RD Bandwidth



Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary. Regressions control for age, age squared and gender, and nearest concession fixed effects. Regressions use a triangular kernel following Cattaneo et al. (2020). Figure shows point estimates and 95% confidence intervals.

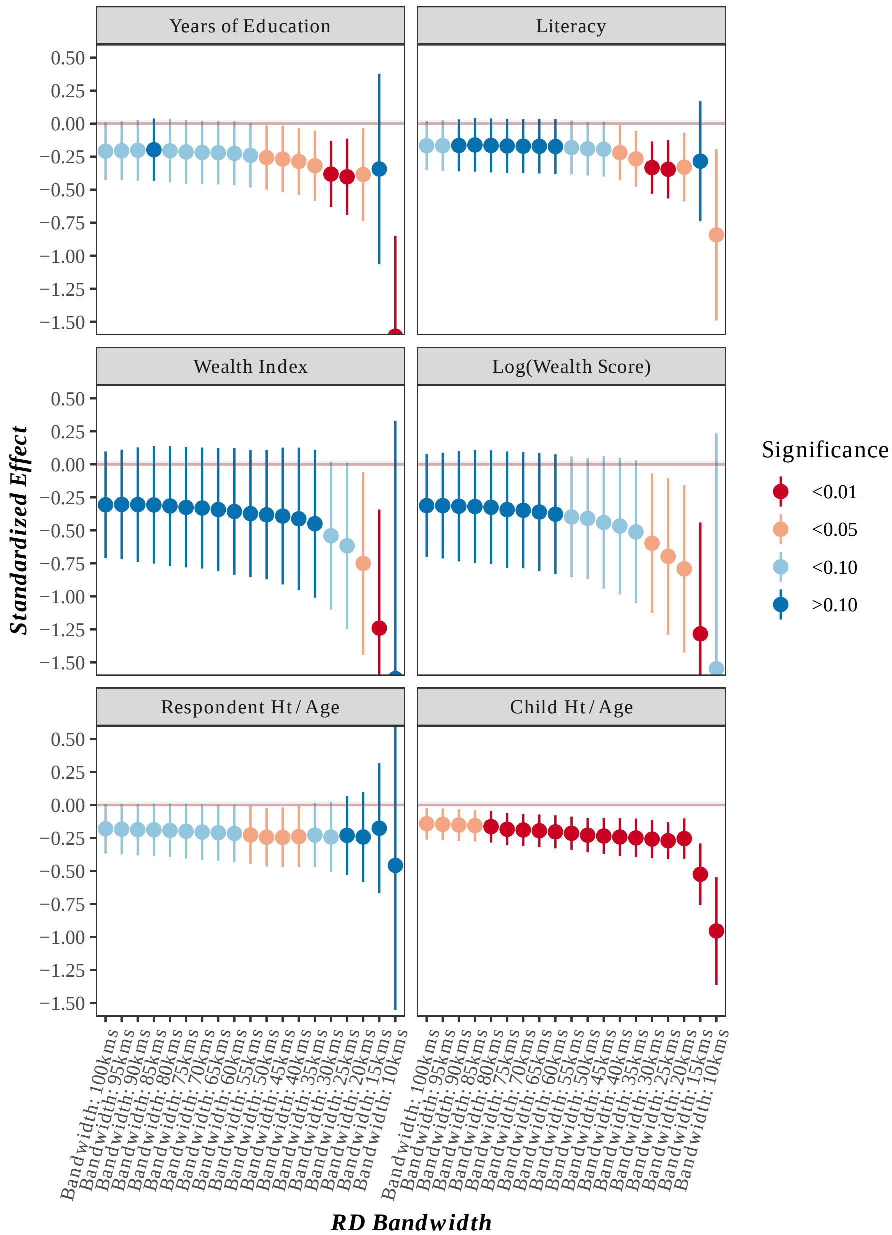
Figure C2: DHS Results – Robustness to Various RD Parameters



Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary. Regressions control for age, age squared and gender, and nearest concession fixed effects. Bandwidth Type represents the optimal bandwidth selection procedure used for each regression: *mserd* choses one common MSE-optimal bandwidth; *msetwo* choses two different MSE-optimal bandwidths (below and above the cutoff); *msesum* choses one common MSE-optimal bandwidth selector for the sum of regression estimates (instead of the difference); *cerrd* choses one common CER-optimal bandwidth; *certwo* two different CER-optimal bandwidths (below and above the cutoff); *cersum* choses one common CER-optimal bandwidth for the sum of regression estimates. See Cattaneo et al. (2020) for more details. Figure shows point estimates and 95% confidence intervals.

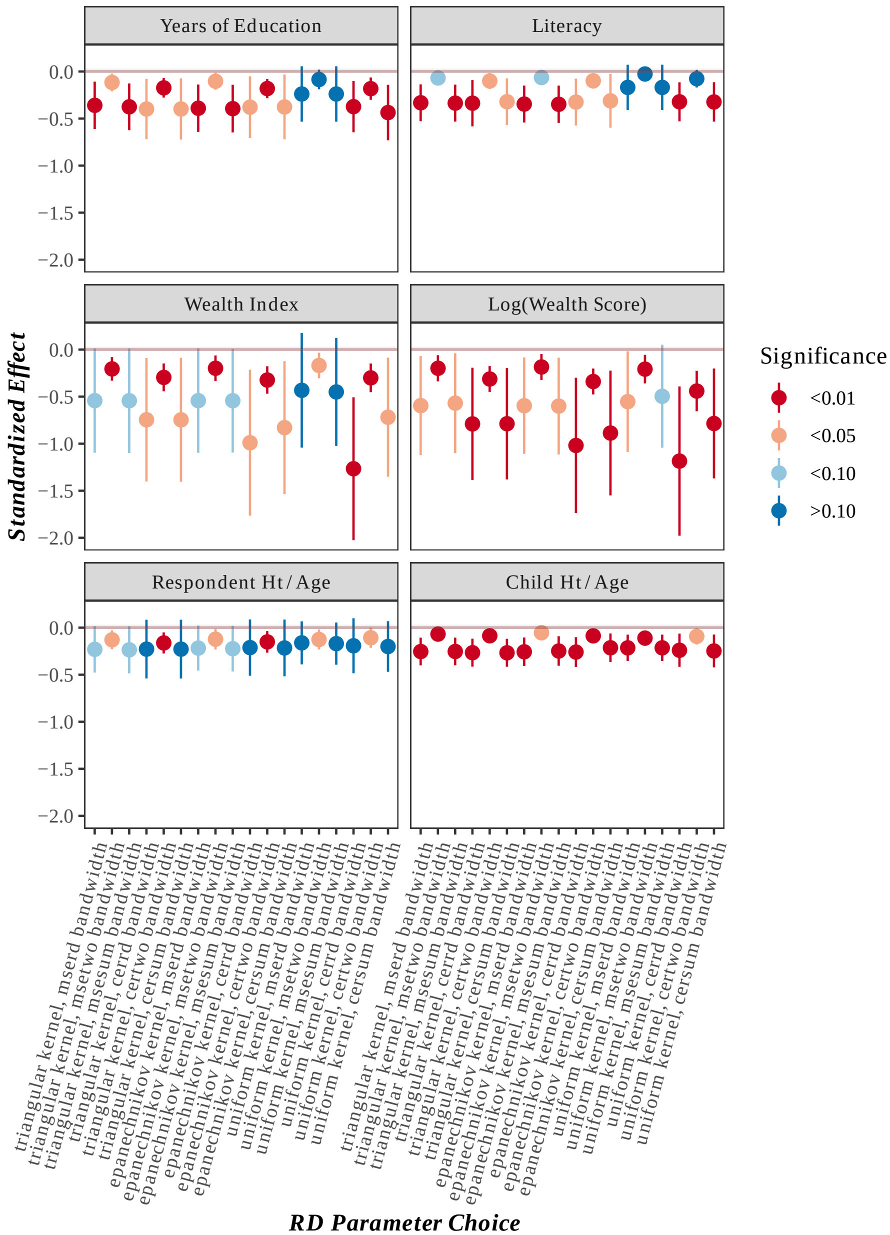
C.6. DHS Results - Robustness to Local Quadratic RD Polynomials

Figure C3: DHS Results – Robustness to RD Bandwidth: Local Quadratic Specification



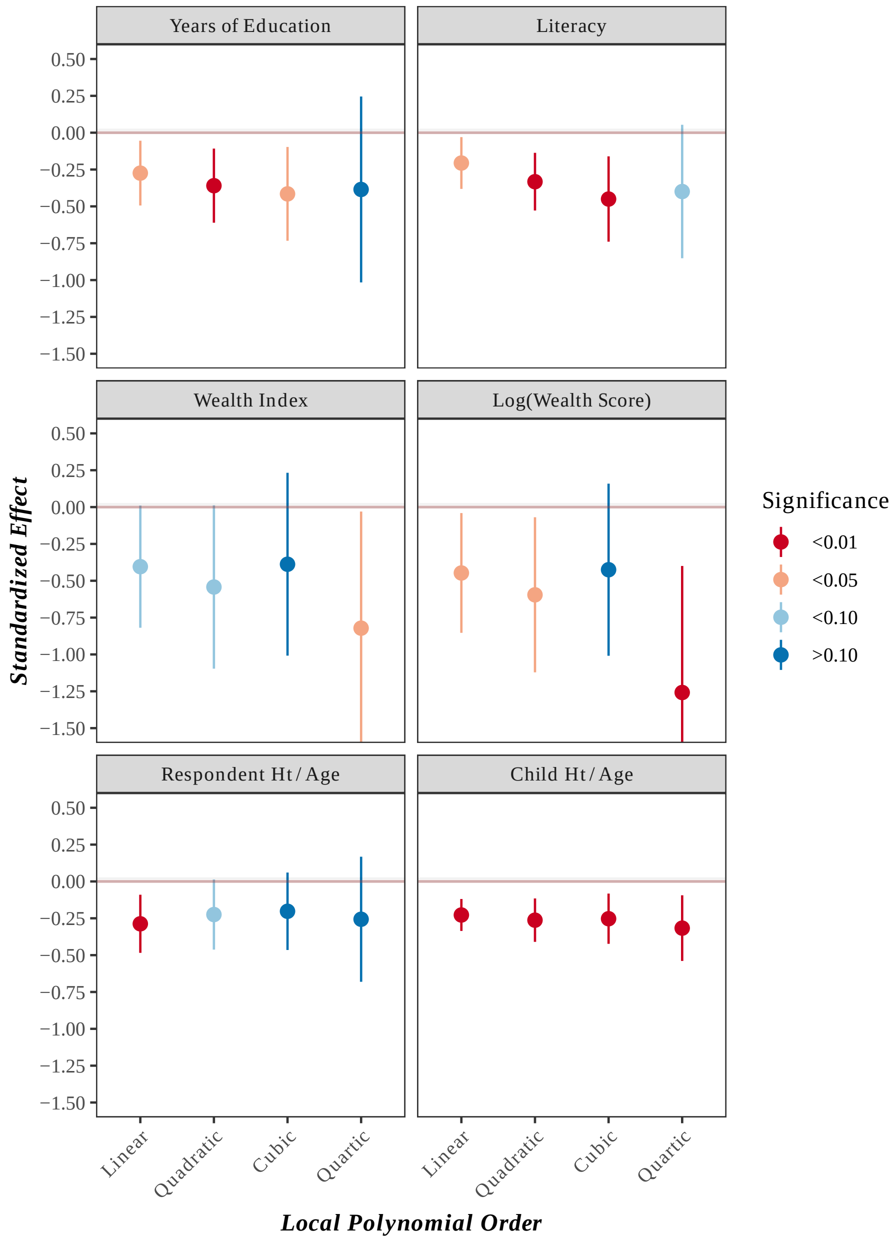
Notes: Standard errors clustered at the DHS cluster level. All regressions include a local quadratic specification estimated separately on each side of the concession boundary. Regressions control for age, age squared and gender, and nearest concession fixed effects. Regressions use a triangular kernel following Cattaneo et al. (2020). Figure shows point estimates and 95% confidence intervals.

Figure C4: DHS Results – Robustness to Various RD Parameters: Local Quadratic Specification



Notes: Standard errors clustered at the DHS cluster level. All regressions include a local quadratic specification estimated separately on each side of the concession boundary. Regressions control for age, age squared and gender, and nearest concession fixed effects. Bandwidth Type represents the optimal bandwidth selection procedure used for each regression: *mserd* choses one common MSE-optimal bandwidth; *msetwo* choses two different MSE-optimal bandwidths (below and above the cutoff); *msesum* choses one common MSE-optimal bandwidth selector for the sum of regression estimates (instead of the difference); *cerrd* choses one common CER-optimal bandwidth; *certwo* two different CER-optimal bandwidths (below and above the cutoff); *cersum* choses one common CER-optimal bandwidth for the sum of regression estimates. See Cattaneo et al. (2020) for more details. Figure shows point estimates and 95% confidence intervals.

Figure C5: DHS Results – Robustness to Various RD Parameters: Higher Order Distance Specifications



Notes: Standard errors clustered at the DHS cluster level. All regressions include a local polynomial specification estimated separately on each side of the concession boundary. Regressions control for age, age squared and gender, and nearest concession fixed effects. egressions use a triangular kernel following Cattaneo et al. (2020). Figure shows point estimates and 95% confidence intervals.

C.7. Alternative RD Specifications: Latitude and Longitude Specifications

Table C6: Rubber Concessions and Education RD Analysis
Alternative RD Specification: Latitude and Longitude Polynomials

	Years of Education			*Literacy*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)
	Panel A: Linear Polynomial in Latitude and Longitude					
Inside Concession	-1.262^{***}	-1.144^{***}	-1.383^{***}	-0.141^{***}	-0.131^{***}	-0.159^{***}
	(0.308)	(0.355)	(0.393)	(0.031)	(0.036)	(0.041)
Observations	5,728	4,274	2,623	5,706	4,266	2,619
Clusters	111	85	52	111	85	52
Bandwidth	5.628	5.109	5.209	0.519	0.465	0.470
Mean Dep. Var.	3.989	3.821	3.787	0.500	0.499	0.499
	Panel B: Quadratic Polynomial in Latitude and Longitude					
Inside Concession	-0.727^{*}	-1.408^{***}	-1.598^{***}	-0.096^{**}	-0.160^{***}	-0.184^{***}
	(0.411)	(0.373)	(0.349)	(0.041)	(0.041)	(0.039)
Observations	5,728	4,274	2,623	5,706	4,266	2,619
Clusters	111	85	52	111	85	52
Bandwidth	5.628	5.109	5.209	0.519	0.465	0.470
Mean Dep. Var.	3.989	3.821	3.787	0.500	0.499	0.499
	Panel C: Cubic Polynomial in Latitude and Longitude					
Inside Concession	-1.087^{***}	-1.584^{***}	-1.687^{***}	-0.125^{***}	-0.180^{***}	-0.196^{***}
	(0.395)	(0.360)	(0.365)	(0.041)	(0.040)	(0.040)
Observations	5,728	4,274	2,623	5,706	4,266	2,619
Clusters	111	85	52	111	85	52
Bandwidth	5.628	5.109	5.209	0.519	0.465	0.470
Mean Dep. Var.	3.989	3.821	3.787	0.500	0.499	0.499

Notes: Standard errors clustered at the DHS cluster level. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C7: Rubber Concessions and Wealth RD Analysis
Alternative RD Specification: Latitude and Longitude Polynomials

	Wealth Index			*Log(Wealth Score)*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)
	Panel A: Linear Polynomial in Latitude and Longitude					
Inside Concession	-0.685^{***}	-0.520^{***}	-0.612^{***}	-0.288^{***}	-0.204^{***}	-0.228^{***}
	(0.123)	(0.127)	(0.185)	(0.061)	(0.056)	(0.078)
Observations	5,737	4,281	2,627	5,737	4,281	2,627
Clusters	111	85	52	111	85	52
Bandwidth	2.287	2.034	2.101	11.041	10.912	10.941
Mean Dep. Var.	1.260	1.060	1.095	0.599	0.443	0.458
	Panel B: Quadratic Polynomial in Latitude and Longitude					
Inside Concession	-0.217	-0.657^{***}	-0.666^{***}	-0.027	-0.259^{***}	-0.251^{***}
	(0.192)	(0.185)	(0.197)	(0.090)	(0.078)	(0.081)
Observations	5,737	4,281	2,627	5,737	4,281	2,627
Clusters	111	85	52	111	85	52
Bandwidth	2.287	2.034	2.101	11.041	10.912	10.941
Mean Dep. Var.	1.260	1.060	1.095	0.599	0.443	0.458
	Panel C: Cubic Polynomial in Latitude and Longitude					
Inside Concession	-0.471^{**}	-0.693^{***}	-0.707^{***}	-0.178^{**}	-0.272^{***}	-0.272^{***}
	(0.182)	(0.195)	(0.196)	(0.081)	(0.081)	(0.081)
Observations	5,737	4,281	2,627	5,737	4,281	2,627
Clusters	111	85	52	111	85	52
Bandwidth	2.287	2.034	2.101	11.041	10.912	10.941
Mean Dep. Var.	1.260	1.060	1.095	0.599	0.443	0.458

Notes: Standard errors clustered at the DHS cluster level. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C8: Rubber Concessions and Health RD Analysis
Alternative RD Specification: Latitude and Longitude Polynomials

	Respondent Ht./Age			*Child Ht./Age*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)
	Panel A: Linear Polynomial in Latitude and Longitude					
Inside Concession	−710.136***	−693.392***	−844.344***	−0.028	−0.030	−0.062**
	(191.431)	(191.889)	(247.385)	(0.017)	(0.018)	(0.024)
Observations	1,870	1,422	888	1,612	1,314	822
Clusters	111	85	52	111	85	52
Bandwidth	2682.521	2592.713	2623.971	0.252	0.247	0.247
Mean Dep. Var.	2612.647	2522.945	2554.350	0.317	0.316	0.321
	Panel B: Quadratic Polynomial in Latitude and Longitude					
Inside Concession	−575.690**	−902.116***	−942.910***	−0.042**	−0.072***	−0.086***
	(224.419)	(211.012)	(235.074)	(0.021)	(0.019)	(0.019)
Observations	1,870	1,422	888	1,612	1,314	822
Clusters	111	85	52	111	85	52
Bandwidth	2682.521	2592.713	2623.971	0.252	0.247	0.247
Mean Dep. Var.	2612.647	2522.945	2554.350	0.317	0.316	0.321
	Panel C: Cubic Polynomial in Latitude and Longitude					
Inside Concession	−842.358***	−1024.087***	−1014.812***	−0.059***	−0.076***	−0.084***
	(209.107)	(202.774)	(233.473)	(0.021)	(0.020)	(0.022)
Observations	1,870	1,422	888	1,612	1,314	822
Clusters	111	85	52	111	85	52
Bandwidth	2682.521	2592.713	2623.971	0.252	0.247	0.247
Mean Dep. Var.	2612.647	2522.945	2554.350	0.317	0.316	0.321

Notes: Standard errors clustered at the DHS cluster level. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.8. DHS Results - "Donut Hole" RD Specifications

Table C9: Rubber Concessions and Economic Development:
Donut Hole Specification

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	-1.443	-1.195^{**}	-0.126	-0.134^{**}
	(1.550)	(0.543)	(0.145)	(0.059)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	2,244	3,745	2,695	3,731
Clusters	23	74	30	74
Bandwidth	22.07	100.00	28.48	100.00
Mean Dep. Var.	4.965	5.037	0.434	0.456
SD Dep. Var.	3.741	3.825	0.496	0.498
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	-1.006	-0.643^{***}	-0.155	-0.216^{**}
	(0.686)	(0.237)	(0.213)	(0.097)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	2,055	3,750	2,017	3,750
Clusters	22	74	20	74
Bandwidth	20.99	100.00	19.49	100.00
Mean Dep. Var.	2.191	2.034	11.011	10.917
SD Dep. Var.	1.086	1.039	0.430	0.431
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.068	-0.086^{**}	-0.040	-0.052
	(0.060)	(0.039)	(0.082)	(0.038)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	605	1,235	489	1,103
Clusters	35	74	22	74
Bandwidth	35.00	100.00	21.02	100.00
Mean Dep. Var.	0.267	0.255	0.265	0.242
SD Dep. Var.	0.258	0.250	0.336	0.313

Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. Regressions do not include clusters within 10 km of the former concession borders before estimating the regressions. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.9. DHS Results - By Concession

C.9.1. ABIR

Table C10: ABIR Concession and Economic Development

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.397	−0.567	−0.049	−0.098
	(0.797)	(0.625)	(0.081)	(0.064)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	919	1,907	1,081	1,904
Clusters	15	39	17	39
Bandwidth	28.51	100.00	33.90	100.00
Mean Dep. Var.	6.106	5.967	0.543	0.555
SD Dep. Var.	3.657	3.776	0.498	0.497
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.115	−0.470*	−0.123	−0.196**
	(0.124)	(0.248)	(0.075)	(0.096)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	773	1,908	773	1,908
Clusters	10	39	9	39
Bandwidth	21.63	100.00	20.83	100.00
Mean Dep. Var.	1.614	1.921	10.759	10.878
SD Dep. Var.	0.813	1.066	0.316	0.449
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.185***	−0.122***	−0.200***	−0.122***
	(0.049)	(0.047)	(0.045)	(0.042)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	270	628	276	594
Clusters	17	39	14	39
Bandwidth	35.00	100.00	27.60	100.00
Mean Dep. Var.	0.262	0.245	0.291	0.263
SD Dep. Var.	0.258	0.243	0.343	0.324

Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for age, age squared, gender, survey year. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.9.2. Anversoise

Table C11: Anversoise Concession and Economic Development

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	−1.484*	−0.662	−0.185**	−0.110*
	(0.767)	(0.595)	(0.078)	(0.059)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	772	1,907	772	1,904
Clusters	10	39	9	39
Bandwidth	21.73	100.00	19.75	100.00
Mean Dep. Var.	5.491	5.967	0.475	0.555
SD Dep. Var.	3.653	3.776	0.500	0.497
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.116	−0.472*	−0.122	−0.197**
	(0.124)	(0.248)	(0.075)	(0.096)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	773	1,908	773	1,908
Clusters	10	39	9	39
Bandwidth	21.63	100.00	20.89	100.00
Mean Dep. Var.	1.614	1.921	10.759	10.878
SD Dep. Var.	0.813	1.066	0.316	0.449
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.185***	−0.122***	−0.200***	−0.122***
	(0.049)	(0.047)	(0.045)	(0.042)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	270	628	276	594
Clusters	17	39	14	39
Bandwidth	35.00	100.00	27.60	100.00
Mean Dep. Var.	0.262	0.245	0.291	0.263
SD Dep. Var.	0.258	0.243	0.343	0.324

Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for age, age squared, gender, survey year. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.10. DHS Results - Without Covariates

Table C12: Rubber Concessions and Economic Development:
No Covariates

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	-1.939^{***}	-1.862^{***}	-0.162^{**}	-0.176^{***}
	(0.746)	(0.583)	(0.078)	(0.060)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	2,183	4,274	2,619	4,266
Clusters	34	85	37	85
Bandwidth	30.45	100.00	36.17	100.00
Mean Dep. Var.	5.215	5.109	0.464	0.465
SD Dep. Var.	3.746	3.821	0.499	0.499
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.442	-0.443	-0.368	-0.291^{**}
	(0.650)	(0.355)	(0.262)	(0.142)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	1,816	4,281	1,816	4,281
Clusters	28	85	27	85
Bandwidth	25.89	100.00	25.07	100.00
Mean Dep. Var.	2.163	2.034	10.957	10.912
SD Dep. Var.	1.122	1.060	0.455	0.443
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.151^{***}	-0.080^{**}	-0.151^{***}	-0.049
	(0.053)	(0.039)	(0.037)	(0.036)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	698	1,422	776	1,314
Clusters	31	85	32	85
Bandwidth	28.15	100.00	29.07	100.00
Mean Dep. Var.	0.282	0.259	0.267	0.247
SD Dep. Var.	0.263	0.252	0.331	0.316

Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.11. DHS Results - Conley Standard Errors

Table C13: Rubber Concession and Economic Development
Conley Standard Errors

	Panel A: Education			
	Years of Education		Literacy	
	(1)	(2)	(3)	(4)
Inside Concession	-1.347^{***}	-1.078^{***}	-0.148^{***}	-0.126^{***}
	(0.446)	(0.360)	(0.045)	(0.038)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	2,183	4,274	2,338	4,266
Bandwidth	40.34	100.00	41.85	100.00
Mean Dep. Var.	5.060	5.109	0.448	0.465
SD Dep. Var.	3.754	3.821	0.497	0.499
	Panel B: Asset Wealth			
	Wealth Index		Log(Wealth Score)	
	(1)	(2)	(3)	(4)
Inside Concession	-0.769^{***}	-0.513^{***}	-0.270^{***}	-0.202^{***}
	(0.245)	(0.137)	(0.098)	(0.060)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	1,303	4,281	1,258	4,281
Bandwidth	25.04	100.00	24.31	100.00
Mean Dep. Var.	2.150	2.034	10.948	10.912
SD Dep. Var.	1.119	1.060	0.453	0.443
	Panel C: Health			
	Respondent Ht./Age Percentile		Child Ht./Age Percentile	
	(1)	(2)	(3)	(4)
Inside Concession	-0.107^{***}	-0.058^{***}	-0.052^{*}	-0.029^{*}
	(0.039)	(0.021)	(0.028)	(0.017)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	531	1,422	548	1,314
Bandwidth	29.20	100.00	28.96	100.00
Mean Dep. Var.	0.278	0.259	0.267	0.247
SD Dep. Var.	0.262	0.252	0.331	0.316

Notes: Standard errors are Conley (1999) spatial standard errors assuming a cut-off window of 50 kms. All regressions include a local linear specification estimated separately on each side of the concession boundary. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for average age, age squared, gender, survey year, and nearest concession fixed effect. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.12. DHS Results - Collapsing on Means at the DHS Cluster Level

Table C14: Rubber Concession and Economic Development
Collapsing at the DHS Cluster Level

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	-1.417^{*}	-1.140^{*}	-0.137^{*}	-0.114^{*}
	(0.733)	(0.587)	(0.070)	(0.059)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	66	85	74	85
Clusters	56.19	100.00	62.20	100.00
Bandwidth	5.070	5.071	0.438	0.460
Mean Dep. Var.	1.732	1.859	0.187	0.193
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.636	-0.621^{**}	-0.267	-0.245^{**}
	(0.408)	(0.285)	(0.170)	(0.117)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	52	85	52	85
Clusters	39.98	100.00	41.55	100.00
Bandwidth	2.051	2.004	10.911	10.900
Mean Dep. Var.	0.701	0.641	0.281	0.271
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.118^{***}	-0.101^{***}	-0.109^{***}	-0.127^{***}
	(0.042)	(0.037)	(0.038)	(0.034)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	64	85	69	85
Clusters	55.58	100.00	63.58	100.00
Bandwidth	0.267	0.260	0.267	0.263
Mean Dep. Var.	0.103	0.098	0.135	0.122

Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for average age, age squared, gender, survey year, and nearest concession fixed effect. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.13. DHS Results - Excluding Observations Near the Congo River

Table C15: Rubber Concession and Economic Development
Excluding Observations within 15 kms of the Congo River

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.293	−0.716	−0.000	−0.078
	(0.765)	(0.662)	(0.040)	(0.064)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	2,234	3,637	2,229	3,629
Clusters	35	72	28	72
Bandwidth	37.85	100.00	31.67	100.00
Mean Dep. Var.	4.809	4.874	0.430	0.444
SD Dep. Var.	3.729	3.815	0.495	0.497
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.249	−0.515*	−0.174**	−0.185
	(0.173)	(0.296)	(0.069)	(0.120)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	1,478	3,643	1,430	3,643
Clusters	17	72	16	72
Bandwidth	21.63	100.00	20.98	100.00
Mean Dep. Var.	2.146	1.959	10.945	10.885
SD Dep. Var.	1.119	1.029	0.463	0.431
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.209***	−0.081*	−0.195***	−0.054
	(0.049)	(0.044)	(0.066)	(0.058)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	508	1,238	732	1,210
Clusters	30	72	26	72
Bandwidth	35.00	100.00	29.87	100.00
Mean Dep. Var.	0.253	0.260	0.314	0.303
SD Dep. Var.	0.260	0.264	0.376	0.367

Notes: Sample excludes observations within 15 kms of the Congo River. Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for average age, age squared, gender, survey year, and nearest concession fixed effect. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C16: Rubber Concession and Economic Development
Excluding Observations within 25 kms of the Congo River

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	0.542	−0.724	0.016	−0.091
	(0.758)	(0.823)	(0.050)	(0.076)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	2,121	3,524	2,164	3,516
Clusters	28	70	27	70
Bandwidth	34.68	100.00	33.68	100.00
Mean Dep. Var.	4.679	4.844	0.425	0.442
SD Dep. Var.	3.741	3.838	0.495	0.497
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.329*	−0.728**	−0.207***	−0.251*
	(0.169)	(0.332)	(0.064)	(0.138)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	1,409	3,530	1,365	3,530
Clusters	15	70	15	70
Bandwidth	20.49	100.00	20.01	100.00
Mean Dep. Var.	2.217	1.967	10.953	10.884
SD Dep. Var.	1.134	1.036	0.475	0.436
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.162***	−0.066	−0.073***	−0.011
	(0.046)	(0.046)	(0.026)	(0.054)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	468	1,198	694	1,172
Clusters	28	70	23	70
Bandwidth	35.00	100.00	28.81	100.00
Mean Dep. Var.	0.248	0.259	0.296	0.301
SD Dep. Var.	0.262	0.265	0.364	0.366

Notes: Sample excludes observations within 25 kms of the Congo River. Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for average age, age squared, gender, survey year, and nearest concession fixed effect. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

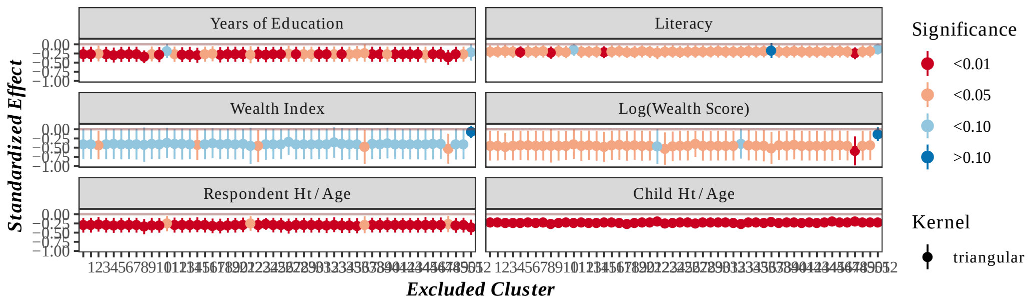
Table C17: Rubber Concession and Economic Development
Excluding Observations within 50 kms of the Congo River

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	0.237	−0.844	−0.036	−0.117
	(0.819)	(0.873)	(0.064)	(0.082)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	1,909	3,195	1,994	3,188
Clusters	24	64	24	64
Bandwidth	35.17	100.00	35.56	100.00
Mean Dep. Var.	4.659	4.859	0.419	0.441
SD Dep. Var.	3.830	3.891	0.494	0.497
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.369***	−0.714**	−0.210***	−0.242*
	(0.138)	(0.326)	(0.062)	(0.139)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	1,524	3,201	1,197	3,201
Clusters	13	64	12	64
Bandwidth	21.50	100.00	20.63	100.00
Mean Dep. Var.	2.277	1.981	10.988	10.886
SD Dep. Var.	1.169	1.047	0.494	0.445
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.175***	−0.069	−0.089***	−0.050
	(0.043)	(0.048)	(0.029)	(0.057)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	394	1,089	585	1,044
Clusters	24	64	20	64
Bandwidth	35.00	100.00	29.51	100.00
Mean Dep. Var.	0.243	0.260	0.301	0.297
SD Dep. Var.	0.266	0.268	0.376	0.365

Notes: Sample excludes the two border-segments closest to the Congo River by excluding DHS clusters within 50 kms of the Congo River. Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for average age, age squared, gender, survey year, and nearest concession fixed effect. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

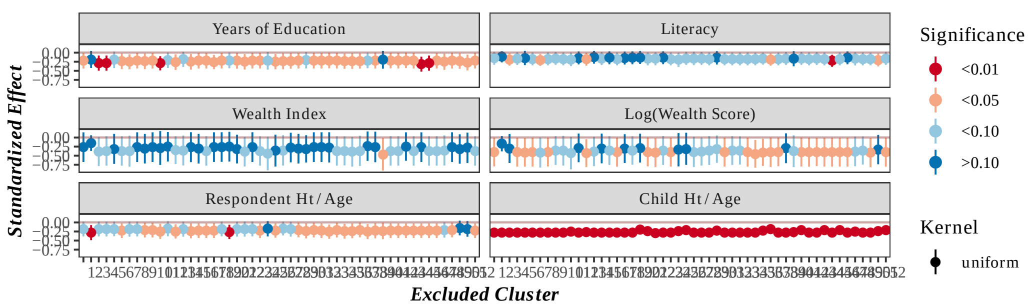
C.14. DHS Results - Robustness to Dropping Clusters

Figure C6: DHS Results – Robustness to Dropping One Cluster At a Time: Local Linear Specification with a Triangular Kernel



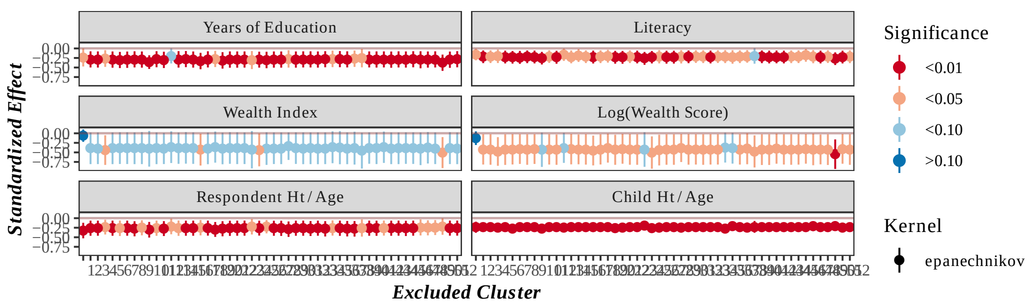
Notes: Standard errors clustered at the DHS cluster level. All regressions include a local quadratic specification estimated separately on each side of the concession boundary estimated within the MSE-minimizing bandwidth from Cattaneo et al. (2020). Regressions control for age, age squared and gender, and nearest concession fixed effects. Figure shows point estimates and 95% confidence intervals.

Figure C7: DHS Results – Robustness to Dropping One Cluster At a Time: Local Linear Specification with a Uniform Kernel



Notes: Standard errors clustered at the DHS cluster level. All regressions include a local quadratic specification estimated separately on each side of the concession boundary estimated within the MSE-minimizing bandwidth from Cattaneo et al. (2020). Regressions control for age, age squared and gender, and nearest concession fixed effects. Figure shows point estimates and 95% confidence intervals.

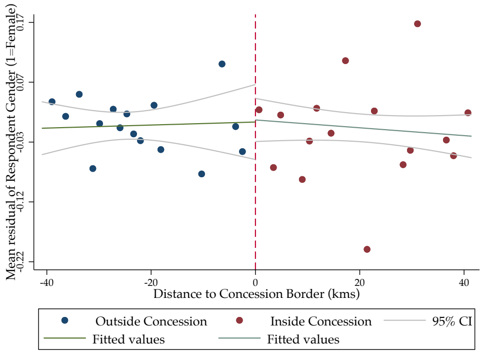
Figure C8: DHS Results – Robustness to Dropping One Cluster At a Time: Local Linear Specification with a Epanechnikov Kernel



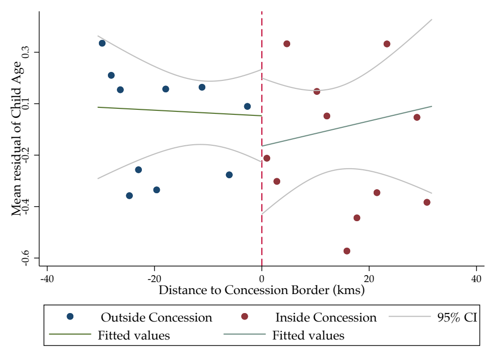
Notes: Standard errors clustered at the DHS cluster level. All regressions include a local quadratic specification estimated separately on each side of the concession boundary estimated within the MSE-minimizing bandwidth from Cattaneo et al. (2020). Regressions control for age, age squared and gender, and nearest concession fixed effects. Figure shows point estimates and 95% confidence intervals.

C.15. *DHS Results - RD Plots for Covariates*

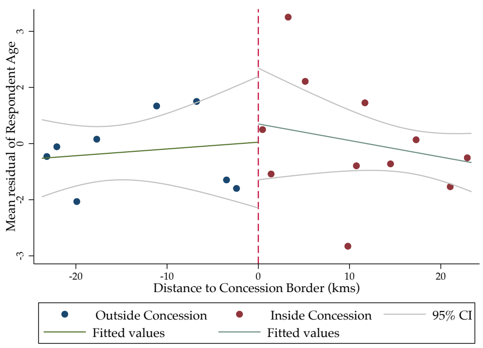
Figure C9: RD Plots for Covariates Outcomes



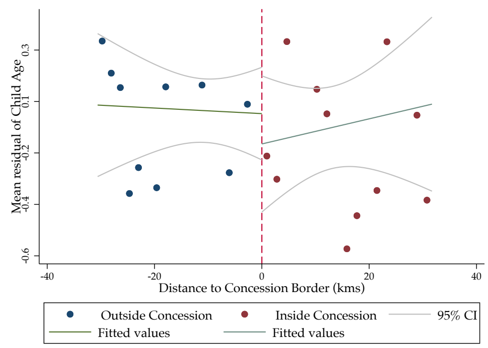
(a) *Respondent Gender*



(b) *Child Gender*



(c) *Respondent Age*



(d) *Child Age*

Notes: The figure shows RD plots for our main covariates and present the mean value of each outcome variable at each 2.5 km bin along the running variable (distance to concession border) as well as with a local linear trend estimated separately on each side of the discontinuity. Each regression is estimated using the optimal bandwidth chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020). Regressions control for survey year, and include a nearest concession fixed effect. Standard errors are clustered at the DHS cluster level and the figure show 95% confidence intervals.

Appendix D. Basin Falsification Exercise

This section explains the basin falsification and basin IV exercises discussed in Section 3.4We first describe the HydroBASINS data and the algorithm used in the construction of the river basin shapefiles. Second, we explain how the falsification exercise was implemented, and present the falsification results. Third, we use the HydroBASINS basins corresponding to the concession basins and present results using these basins and a 25 buffer around them as an IV for being inside the former concessions (and present placebo basin buffers and show the results are strongest for the 25 km buffer, consistent with the way the concessions were defined).

D.1. River Basin Data

The data used in the falsification exercise is from Lehner and Grill (2013).[36] The data used is called the "HydroBASINS" data. This data provides a set of polygon shapefile layers that depict

[36] The data is available online at WWW.HYDROSHEDS.ORG.

watershed boundaries and delineate sub-basins at a global scale.

The significant innovation in this data is the sub-basin delineation procedure. Two important features required for a consistent mapping of river basins at a global scale are (1) a consistent method for sub-basin breakdown, i.e. the decision of when and how to subdivide a larger basin into multiple tributary basins, and (2) a method for grouping sub-basins together. For example, take the Mongala river basin in this paper, which defined the limits of the Anversoise river concession. The Mongala River basin is part of the larger Congo River basin, but also constitutes its own sub-basin, the Mongala River Basin.

The HydroBASINs data proceeds as follows for sub-basin breakdown. First, it breaks out sub-watersheds at any confluence where the inflowing branches (i.e., a tributary and its main stem) exceed a certain size threshold. In particular, hydroSHEDs divides a basin into two sub-basins at every location where two river branches meet which each have an individual upstream area of at least 100 km^2(Lehner and Grill, 2013).

The second critical feature of the HydroSHEDs data is the way the sub-basins are grouped or coded to allow for the breakout of nested sub-basins at different scales. The "Pfafstetter" coding system is used due to its relative simplicity and ease of application. Pfafstetter coding in this case means that a larger basin is sequentially subdivided into 9 smaller units (the 4 largest tributaries, coded with even numbers, and the 5 inter-basins, coded with odd numbers). Thus, the next finer resolution of a sub-basin delineation is achieved at the next Pfafstetter level by adding one digit to the code of the previous level as depicted in the Figure D1 from Lehner and Grill (2013). The HydroBASINS data uses the Pfafstetter coding system for 12 levels globally.

D.2. Falsification Exercise: Implementation

One possible concern with the results presented in Section 3.3 is that because the concession borders were drawn using major river basins, the results may reflect some inherent characteristic of river basins, rather than exposure to the concessions. To assess this claim, we conduct a falsification exercise where we run our main specification across all major river basins in DRC using the HydroBASINS data from Lehner and Grill (2013) to examine how our estimated effects for the former concessions correspond to the estimated effects for all other major river basins in DRC.

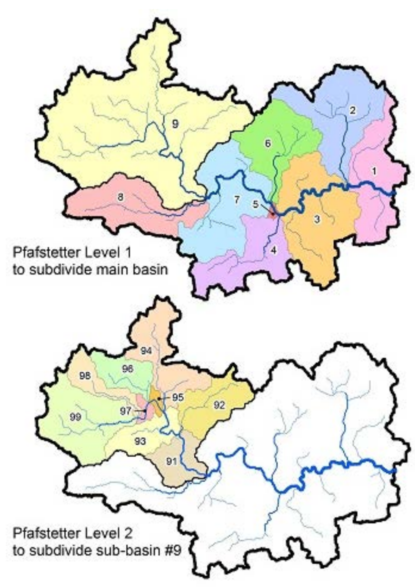


Figure D1: Example of the Pfafstetter Coding used in HyrdroBASINS

In particular, we re-run our main analysis using all other river basin limits from HydroBASINS that are of similar size and importance as our two main river basins. Specifically, the river basins corresponding to Anversoise and ABIR concession boundaries are level 5 and 6 river basins in the HydroBASINS data, respectively. Thus, the falsification exercise uses all level 5 and level 6 HydroBASINS layers to only consider river basins of similar size and importance to the main basins of interest.

We take all of these river basins – excluding the basins corresponding to Anversoise and ABIR concession boundaries – and use the DRC DHS 2007 and 2014 to calculate distance to each basin and whether or not a DHS observation falls within the river basin polygon. We exclude all DHS observations in Kinshasa and Lubumbashi since Kinshasa and Lubumbashi – the two largest cities in DRC – are major outliers in the DHS data. We then estimate our two specifications. (1) Our baseline specification – local linear in distance to the borders estimated separately on each side – within a bandwidth of 25 km (the average optimal bandwidth across all basins), our preferred specification as detailed in Section 3.2. (2) to complement the specification using distance to the border, we also show our results using a parsimonious linear latitude-longitude RD specification – linear latitude-longitude – within a bandwidth of 100 km from the river basin border that includes controls for age, age-squared, gender, survey-year fixed effects and district fixed effects. The district fixed effects serve as border-segment fixed effects. The advantage of using districts fixed effects as border-segment fixed effects is that we can construct border-segment fixed effects for all of DRC in a non-arbitrary manner.

As the dependent variables for the analysis, we use years of education and the log of the wealth score. We limit our analysis to river basins that have at least five DHS clusters within the basin so that the RD estimate is well-estimated, leaving us with 29 river basins in total. We record the RD estimates for each of these river basins and present them visually in the next section below.

Figure D4a presents a map of all level 5 river basins in DRC and Figure D4b presents a map of all level 6 river basins in DRC from the HydroBASINS data. Figure D5 presents a map of the HydroBASINS river basins corresponding to the ABIR and Anversoise concessions along with the actual concession borders. It shows that both boundaries are roughly similar, consistent with Section 2 on how these concession boundaries were drawn.[37]

D.3. Falsification Exercise: Results

Figure D2 presents the empirical cumulative distribution of the RD estimates for education for all major river basins in DRC, excluding the basins corresponding to the Anversoise and ABIR concession boundaries. On average, being inside a river basin is associated with more years of education and higher asset wealth. To highlight where the corresponding RD estimates for ABIR and Anversoise would fall relative to these estimated basin effects, we include in solid-red the RD estimate corresponding to the ABIR concession border and in dashed-blue the RD estimate corresponding to the Anversoise concession border. The Anversoise estimate falls in the bottom 2.5% of this river basin RD estimate distribution for both outcomes of interest, while ABIR falls in the 0.05% and 0.10% of the distributions for log wealth score and years of education, respectively, when using our baseline specification.

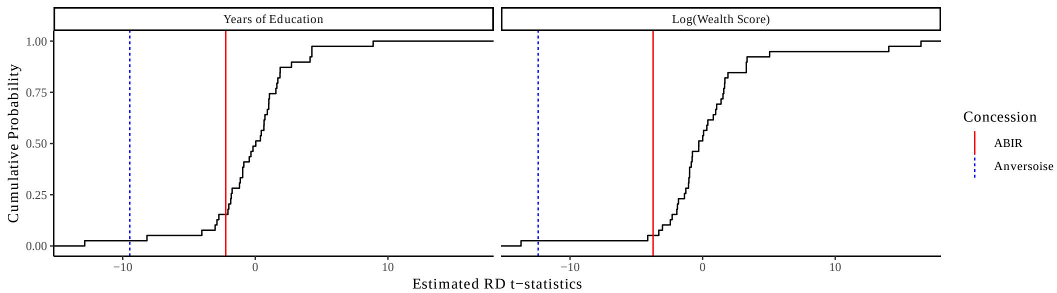
We also present results using a linear latitude and longitude specification in Figure D3. The results are very similar: the Anversoise estimate falls on the far-left of the distribution, and the ABIR estimate is also on the far-left of the distribution and only a few of the river basin estimates are more negative. In particular, the ABIR estimate falls in the bottom 3.44% of this river basin RD

[37] It is important to note that we wouldn't expect the borders to match the modern river basins as rivers can move across decades, and the concession borders were drawn with imperfect maps at a time when the interior of DRC was not mapped in great detail.

estimate distribution while Anversoise falls in the 0.0% of this distribution using the parsimonious latitude and longitude specification.

The results of this basin falsification exercise presents important evidence that the results presented in Section 3.3 are not an artifact of the concessions being drawn using river basins, but instead represent the impacts of the labor coercion during the rubber period.

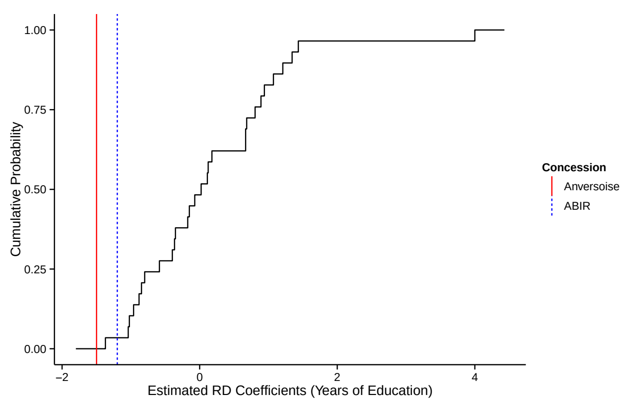
Figure D2: Empirical Cumulative Distribution of RD Estimates for Major River Basins in DRC



Cumulative Distribution and Concession Estimates

Notes: The estimates use our baseline RD specification – local linear specification – within a bandwidth of 25 km from the river basin borders. The solid-red line presents the RD estimate t-statistic corresponding to the ABIR concession border and the dashed-blue line presents the RD estimate t-statistic corresponding to the Anversoise concession border.

Figure D3: Empirical Cumulative Distribution of RD Estimates for Major River Basins in DRC - Relative to Concession Boundary Effects: Linear Latitude and Longitude Specification



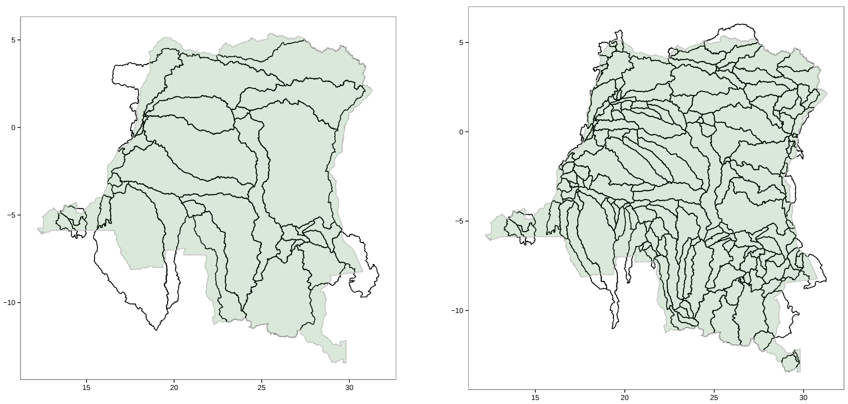
Cumulative Distribution and Concession Estimates

Notes: The estimates use a linear latitude and longitude RD specification within a bandwidth of 100 km from the river basin borders. The solid-red line presents the RD estimate corresponding to the Anversoise concession border and the dashed-blue line presents the RD estimate corresponding to the ABIR concession border.

D.4. Basins as Instruments: Implementation & Results

We use the data from Hydrobasins (Lehner and Grill, 2013) on the geographic extent of two basins used to define the rubber concessions – the Mongala basin and the Maringa-Lopori basin – to create a prediction of the actual boundary line. We do this to address concerns that the actual boundaries may have been manipulated. Figure D5 provides a map of the two relevant basins and our concession borders.

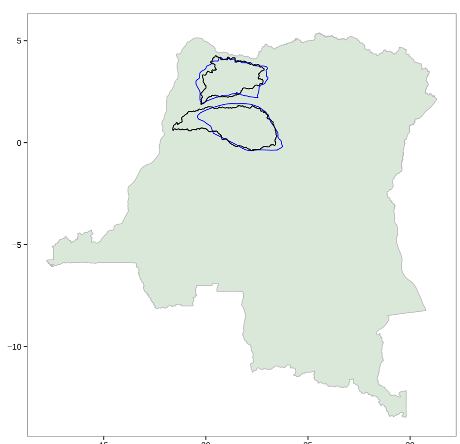
Figure D4: Major River Basins in DRC from HydroBASINS



(a) Level 5 River Basins

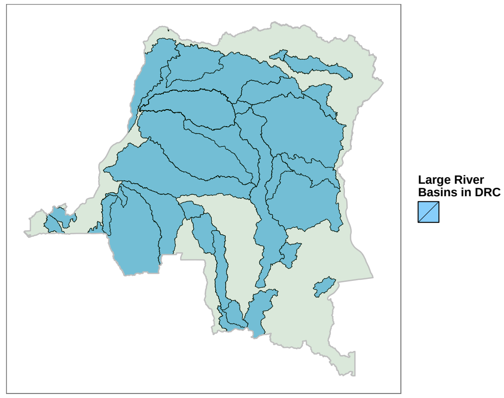
(b) Level 6 River Basins

Figure D5: Concession Borders and Concession River Basins from HydroBASINS



Notes: Concession Borders are outlined in blue and River Basins from HydroBASINS corresponding to the river basins used to define the concessions are outlined black.

Figure D6: River Basins from HydroBASINS - Sample in Falsification Exercise



Notes: The map shows the river basins outlines from HydroBASINS corresponding to the river basins used the falsification exercise. These correspond to all level 5 and level 6 river basins for DRC with at least ten clusters in the DHS data for DRC.

Using this data, we calculate the distance from each DHS cluster to the two basins (with a 25 km buffer around the basins), and determine whether they fall inside or outside the basins. We then estimate an instrumental variable specification, using "inside basin" as an instrument for the actual classification (i.e. whether or not a cluster falls within the former concession boundaries). Table D1 presents reduced-form RD estimates; Table D2 presents the IV estimates. Using these basin borders, we find (i) that the basin classification is a strong predictor of whether or not a DHS cluster falls within the concession borders, and (ii) the estimates effects are quite similar to our main baseline estimates presented in Table 2: areas inside the former concession boundaries tend to have lower education, wealth and less health outcomes.

D.5. Basins as Instruments: Using Placebo Basin Buffers

As a placebo test, we extend the analysis and re-calculate the reduced-form estimates using several different "placebo" buffers around the two relevant basins. We present the results in Figure D7. Consistent with the actual rule used to define the rubber concession borders, we find that the largest negative effects occur precisely using the buffer the colonialists used to define the concession borders (25 kms) and do not find similar negative estimates for other placebo buffers (and generally tend to find positive estimated effects).

Table D1: Rubber Basins and Economic Development
RD Reduced-Form Estimates Using Basin Borders

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Basins	-1.539^{**}	-0.876	-0.142^{**}	-0.062
	(0.656)	(0.534)	(0.064)	(0.055)
Observations	3,083	4,630	3,077	4,616
Clusters	27	85	27	85
Bandwidth	25.00	100.00	25.00	100.00
Mean Dep. Var.	5.119	5.109	0.462	0.465
SD Dep. Var.	3.694	3.821	0.499	0.499
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Basins	-0.699^{***}	-0.444^{***}	-0.224^{***}	-0.192^{***}
	(0.172)	(0.153)	(0.074)	(0.073)
Observations	3,089	4,637	3,089	4,637
Clusters	27	85	27	85
Bandwidth	25.00	100.00	25.00	100.00
Mean Dep. Var.	2.150	2.034	10.957	10.912
SD Dep. Var.	1.119	1.060	0.455	0.443
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Basins	-0.076	-0.074^{**}	-0.130^{***}	-0.166^{***}
	(0.051)	(0.035)	(0.021)	(0.042)
Observations	1,033	1,532	457	1,010
Clusters	27	85	27	85
Bandwidth	25.00	100.00	25.00	100.00
Mean Dep. Var.	0.280	0.259	0.268	0.247
SD Dep. Var.	0.256	0.252	0.332	0.316

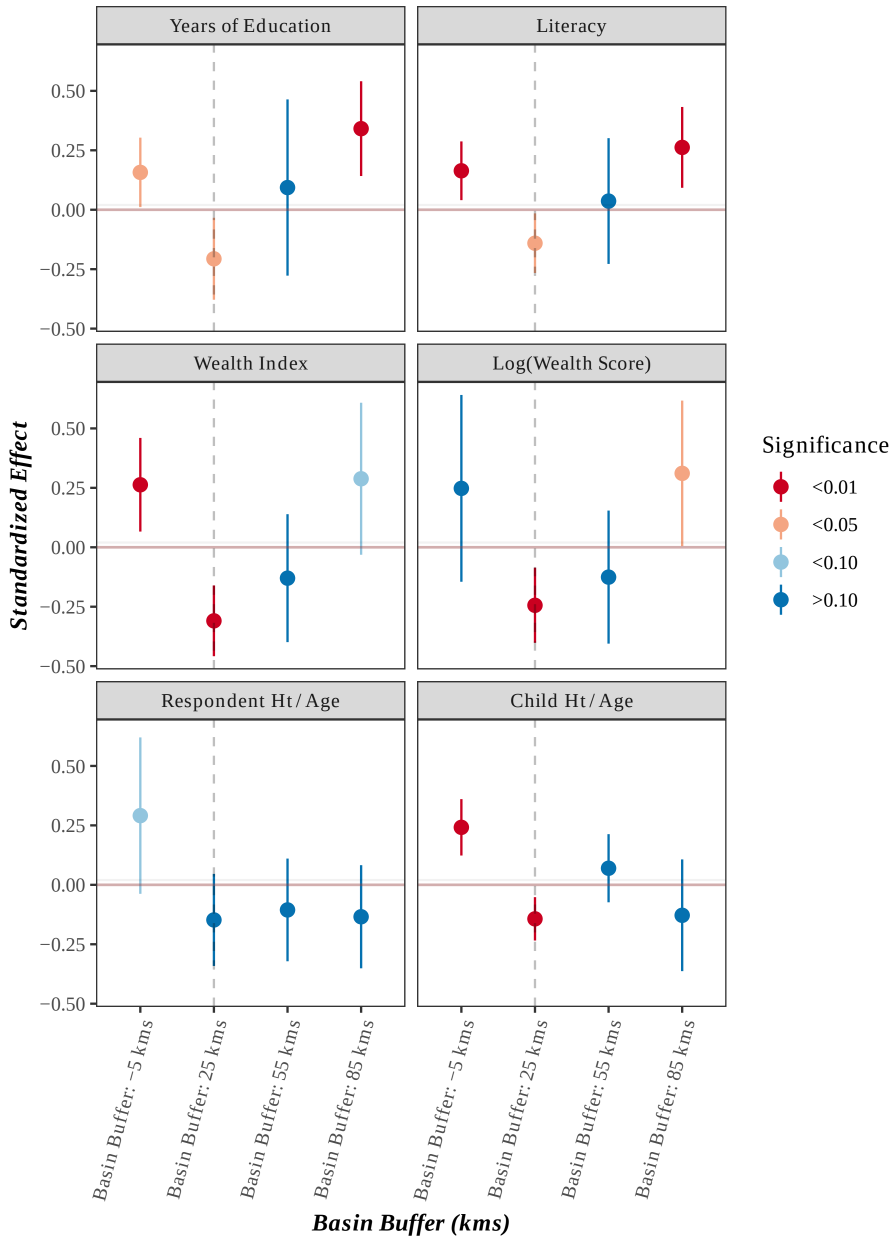
Notes: Standard errors clustered at the DHS cluster level. Regressions use distance to the Mongala and Maringa-Lopori basin as defined by Hydrobasin (with a 25 km buffer around each basin). All regressions include a local linear specification estimated separately on each side of the basin boundary and use a triangular kernel. Regressions control for average age, age squared, gender, survey year, and nearest concession fixed effect. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D2: Rubber Concession and Economic Development
Instrumental Variable Estimates Using Basin Borders

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	−1.859**	−1.743	−0.208***	−0.152
	(0.733)	(1.136)	(0.075)	(0.130)
Observations	1,814	2,623	1,811	2,619
Clusters	37	52	37	52
Bandwidth	35.00	50.00	35.00	50.00
Mean Dep. Var.	5.131	5.209	0.464	0.470
SD Dep. Var.	3.735	3.787	0.499	0.499
F Statistic	36.93	42.75	36.92	42.97
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	−1.126***	−1.373***	−0.568***	−0.766***
	(0.344)	(0.381)	(0.142)	(0.182)
Observations	1,816	2,627	1,816	2,627
Clusters	37	52	37	52
Bandwidth	35.00	50.00	35.00	50.00
Mean Dep. Var.	2.112	2.101	10.944	10.941
SD Dep. Var.	1.109	1.095	0.447	0.458
F Statistic	36.96	42.82	36.96	42.82
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.143***	−0.053	−0.284***	−0.237***
	(0.052)	(0.127)	(0.037)	(0.056)
Observations	617	888	607	822
Clusters	37	52	37	52
Bandwidth	35.00	50.00	35.00	50.00
Mean Dep. Var.	0.266	0.262	0.266	0.738
SD Dep. Var.	0.259	0.255	0.331	0.440
F Statistic	33.41	41.93	40.22	39.85

Notes: Standard errors clustered at the DHS cluster level. Regressions use distance to the Mongala and Maringa-Lopori basin as defined by Hydrobasin (with a 25 km buffer around each basin) as an instrument for being inside a former concession. Regressions control for average age, age squared, gender, survey year, and nearest concession fixed effect. All regressions include a local linear polynomial in distance to the basin borders estimated separately on each side of the basin boundary and use a uniform kernel. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure D7: DHS Results – RD Estimates Using Basis & Placebo Basin Buffers



Notes: Standard errors clustered at the DHS cluster level. Regressions use distance to the Mongala and Maringa-Lopori basin as defined by Hydrobasin, with an additional buffer around each basin denoted on the x-axis. Vertical line denotes basin buffer distance historically used to define the concession areas. All regressions include a local quadratic specification estimated separately on each side of the basin boundaries within a 25 km bandwidth from the boundaries. Regressions control for age, age squared and gender, and nearest basin fixed effects. Regressions use a triangular kernel following Cattaneo et al. (2020). Figure shows point estimates and 95% confidence intervals.

Appendix E. Differences in Subsequent Colonial Policies

In this section, we explore several alternative mechanisms that may explain our observed results. In particular, we explore whether the rubber extraction period (i) affected subsequent Belgian colonial infrastructure investments and missionary presence (ii) altered migration patterns and induced selective migration. In Appendix I.6 we examine differences in market access today using modern road network data, and in Appendix I.7 we analyze differences in population density.

One potential explanation for the differences in development today is that the subsequent Belgian colonial policies were different inside and outside the former concessions. We gathered archival data from the Académie Royale des Sciences d'Outre-Mer (1954) and Rouck (1945) to assess whether colonial policies and investments were different in the former concessions relative to areas just outside the border. In particular, we examine missionary presence and colonial infrastructure investments.

E.1. Missionary Presence

In particular, during the colonial period, Catholic and Protestant missions were the primary providers of education (Hochschild, 1998). The differences in education found in Table 2, could be explained by differences in missionary presence if missionaries or colonial officials decided not to engage as much with the former concession areas. We use data from Nunn (2010) on missionary posts in 1924 and colonial maps from 1897 (from Rouck (1945)) and 1953 (from Académie Royale des Sciences d'Outre-Mer (1954)) to test whether areas inside the concessions had fewer missionary posts.

Panel A of Table E1 presents results from estimating equation (1) on missionary presence in 1897, 1924, and 1953 (see Appendix Figure B5 for a map of mission stations in 1924). We find no evidence that areas inside and outside the concessions had significantly different missionary presence during the colonial period. Additionally, we find no differential Protestant or Catholic presence, nor do we find any differences in type of mission station (e.g. with health center, school or neither).[38] This suggests that differences in outcomes are not driven by subsequent missionary interventions in the areas, nor by the different policies pursued by Protestants and Catholic missions during the colonial era.

E.2. Colonial Infrastructure Investments

Even though the Belgian colonial government was not primarily responsible for the provision of schooling, the government did provide infrastructure investment and other public goods (Van Reybrouck, 2014). If the colonial government chose to invest less in former concessions areas - perhaps due to lower population density as a result of the rubber period - then differences in colonial investments during this period could be a channel through which the rubber areas remain less developed today.

Using colonial data from the Académie Royale des Sciences d'Outre-Mer (1954), we test whether areas inside the former concessions had fewer telecommunication stations and health centers in 1953, and lower road network density in 1968 using maps from the Army Map Services, Corps of Engineers, U.S. Army (1968). The DRC achieved independence in 1960, but dealt with political instability in the subsequent years (Van Reybrouck, 2014); thus, even though the road density data is from after independence, it serves as a reasonable proxy for colonial road investments. The estimates are presented in Panel B of Table E1. We find little evidence that colonial investments in these goods were different inside and outside the concessions: areas

[38] These results are not presented but are available upon request.

Table E1: Rubber Concessions, Missionary Stations, and Colonial Investment

	Panel A: Number of Missionary Stations in:					
	1897		*1924*		*1953*	
	(1)	(2)	(3)	(4)	(5)	(6)
Inside Concession	−0.003	0.002	−0.008	0.011	0.019	0.148
	(0.004)	(0.002)	(0.065)	(0.029)	(0.095)	(0.092)
Bandwidth Choice	Optimal	Wide	Optimal	Wide	Optimal	Wide
Observations	556	853	438	853	485	853
Clusters	25	29	23	29	23	29
Bandwidth	48.00	100.00	30.64	100.00	26.87	100.00
Mean Dep. Var.	0.002	0.001	0.021	0.026	0.545	0.251
SD Dep. Var.	0.045	0.034	0.162	0.179	0.577	0.473
	Panel B: Colonial Infrastructure Investment					
	Number of Telecomm Stations in 1953		*Number of Health Centers in 1953*		*Road Network Density in 1968*	
	(1)	(2)	(3)	(4)	(5)	(6)
Inside Concession	0.049	0.018	−0.034	−0.054	−6.396	−1.486
	(0.050)	(0.021)	(0.076)	(0.041)	(8.067)	(5.397)
Bandwidth Choice	Optimal	Wide	Optimal	Wide	Optimal	Wide
Observations	452	853	504	853	554	853
Clusters	23	29	23	29	23	29
Bandwidth	26.33	100.00	31.79	100.00	33.83	100.00
Mean Dep. Var.	0.030	0.034	0.198	0.245	34.478	35.394
SD Dep. Var.	0.170	0.181	0.459	0.513	33.736	34.370
	Panel C: Market Access					
	Number of Bridges in 2010		*Road Density in 2010*		*Road Density per Capita in 2010*	
	(1)	(2)	(3)	(4)	(5)	(6)
Inside Concession	0.088	−0.037	−5.424	−1.422	−0.380	−0.022
	(0.063)	(0.039)	(10.909)	(7.799)	(0.665)	(0.628)
Bandwidth Choice	Optimal	Wide	Optimal	Wide	Optimal	Wide
Observations	508	853	478	853	651	845
Clusters	23	29	23	29	24	29
Bandwidth	28.83	100.00	32.33	100.00	43.37	100.00
Mean Dep. Var.	0.062	0.055	50.438	57.469	3.982	4.524
SD Dep. Var.	0.289	0.266	41.955	48.047	4.201	4.900

Notes: We present standard errors clustered at the territory level in ( ). All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. The 1924 data are from Nunn (2010). Data from 1897 are from Rouck (1945) and data from 1953 is from the Académie Royale des Sciences d'Outre-Mer (1954). Number of Missionary Stations in each year is a measure of the number of missions in each 20 km by 20 km grid cell for each year that we have a map with the exact locations of missions. In Panels B and C Columns 1-3 control for density of navigable rivers and columns 4-9 control for the percentage of each grid cell that is a river. Data is from 2010 available from the Referentiel Geographique Commun for DRC. *Number of Telecomm Stations in 1953* is defined as the total number of colonial telecommunications stations located in each 20 km by 20 km grid cell in 1953. *Number of Health Centers in 1953* is defined as the total number of colonial health centers stations located in each 20 km by 20 km grid cell in 1953. Data are from Académie Royale des Sciences d'Outre-Mer (1954). *Road Network Density in 1968* is defined as total length in meters of roads in each 20 km by 20 km grid divided by the grid's total surface area in kilometers squared for roads in 1968. Data is from 2010 available from the Referentiel Geographique Commun for DRC. *Number of Bridges* variable is defined as the total number of bridges located in each 20 km by 20 km grid cell. *Road Density* is defined as total length in meters of roads in each 20 km by 20 km grid divided by the grid's total surface area in kilometers squared. Columns 7-9 normalize by population levels using data from Landscan 2007 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

inside the concessions had similar numbers of telecommunication stations and health centers in 1953, and similar (though sightly lower) road network density in 1968.

Appendix F. Migration

F.1. Selective Migration

As mentioned in Section 3.2, a potential channel of interest is selective migration. Selective migration would be a plausible channel of persistence if all of the most capable individuals are leaving the former concession areas and moving to places outside of the former concession, and the relevant determinants of income are highly heritable. Examining the plausibility of this last assumption about the heritability of the relevant determinants of income, especially in a rural setting, is interesting but outside the scope of the paper. Unfortunately, detailed micro-level information on migration rates does not exist for our area of interest. However, in this section we conduct three exercises to examine what the magnitude of present day selective migration would have to be to explain our observed results and to test for heterogeneity in effect sizes by ease of migration.

First, we conduct a trimming exercise with the DHS data to examine whether selective migration might be responsible for the differences in development outcomes between former concessions and areas just outside the former concessions. The intuition is that all of the "good" people inside the concession areas leave and locate just outside the concession areas. Thus, the areas outside the concession appear more developed. We consider how much of the sample we would have to trim in order for our results to lose significance. Specifically, we ask what percentage of the most well-off individuals who reside outside of the concessions would we need to omit so that we no longer observe statistically significant differences between former concession and non-concession areas, under the strong assumption that the x% of the most well-off from outside the concession are actually from inside the concession and that the "good" individuals from outside are not migrating to even better locations. When we examine our education and income results, we find that x would have to be between 16% and 26% to explain the differences we observe. This would imply that for selective migration to fully explain the results, only the "best" people from inside are leaving and that the top one-fifth of the individuals we observe outside the concession are all came from inside the concession.

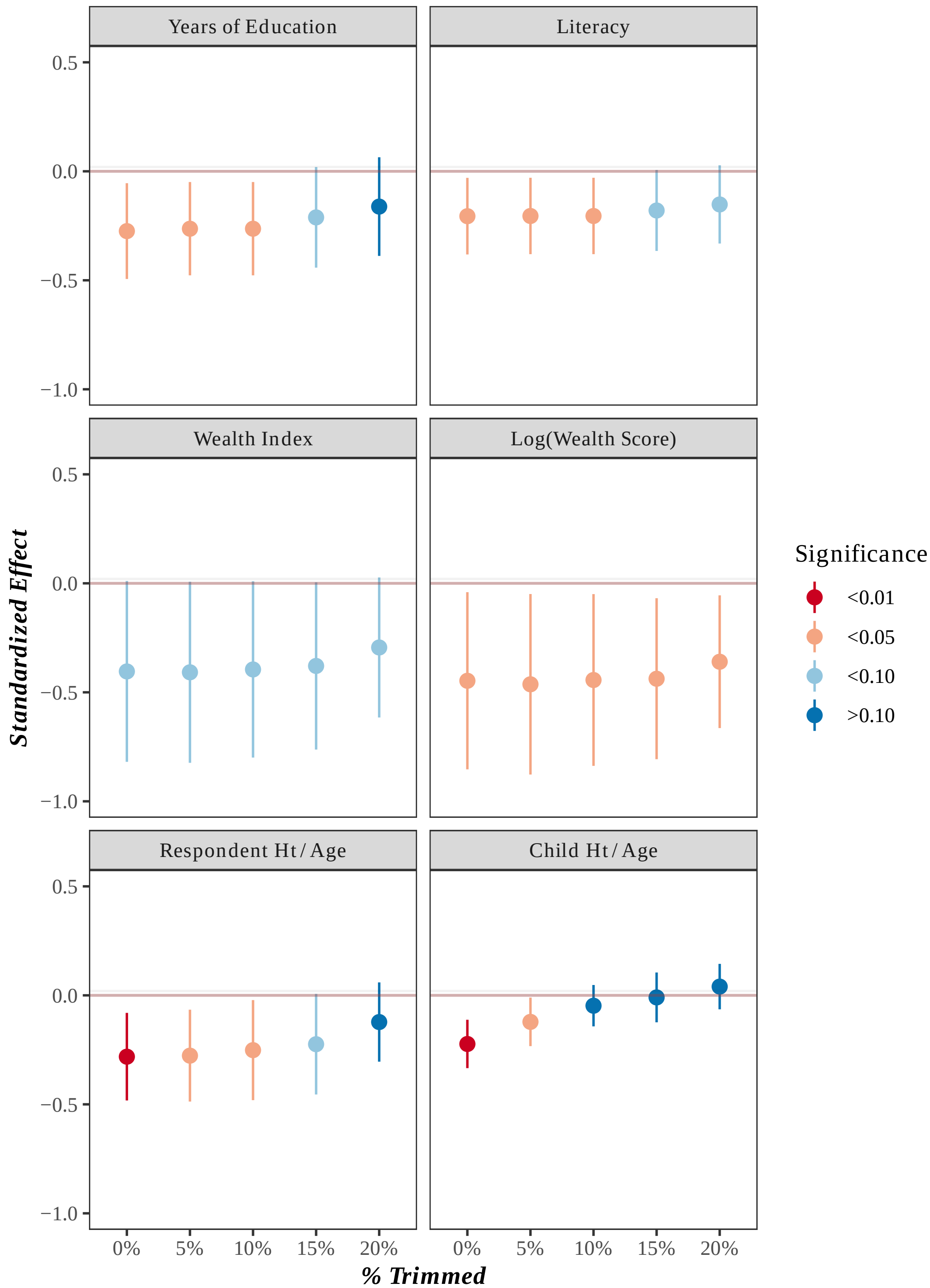
We present the results if we trim the top 0, 5, 10, 15, and 20% of the most well-off individuals and present the results in Figure F1. After trimming the top 10% of the sample outside the concessions but within 200 km, the estimates of the effect of the former rubber concessions remain of similar magnitude and statistical significance. These estimates demonstrate that even under a strong assumption of high levels of selective migration our results remain consistent. However, for higher trimming rates (e.g. 15%), the results tend to no longer be statistically significant. As a point of reference, Dell (2010) omits the top 4.8% using information on migration rates in Peru. A 10% migration rate is a much higher estimate of migration than the differences in population flows we observe using Landscan data from 2007 compared to 2013, or using a measure for population density in 1954 compared to modern Landscan data. In both cases, population growth differences between inside and outside the concessions are a maximum of 4%.

Finally, we examine whether our estimated results differ between places inside the former concessions where it would be easier to migrate compared to places where it would be more difficult to migrate. If selective migration is easy such that more of the "good" people are leaving, we would expect that the RD estimates for villages where it is easier to migrate to be larger and more negative than the estimate for villages for which migration is more difficult. As proxies for ease of migration, we use (i) the colonial road network and (ii) ethnic group boundaries. For (i), we calculate the shortest distance along a road to leave the concession for each village. We then compare villages that have a below median road distance to leave the former concessions to those villages that have an above median road distance to leave the former concessions. Almost all villages happen to be next to a road segment, so this exercise is not comparing villages with roads to those without roads. The intuition is that for villages with a shorter road segment out of the concession, it is easier to migrate. For (ii), we use maps of ethnic groups boundaries from Murdock (1959) and compare villages inside ethnic homelands with an above median share of their ethnic homeland outside of the concession to those villages within ethnic homelands with a below median share of their ethnic homeland outside the concession. The intuition is that leaving the concession is easier for individuals with an above median share of their ethnic homeland outside of the concession. The results for exercise (i) are described and presented in Appendix F.3. The results for exercise (ii) are detailed and presented in Appendix F.4. In both cases, we find that villages where it is easier to migrate are not significantly worse off than villages where it is less easy to migrate, offering additional suggestive evidence that selective migration is not enough to explain the differences in development presented in Section 3.3 In fact, in general, those areas where it is harder to migrate have larger and more negative effect sizes.

The results from these exercises suggest that migration today is likely not the main channel behind the differences between former concessions and neighboring areas. This finding is consistent with a growing literature that highlights a lack of selective migration in developing country settings. For instance, Bazzi, Gaduh, Rothenberg and Wong (2016) examine the Transmigration Program in Indonesia that relocated two million migrants from rural Java and Bali to new rural settlements in the Outer Islands. They find that there has been little selective migration away from the settlements. The results are also consistent with qualitative evidence from our visits to the area: migration to other rural areas is challenging due to poor infrastructure and difficulties in gaining access to land and resources in a different village if one does not originate from that community.

F.2. Trimming for Selective Migration

Figure F1: DHS Results - Trimming for Selective Migration

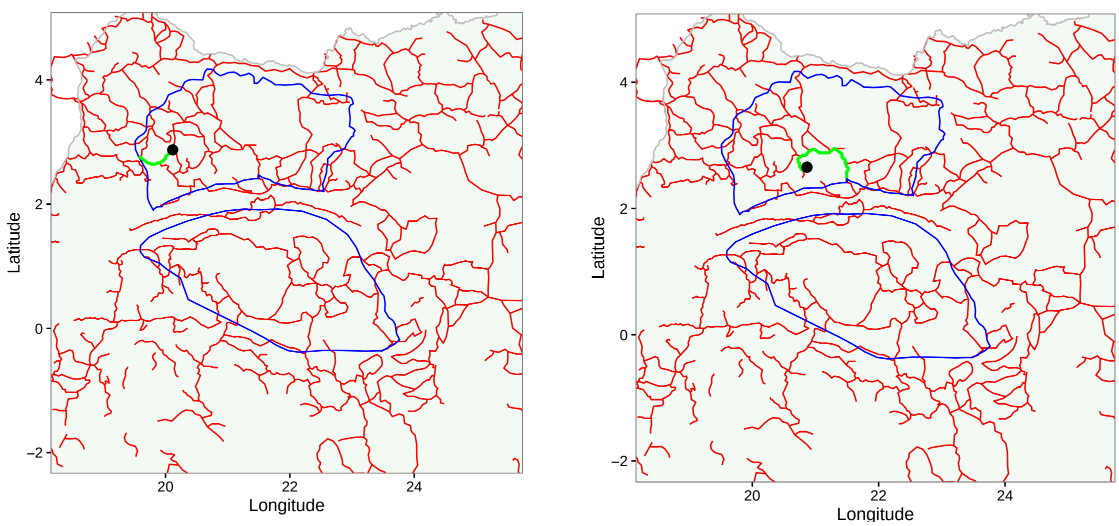


Notes: We trim our sample by dropping the top x% of observations outside of the former concessionsand then estimate our main specification for each dependent variable (where x is denoted on the x-axis). All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. Standard errors are clustered at the DHS cluster level. Figure plots 95% confidence intervals.

F.3. Heterogeneity by Ease of Migration Based on the Colonial Road Network

In this section, we examine how our results differ depending on the ease of migration to outside the former concession. In particular, we examine villages that happen to be on a road segment that has a easy access (i.e. shorter distance) to the concession border relative to those villages who happen to be on roads with harder access to the border. For intuition, consider Figure F2. Here we are considering two villages in the former Anversoise concession with similar euclidean distances to the border but differing road route distances to the border: one village is located on a road that makes it easier to leave along the road network (Figure F2a, optimal path highlighted in green) while the other village is on a road that makes it harder to leave the former concession (Figure F2b). By comparing these two villages, we can examine whether people in the villages where it's easier to get out have worse development outcomes than those from villages where it's harder to get out.

Figure F2: Road Networks and Road Distance to Concession Borders: Examples

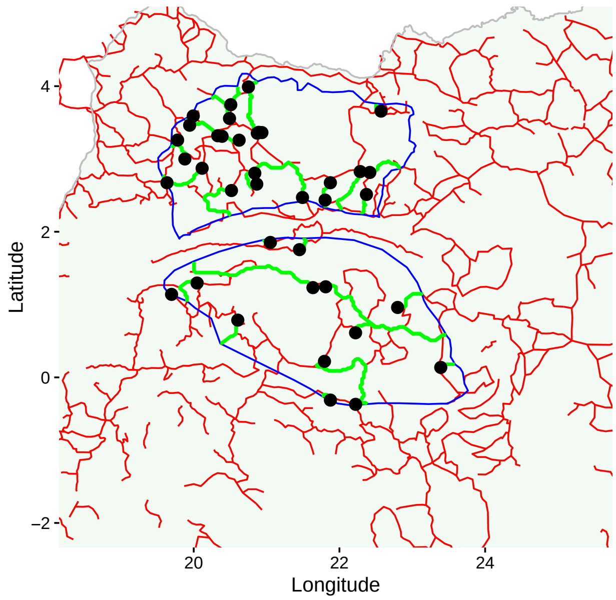


(a) Short Road Distance

(b) Long Road Distance

We proceed in the following way. First, for each village inside the former concessions, we calculate the shortest cost road route from its location to the border. These optimal shortest cost road segments for each cluster in the DHS data are highlighted in the map in Figure F2a in green. Second, for each sub-sample of villages within an Euclidean distance bandwidth away from the concession borders, we split the sample into villages above and below median road distance to the concession borders and estimate our baseline specification. Note that almost all villages happen to fall right next to some road, so this exercise is not simply comparing villages with roads to those without roads. Instead, it is comparing villages with easier access to out-migrate to outside the concessions relative to villages where it is harder to out-migrate due to differences in the road network structure.

Figure F3: Road Networks and Road Distance to Concession Borders: All Clusters



If selective migration is very large, we would expect that the RD estimates for villages with shorter road distance to the former borders (i.e. where it is easier to out-migrate) to be larger than the estimates for villages longer road distances to the former borders (i.e. where it is harder to out-migrate). Table F1 presents the estimates from splitting the sample within each bandwidth as described above. The results suggest that the estimates for the negative impacts of the rubber concession on education and wealth are very similar for both samples; in fact, the negative estimated effect of the rubber concession seems to be slightly larger in places where the road network makes it harder to out-migrate. Thus, areas with easier access to out-migrate are not significantly worse off than villages with less access to out-migrate, offering suggestive evidence against selective migration being a crucial explanation for the results in Section 3.3.

Table F1: Ease of Migration and Economic Development:
Colonial Road Network

	Panel A: < 10 km Road Distance to Border (Easier to Migrate)			
	Years of Education		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	-1.224	-1.024^{*}	-0.400^{*}	-0.491^{***}
	(0.938)	(0.597)	(0.205)	(0.145)
Bandwidth Choice	Narrow	Wide	Narrow	Wide
Observations	1,121	2,659	1,121	2,660
Clusters	37	85	37	85
Bandwidth	35.00	100.00	35.00	100.00
Mean Dep. Var.	5.131	5.109	10.944	10.912
SD Dep. Var.	3.735	3.821	0.447	0.443

	Panel B: > 10 km Road Distance to Border (Harder to Migrate)			
	Years of Education		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	-1.794^{***}	-1.022	-0.335^{**}	-0.154
	(0.363)	(0.702)	(0.170)	(0.132)
Bandwidth Choice	Narrow	Wide	Narrow	Wide
Observations	1,636	4,096	1,638	4,103
Clusters	37	85	37	85
Bandwidth	35.00	100.00	35.00	100.00
Mean Dep. Var.	5.131	5.109	10.944	10.912
SD Dep. Var.	3.735	3.821	0.447	0.443

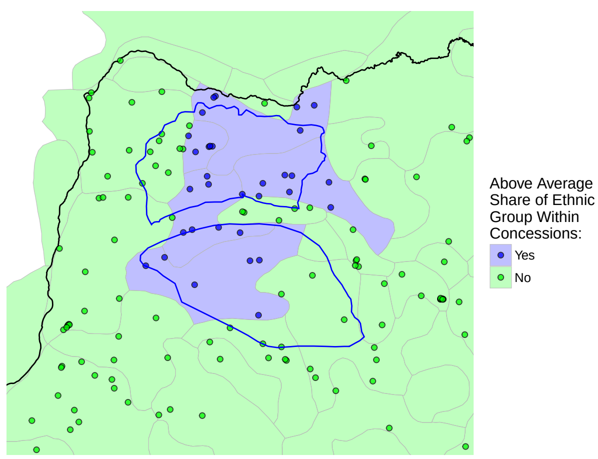
Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Bandwidths are reported in kms. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

F.4. Heterogeneity by Ease of Migration Based on Ethnic Group Boundaries

In this exercise, we use the intersection the rubber concession borders with ethnic group boundaries to examine heterogeneity of our estimated effects by the amount of ethnic group area is outside the concessions. The intuition is to analyze how our results differ depending on whether a village is part of an ethnic group with many possible villages for individuals to migrate outside the concessions, relative to villages part of an ethnic group that does not have many possible villages for individuals to migrate to outside the concession. The idea is that it may be difficult for individuals to migrate to villages where they are not the main ethnic group (due to cultural differences or language differences for instance).

We proceed in three steps. First, using the ethnic group boundaries map from Murdock (1959) and the rubber concession boundaries, for each ethnic group with at least one DHS village within the concessions, we calculate the share of its area that falls within the concessions. Second, we split our DHS sample within the former concessions into those part of an ethnic group with greater than average share of its ethnic group residing outside the concession and those with a lower than average share of its ethnic group residing outside the concession. The average village inside the former concessions is part of an ethnic group with approximately 50% of its area inside the former concessions. Third, we estimate our main specification for education and wealth for these two samples and compare the estimates. Figure F4 demonstrates which ethnic groups near the former concessions have an above and below average share of their area outside the former concessions, and which DHS villages belong to each group.

Figure F4: Ethnic Group Boundaries and Rubber Concessions



If selective migration is very large, we would expect that the RD estimates for villages with a higher share of their ethnic group residing outside former borders (i.e. where it is easier to out-migrate) to be larger than the estimates for villages with a lower share of their ethnic group residing outside the former borders (i.e. where it is harder to out-migrate). Table F2 presents the estimates from splitting the sample as described above. Interestingly, as in Section F.3 the estimates suggest that the (negative) impacts of the rubber concession on education and wealth are very similar for both samples and that the effect seems to be slightly more negative in places where the ethnic boundary locations make it harder to out-migrate. Thus, the results offer suggestive evidence once more that selective migration is unlikely to be a critical explanation for the DHS results from Section 3.3.

Table F2: Ease of Migration and Economic Development:
Ethnic Group Boundaries

	Panel A: < Avg. Share of Ethnic Group Inside Concession (Easier to Migrate)			
	Years of Education		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	-1.961^{***}	-1.410^{**}	-0.370^{**}	-0.266^{*}
	(0.310)	(0.691)	(0.188)	(0.137)
Bandwidth Choice	Narrow	Wide	Narrow	Wide
Observations	1,587	3,502	1,589	3,507
Clusters	37	85	37	85
Bandwidth	35.00	100.00	35.00	100.00
Mean Dep. Var.	5.131	5.109	10.944	10.912
SD Dep. Var.	3.735	3.821	0.447	0.443
	Panel B: > Avg. Share of Ethnic Group Inside Concession (Harder to Migrate)			
	Years of Education		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	-2.006^{***}	-1.592^{***}	-0.362^{***}	-0.289^{**}
	(0.641)	(0.568)	(0.127)	(0.132)
Bandwidth Choice	Narrow	Wide	Narrow	Wide
Observations	1,170	3,253	1,170	3,256
Clusters	37	85	37	85
Bandwidth	35.00	100.00	35.00	100.00
Mean Dep. Var.	5.131	5.109	10.944	10.912
SD Dep. Var.	3.735	3.821	0.447	0.443

Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Bandwidths are reported in kms. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Avg. Share of Ethnic Group Inside Concession* is 0.49. *Share of Ethnic Group Inside Concession* is the share of the Murdock ethnic group polygon that falls inside either rubber concession. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix G. Field Work

G.1. Sampling Procedure

The data were collected between July and August 2015 in Gemena, the capital of Sud-Ubangi province (formerly a part of Equateur province). We used Google satellite imagery from June 2015 to develop a sampling frame. We divided Gemena in to enumeration areas - "polygons" - whose shape was determined by natural boundaries, such as roads and rivers. We estimated the population size within each polygon by counting the number of houses. See Figures G.1, G.1 and G.1 for maps showing the satellite image of Gemena, Gemena divided into sampling polygons, and the sampled polygons and households that participated in the survey (indicated by navy blue dots).

Random sample

For the random sampling, 40 out of the 89 polygons were randomly selected to be visited by survey enumerators. The probability of a polygon being chosen was proportional to its estimated population size. Thus, more populated polygons had a higher probability of being selected. In other words, we used a probability-proportional-to-size (PPS) sampling method. The target number of observations for the study was 520. The number of households visited within each polygon is constant. Thus, 13 households were chosen to be visited in each randomly selected polygon.[39] To ensure geographic coverage of the polygon, enumerators followed a skip pattern within each polygon that depended on the number of houses within that polygon. Due to differences in the size of polygons, this generated a different skip pattern for each polygon. The polygons chosen, their approximate population size, and the skip pattern for each one is shown in Figure G.1 below. Using this sampling method, we visited 506 households within 40 randomly selected polygons during July and August of 2015.

G.2. Data Collection

For each household that was visited, survey team members asked to speak to the head of the household. If the head of the household was not available, the enumerator asked to interview an adult member of the household, with a preference for older household members. In order to avoid survey fatigue and improve engagement, we split up our survey into two parts, the first one conducted on the first visit and the second part in the second visit, along with some behavioral games. Below, we explain each visit in-depth.

First Visit - Main Survey

The first survey sections consisted of questions intended to identify the respondent's ethnic group and village and territories of origin and of birth. The survey collected information on basic demographics, migration, the institutions of the individual's village of origin, politics, values, and religion. In addition, for a final section of the survey, enumerators asked to speak to the oldest member of the household to collect data on her views of various historical events in the region. Respondents' were also asked to take a picture, if willing, in order to make relocating them for the second visit easier. The survey also contained detailed question on respondents' village of origin – on the status of the village, public goods available, trade and political institutions – but these were only asked to respondents who were at least somewhat familiar with their village of

[39] One polygon had to be dropped once enumeration began as most of it was a military complex and our enumerators were not granted access inside of it.

Figure G1: Satellite Image of Gemena

Figure G2: Satellite Image of Gemena with Sampling Polygons

Figure G3: Satellite Image of Gemena with Sampled Polygons and Households Visited

origin to improve the quality of the data. For this first visit, 506 households were visited and 503 households agreed to participate in the first survey.[40]

Second Visit - Survey and Behavioral Games

During the second visit, enumerators were asked to revisit the original households if the respondents' village of origin was within 200 kms of the former rubber concessions. The villages were located on maps by the respondent and enumerator during the first visit, and the distance to the former borders was done in R between the two visits. In total, there were 484 households out of 503 who fit this criteria and were therefore selected for the second visit. Enumerators were tasked with finding the same respondent as in the first visit to conduct a short survey, a variant of the Reverse Dictator Game (DG), a time and risk module, and an Implicit Association Test (IAT) on views of local chiefs. The Reverse DG and the IAT are explained in detail in Sections G.3 and G.4. The survey was always conducted first, and then the order of the three subsequent activities was randomized. When enumerators were unable to track down the original respondent due to travel or illness,they were asked to first attempt to locate another member of the same household and arrange two visits (to conduct the survey from the first part and the second visit); if they could not locate another household member, they were asked to use the same sampling method and skip pattern for that polygon to attempt to find a replacement household. This method of replacement resulted in 29 households being replaced in order to reach our target of 484 households for the second visit.

Third Visit - Payments from Behavioral Games

Finally, enumerators conducted a final visit that only consisted of payment for the outcomes of the Reverse DG. No survey was conducted in this visit; respondents only had to sign a receipt of payment and, if willing, take a photo with their sealed envelope containing the payment.[41]

[40] Three households refused to participate in the survey and were therefore not included in the survey.

[41] The pictures were taken to verify payment.

Figure G4: Sampling Frame

(1) Polygon ID	(2) Number of Households in Polygon	(3) Skip Pattern: Visit Every x Households
62	128	10
19	120	9
70	63	5
18	130	10
40	133	10
74	305	23
3	80	6
45	155	12
14	102	8
77	146	11
56	75	6
27	488	38
15	361	28
2	120	9
26	257	20
69	97	7
28	95	7
78	289	22
23	334	26
87	136	10
41	253	19
59	96	7
24	462	36
35	110	8
46	201	15
1	48	8
8	179	14
73	195	15
42	187	14
83	76	6
60	171	13
50	65	5
55	175	13
48	80	6
86	31	5
38	191	15
84	184	14
29	180	14
13	239	18
16	168	13

G.3. Reverse Dictator Game

Description of Game

During the second visit, we asked participants to play a variation of the Dictator Game (DG) proposed by (Jakiela, 2011) to experimentally measure an individual's *respect for earned property rights*. In this section we first explain the basic outline of the game and then provide the detailed experimental instructions we used.

In the standard DG, one player (*Player 1*) is allocated an amount of money (*budget*) and asked to allocate it between themselves (the"dictator") and another subject (*Player 2*). In the (Jakiela, 2011) variant, there are two differences from the standard DG: (i) instead of having an external *budget* endowed to *Player 1*, *Player 1* must perform a real effort task to earn the *budget*, and (ii) instead of *Player 1* being the "dictator", now *Player 2* is the "dictator" and gets to decide how to divide *Player 1*'s earned income between themselves and *Player 1*.

Variation (i) of the DG has been used before by Hoffman et al. (1994) and Cherry et al. (2002) subjects tend to be much less generous when they earned their own income, which Farh and Irlenbusch (2000) refer to as *earned property right*. Variation (ii) on its own changes the standard DG to what is known as a Reverse DG, which has been used many times before List (2007, see). Jakiela (2011) combines these two variations to get a measure of respect for earned property rights and finds that subjects in the US tend to others' respect earned income much more than subjects in Kenya. The amount *Player 2* decides to take from *Player 1*'s earned income therefore represents a measure for the *respect for earned property rights*.

In our experiment, every respondent is matched to an anonymous, randomly selected individual from Gemena. This individual was chosen from within our sample and matches remained completely anonymous to everyone on the team except for the authors. This removed strategic considerations from the decisions of the participants on how much income to take from others. Additionally, every respondent plays the game twice: once as *Player 2* and then as *Player 1*. Respondents first learned about the general structure of the experiment, the details of the earning task, and then decided whether to participate or not.

Before performing the effort task (i.e. played as *Player 1*), subjects decide how they wanted to take from an anonymous *Player 1*'s income. We used the strategy method to elicit these divisions: for each of 20 possible earnings, respondents would enter the amount they wished to take for themselves. The share of earned income that *Player 2* decides to take from *Player 1*'s earned income is our measure for the *respect for earned property rights*.

For the earnings task, we selected a task that could be easily understood by all respondents and in which more effort was rewarded by more income: subjects played a "clicking-game" on touch screen tablets. In this "clicking-game", a small blue dot appears in a random location on the screen every three seconds and the respondent has one second to click on it before it disappears.Importantly, this effort task did not rely on physical strength or skill for effort but instead relied more on concentration and perseverance for effort. It is purposefully a very boring game. The game lasted five minutes and respondents were paid based on the number of successful "clicks", earning 100 Congolese Francs (approximately $0.10) for each 10 successful clicks. Respondent were very engaged in the task as participants earned on average 700 CF in this task, a significant amount for this region of DRC equal to about 1 days wage. Figure G.3 provides a picture of the basic layout of the game. The game was preformed on seven-inch Samsung Galaxy II tablets.

The game was conducted in private between the participant and the enumerator at the home of the respondent. Out of the 484 second visit participants, 482 total individuals participated in this reverse dictator game variation. Two individuals chose to skip participation, one due to poor eyesight and the other because her husband refused giving her and the enumerator permission to conduct the game in privacy. On average, participants chose to take close to 40% of the other

Figure G5: Clicking Task

individuals' earnings when playing this game. Below, we present the experimental instructions we used to conduct the game. These are presented in English; we translated these instructions to french and then Lingala using back translation to verify the accuracy of the translations.

Experimental Instructions - Reverse Dictator Game

[Find a private place to meet with the same respondent whom you interviewed for the survey. It is very important that the player will not be watched by members of his household or other people while he or she is playing the games.]

Now I will explain how to play this game. It is very important to pay attention because only those who understand the rules of the game well will be able to play. Let me remind you that this project is completely voluntary and you are free to leave at any time if you decide that you do not want to participate in this game.

This game is played in pairs: there is a player 1 and a player 2. Importantly, you will play both roles today.

You will play with someone chosen randomly from the population of Gemena. Neither you nor I will know exactly who you are playing with. Only one person in our research office will know who plays with who, and he will never tell anyone.

In this game, Player 1 earns money by performing some task and Player 2 will decide how they want to divide the money that Player 1 earns between them and the other player. Player 1 will earn money by performing a clicking task that will explain soon. The performance in this task will determine how much money is given to each pair of players. For example, if Player 1 earns 1000 FC, then Player 2 will decide how to divide the 1000 FC given to the players; if Player 1 earns 100 FC, then Player 2 will decide how to divide the 100 FC given.

Let me explain how the task is paid out. In this game, Player 1 will earn money by playing a clicking game on the tablet. In this game, a blue button will appear on the screen and the Player 1 must click the button to earn a point. The button will appear in different parts of the screen and it is Player 1's job to find it and click it before it disappears. Each time Player 1 presses the button the number of clicks on the screen increases by one. [Show player the image of the clicking game.]

Every time Player 1 clicks the blue button, the number goes up by one - it never decreases - and the button will change color. So, the number at the top of the tablet will show the number of times Player 1 has clicked. In this game, Player 1 will be paid money based on the number of times she clicks. We'll give

Player 1 five minutes in which Player 1 can click as much as Player 1 can. Player 1 has to play the game with only one hand. Player 1 cannot switch hands, or click with two hands. The more times Player 1 clicks the blue button during the five minutes, the more money Player 1 will be paid.

PAYOUTS

How much Player 1 is paid depends on how much Player 1 clicks. For each number of times she might click, she will earn 10 FC per click. We will then round this number to the closest hundred value. This sheet shows you how much she will be paid. So, Player 1 will be paid 100 FC if Player 1 clicks between 50 times and 149 times. If Player 1 clicks between 150 and 249 times, she will be paid 200 FC. The more times Player 1 clicks, the more money Player 1 earns.

- 0-4 Cliquers: 0 FC
- 5-14 Cliquers: 100 FC
- 15-24 Cliquers: 200 FC
- 25-34 Cliquers: 300 FC
- 35-44 Cliquers: 400 FC
- 45-54 Cliquers: 500 FC
- 55-64 Cliquers: 600 FC
- 65-74 Cliquers: 700 FC
- 75-84 Cliquers: 800 FC
- 85-94 Cliquers: 900 FC
- 95-104 Cliquers: 1000 FC
- 105-114 Cliquers: 1100 FC
- 115-124 Cliquers: 1200 FC
- 125-134 Cliquers: 1300 FC
- 135-144 Cliquers: 1400 FC
- 145-154 Cliquers: 1500 FC
- 155-164 Cliquers: 1600 FC
- 165-174 Cliquers: 1700 FC
- 175-184 Cliquers: 1800 FC
- 185-194 Cliquers: 1900 FC
- >195 Cliquers: 2000 FC

[Check that the player has understood how Player 1 will be paid depending on the number of clicks.]

Player 2 must then decide how to divide the money between himself and player 1. Player 2 must take between 0 and 1000 FC from player 1, but the total amount possible will depend on the effort made by Player 1. Player 2 takes home what he takes from player 1, and player 1 takes home the rest. Now, we are going to run through some examples to show how this game can be played.

[Take the money in your hands for these demonstrations and push the offer made to player 2 across a line on the floor.]

1. Here is the first example. Imagine that Player 1 earns 1000 FC. Player 2 then chooses to take 900 FC from Player 1. Then, Player 2 will go home with 900 FC. Player 1 will go home with 100 FC (1000 FC minus 900 FC equals 100 FC).
2. Here is another example. Imagine that Player 1 earns 600 FC. Player 2 then chooses to take 200 FC from Player 1. Then, Player 2 will go home with 200 FC. Player 1 will go home with 400 FC (600 FC minus 200 FC equals 400 FC).

3. Here is another example. Imagine that Player 1 earns 1000 FC. Player 2 then chooses to take 500 FC from Player 1. Then, Player 2 will go home with 500 FC. Player 1 will go home with 500 FC (1000 FC minus 500 FC equals 500 FC).
4. Here is another example. Imagine that Player 1 earns 700 FC. Player 2 then chooses to take 700 FC from Player 2. Then, Player 2 will go home with 700 FC. Player 1 will go home with 0 FC (700 FC minus 700 FC equals 0 FC).
5. Here is another example. Imagine that Player 1 earns 500 FC. Player 2 then chooses to take 0 FC from Player 1. Then, Player 2 will go home with 0 FC. Player 1 will go home with 500 FC (500 FC minus 0 FC equals 1000 FC).

Now please respond to the following test questions to be sure that you have understood. Then, I will tell you if you are a player 1 or a player 2 and you will begin to play. You will play as both Player 1 and Player 2 today, and we will return with your payouts in the next week.

[Use the following list as test questions. If it is necessary to ask more test questions, start again with the first example above and write "test questions repeated" in the notes section.]

For all the following questions, imagine Player 1 has earned 1000 FC:

Test Questions

1. Imagine that Player 2 chooses to take 1000 FC from Player 1. How much will Player 2 go home with? [1000] And how much will Player 1 go home with? [0]
2. Now imagine that Player 2 chooses to take 400 FC from Player 1. How much will Player 2 go home with? [400] How much will player 1 go home with? [600]
3. Now imagine that Player 2 chooses to take 600 FC from Player 1. How much will Player 2 go home with? [600] How much will player 1 go home with? [400]
4. Now imagine that Player 2 chooses to take 100 FC from Player 1. How much will Player 1 go home with? [900] How much will player 2 go home with? [100]
5. Now imagine that Player 2 chooses to take 800 FC from Player 1. How much will Player 1 go home with? [200] How much will player 2 go home with? [800]
6. Now imagine that Player 2 chooses to take 300 FC from Player 1. How much will Player 1 go home with? [700] How much will player 2 go home with? [300]

Now that you fully understand the game, do you still want to participate?

For this activity, first you are Player 2. The Player 1 you play with will be someone chosen randomly from the population of Gemena who has performed the clicking task. Remember that only one person in our research office will know who plays with who, and he will never tell anyone. Now I will ask you how much money you would take from Player 1 depending on how much they earned at the task:

1. If Player 1 earned 1000 FC, how much of this amount, if anything, would you take from Player 1?
2. If Player 1 earned 1000 FC, how much of this amount, if anything, would you take from Player 1?
3. If Player 1 earned 1000 FC, how much of this amount, if anything, would you take from Player 1?
4. If Player 1 earned 1000 FC, how much of this amount, if anything, would you take from Player 1?
5. If Player 1 earned 1000 FC, how much of this amount, if anything, would you take from Player 1?
6. If Player 1 earned 1000 FC, how much of this amount, if anything, would you take from Player 1?
7. If Player 1 earned 1000 FC, how much of this amount, if anything, would you take from Player 1?
8. If Player 1 earned 1000 FC, how much of this amount, if anything, would you take from Player 1?
9. If Player 1 earned 1000 FC, how much of this amount, if anything, would you take from Player 1?

10. If Player 1 earned 1000 FC, how much of this amount, if anything, would you take from Player 1?
11. If Player 1 earned 1000 FC, how much of this amount, if anything, would you take from Player 1?
12. If Player 1 earned 900 FC, how much of this amount, if anything, would you take from Player 1?
13. If Player 1 earned 800 FC, how much of this amount, if anything, would you take from Player 1?
14. If Player 1 earned 700 FC, how much of this amount, if anything, would you take from Player 1?
15. If Player 1 earned 600 FC, how much of this amount, if anything, would you take from Player 1?
16. If Player 1 earned 500 FC, how much of this amount, if anything, would you take from Player 1?
17. If Player 1 earned 400 FC, how much of this amount, if anything, would you take from Player 1?
18. If Player 1 earned 300 FC, how much of this amount, if anything, would you take from Player 1?
19. If Player 1 earned 200 FC, how much of this amount, if anything, would you take from Player 1?
20. If Player 1 earned 100 FC, how much of this amount, if anything, would you take from Player 1?

Now that you have told me what amounts you would take from Player 1, our research office will calculate your payoff after comparing your responses with the amount earned by Player 1. I will return in one or two weeks with your payment for these activities.

Now, you are a player 1. The player 2 you play with will be someone chosen randomly from the population of Gemena. You will never know with whom you are playing, and this player 2 will never know that he is playing with you.

You will earn money by clicking a game on the tablet like this (demonstrate). We will then match you with a Player 2 who will decide how much to take from what you earn, and we will return with your payment in two or three weeks.

[Ask the player if he/she remembers how the clicking and payouts work. If he/she is uncertain, explain the following:]

Let me explain how the task is paid out. In this game, Player 1 will earn money by playing a clicking game on the tablet. In this game, a blue button will appear on the screen and the Player 1 must click the button to earn a point. The button will appear in different parts of the screen and it is Player 1's job to find it and click it before it disappears. Each time Player 1 presses the button the number of clicks on the screen increases by one. **[Show player the image of the clicking game.]**

Every time Player 1 clicks the blue button, the number goes up by one - it never decreases - and the button will change color. So, the number at the top of the tablet will show the number of times Player 1 has clicked. In this game, Player 1 will be paid money based on the number of times she clicks. We'll give Player 1 five minutes in which Player 1 can click as much as Player 1 can. Player 1 has to play the game with only one hand. Player 1 cannot switch hands, or click with two hands. The more times Player 1 clicks the blue button during the five minutes, the more money Player 1 will be paid.

PAYOUTS

How much Player 1 is paid depends on how much Player 1 clicks. For each number of times she might click, she will earn 10 FC per click. We will then round this number to the closest hundred value. This sheet shows you how much she will be paid. So, Player 1 will be paid 100 FC if Player 1 clicks between 5 times and 15 times. If Player 1 clicks between 15 and 24 times, she will be paid 200 FC. The more times Player 1 clicks, the more money Player 1 earns.

- 0-4 Cliquers: 0 FC
- 5-14 Cliquers: 100 FC
- 15-24 Cliquers: 200 FC
- 25-34 Cliquers: 300 FC
- 35-44 Cliquers: 400 FC
- 45-54 Cliquers: 500 FC

- 55-64 Cliquers: 600 FC
- 65-74 Cliquers: 700 FC
- 75-84 Cliquers: 800 FC
- 85-94 Cliquers: 900 FC
- 95-104 Cliquers: 1000 FC
- 105-114 Cliquers: 1100 FC
- 115-124 Cliquers: 1200 FC
- 125-134 Cliquers: 1300 FC
- 135-144 Cliquers: 1400 FC
- 145-154 Cliquers: 1500 FC
- 155-164 Cliquers: 1600 FC
- 165-174 Cliquers: 1700 FC
- 175-184 Cliquers: 1800 FC
- 185-194 Cliquers: 1900 FC
- >195 Cliquers: 2000 FC

[Check that the player has understood how he/she will be paid depending on the number of clicks. Then, go to the "Effort Task" game on the home screen. Enter the respondent's information and give the Player the tablet. Have them play the game and, after the 5 minutes, then return to geoodk.]

[Read the conclusion only after having administered the activity.] Thank you for participating in this game. The player 2 you will be playing with will be drawn randomly from the population of Gemena and will decide how to divide the money you just earned. I will return in one or two weeks to give you this money.

G.4. Implicit Association Test (IAT)

During the second visit, we also conducted a Single-Target Implicit Association Test (ST-IAT) on local chief authority to measure implicit attitudes towards local chiefs. The ST-IAT was developed by Bluemke and Friese (2008) and is a variant of the original IAT. The ST-IAT was created to measure the positivity or negativity of individuals' implicit association of a single target – in our case, this is local chiefs. ST-IATs have been used very recently in similar settings in the DRC by Lowes, Nunn, Robinson and Weigel (2015) and Lowes et al. (2017).

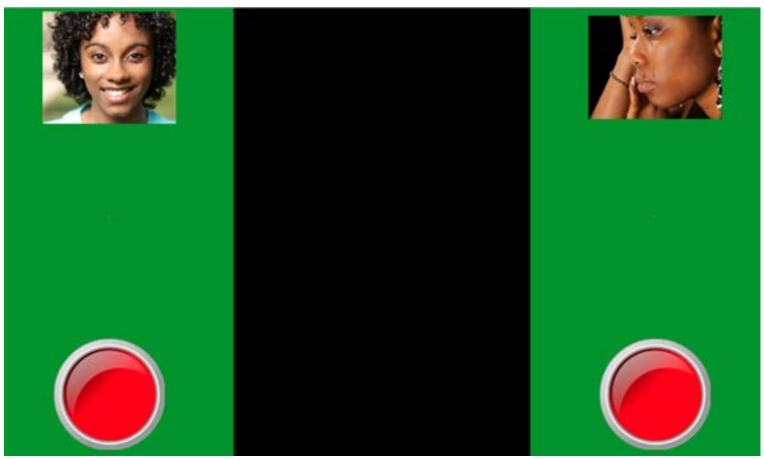
The ST-IAT (henceforth IAT) asks respondents to sort words into two groups, one group on the left side and the other on the right side of the screen. Three different sets of words are presented audibly: words related to happiness, words related to sadness, and words related to local chiefs. The IAT consists of two blocks: in one happy words and chiefs words are sorted left and sad words to the right, and the other happy words are sorted to the left and chief words and sad words are sorted to the right. The order of the blocks is randomized across individuals.

The intuition behind the IAT is that if a respondent has a positive view of chiefs, he/she will have an easier time sorting chief words to the left with happy words than to the right with sad words. The respondent would be using a subconscious heuristic that good things go left and bad things go right (Lowes et al., 2015) . If a respondent does not have a positive association with chiefs, then this heuristic will not apply and the opposite heuristic will be used; he/she will find it easier to sort chief words to the right instead. By examining the difference in the speed at which the respondent sorts the words across the two blocks we can infer their implicit view of chiefs.

Formally, we follow Lowes et al. (2015) and calculate the standard *D-Score* as our inferred measure of the implicit view of chiefs for a given respondent. The *D-Score* is defined as: *D-Score* $= [Mean(latency^{-ve}) - Mean(latency^{+ve})]/SD(latency^{+ve\,and\,-ve})$, where $Mean(latency^{-ve})$ is the average response time in milliseconds for the block in which the chief words are meant to go right, $Mean(latency^{-ve})$ is the average response time for the block in which the chief words are meant to go left, and $SD(latency^{+ve\,and\,-ve})$ is the standard deviation in response times across both blocks. In this *D-Score*. more positive values will indicate more positive implicit views.

The IATs were played on seven-inch Samsung Galaxy II tablets with Panasonic RP-HT21 Lightweight Headphones connected to them. The respondents always played a practice block first that asks them to sort only happy words and sad word; this allows them to get used to the interface, the headphones and the tablet. To sort a word to the left (right), the participant presses the red button on the left (right) side of the screen, presented in Figure **??**. In every block of the IAT, participants had to obtain a 75% success rate in sorting words to the correct side in order to advance to the next block. If they did not meet this success rate, they had to repeat the block. Figure **??** presents a screenshot of the practice block set-up. There is an image of a happy person on the left and an image of a sad person on the right to help the participant with the sorting directions.

Figure G6: IAT Screenshot - Practice Block



After the practice block, the participant engages in two more blocks where three different sets of words are presented audibly: words related to happiness, words related to sadness, and words related to local chiefs. These serve as the main two block of the the IAT: in one happy words and chiefs words are sorted left and sad words to the right, and the other happy words are sorted to the left and chief words and sad words are sorted to the right. The order of the blocks is randomized across individuals. In all blocks, happy words are sorted left and sad words are sorted left. Each time a participant gets to sort a new word is called a trail, and each block consists of 24 trails: 8 trails with happy words, 8 trails with sad words, and 8 trails with chief words. The order of the trails was randomized within each block. The full list of words used is presented in Table G1. Figures G.4 and G.4 present screenshots of the two blocks: the first image of the block in which chiefs words are sorted with happy words, and the second one of the block in which chief words are sorted with sad words. The screen displays the word "kapita" – the word for local chief in Lingala – on the left or right side of the screen to help the participant with the sorting.

As stated earlier, examining the difference in the speed at which the respondent sorts the words across the two blocks allows us to infer their implicit view of chiefs using the standard *D-Score* as our inferred measure of the implicit view of chiefs for a given respondent. We follow Lowes et al. (2015) and ignore data from practice blocks and repeated blocks where the participant did not achieve a 75% success rate. We windsorize the response times (also known as latency) to 3,000

Figure G7: IAT Screenshot - Block B

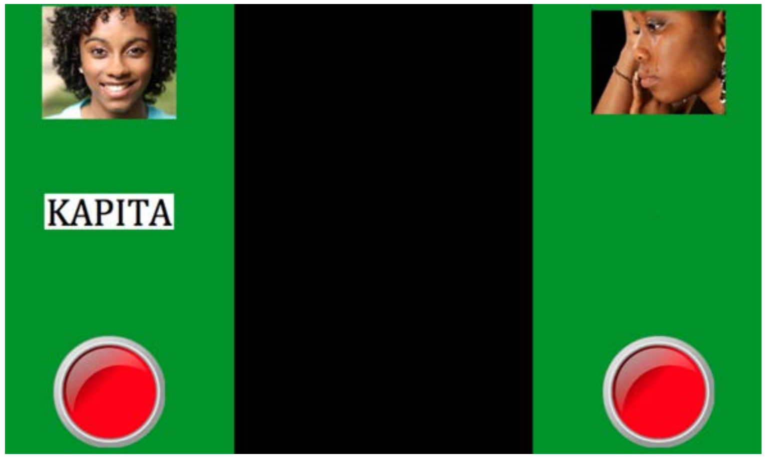
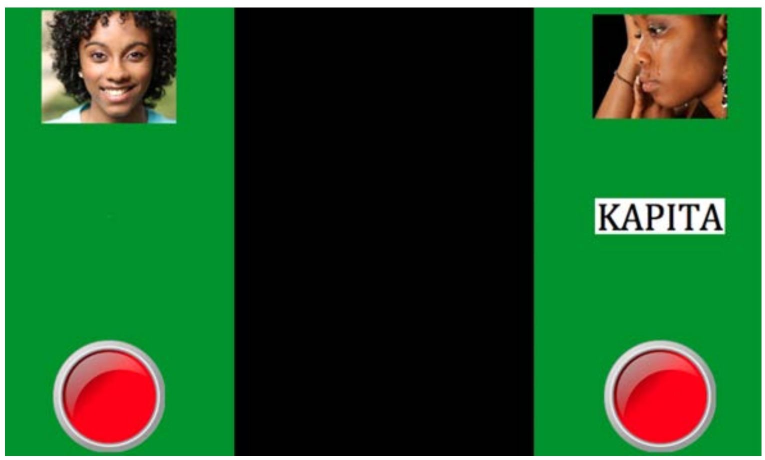


Figure G8: IAT Screenshot - Block C



milliseconds. We account for incorrect replacing their latency with the block mean plus the block standard deviation.

The IAT instructions we used borrow heavily from Lowes et al. (2015). Lowes et al. (2015) conducted ST-IATs in the DRC during 2014 and their findings confirm that the single-target IAT succeeds in capturing participants' implicit attitudes in a very similar setting. These instructions are presented below in English. We translated the instructions, the tablet audio, and the tablet-game prompts into Lingala using back translation methods to verify the accuracy and consistency of the translations.[42]

During the second visit, 459 participated completely in the activity. Many participants refused to participate, either due to poor eyesight, hearing problems, and/or sickness. Some participants refused to complete the activity due to failing to achieve a 75% success rate and grew fatigued of having to repeat blocks. Thus, we have 459 IAT observations from our second visits to test for implicit views of chiefs.

[42] We only performed the IAT in Lingala as everyone in our sample for the second visits spoke Lingala, the Lingala for the game is very basic, and the chief authority words have more meaning in Lingala than their equivalents in French for the local context.

Table G1: Words used in the IAT

Happy Words		*Sad Words*		*Chief Authority Words*	
English	Lingala	English	Lingala	English	Lingala
Love	Bolingo	Pain	Bolosi	Chief	Kapita
Generosity	Boboto	Failure	Kokueya	Chiefs	Bakapita
Laughter	Koseka	Suffering	Kpokoso	Village Chief	Mokonzi
Joy	Sai	Bad	Kitkote	Village Chiefs	Bakonzi
Glory	Kembo	Horrible	Somo	Sub-tribe Chief	Kumu
Happiness	Esengo	Wrong	Mabe	Sub-tribe Chiefs	Bakumu
Pleasure	Kosepela	Wicked	Motomboki	Local Leader	Mokambi
Sympathetic	Motema Kitoko	Terrible	Yakobangisa	Arbiter of Village Conflict	Mokubua

Notes: English and French words for the Happy Words and Sad Words come from Lowes, Nunn, Robinson and Weigel (2015).

Experimental Instructions - IAT

[Find a private place to meet with the same respondent whom you interviewed for the survey. It is very important that the player will not be watched by members of his household or other people while he or she is playing the games.]

Now I will explain how to play this game. It is very important to pay attention because only those who understand the rules of the game well will be able to play.

Let me remind you that this project is completely voluntary and you are free to leave at any time if you decide that you do not want to participate in this game.

First Block:

We are going to play a computer game. Before we play I would like to ask you to put on these headphones. If they are too loud or are uncomfortable, please let me know so I can adjust them.

[Ask participant if he/she will put on headphones]

[If participant doesn't want to put on headphones then wait until the start of the second block to ask if they changed their mind. If they still decline, unplug the headphones and use the computer's internal speakers. But make sure the volume isn't so loud that other people can hear.]

You are going to hear sounds when an dot appears in the middle of the screen one at a time. Some words will be good words and some words will be bad words. If you hear a good word like happy or nice I want you to press the left button as fast as you can. There is a smiley face on the left side to remind you to press the left button when you hear a good word. But if you hear a bad word like wicked or mad I want you to press the right button as fast as you can. There is a frowny face on the right side to remind you to press the right button when you hear a bad word.

Now, there are a few things to remember.

1. Please use one finger for each button. [Demonstrate by holding one finger by both buttons and pressing each one at a time.]

2. After you press the button be sure to take your finger off of it because if you hold it down [demonstrate holding it down], the button will stop working.

3. Please play the game as fast as you can. It is okay if you make mistakes, I just want to see how quickly you can play. But if you do press the wrong button, just press the correct one and keep playing.

Now that you fully understand the game, do you still want to participate?

[If the person indicates yes, administer the game.]

[Have participant put one finger by each button before beginning the first block.]

Are you ready to play the game?

[Make sure the participant has one finger by each button and is ready to begin before starting.]

Some sounds will be words related to happiness and words related to sadness. If you hear a sound related to happiness, I want you to press the red button on the left as quickly as you can; There is a picture of a smiling face on the left side to remind you to press the red when you hear words related to happiness. Finally, if you hear a sound related to sadness, I want you to press the right button as quickly as you can; There is a picture of a sad face on the right side to remind you to press the red button on the right when you hear words related to sadness.

Remember: please try to play the game as fast as you can. It is okay if you make mistakes.

Second Block:

This next activity will be a bit more complicated. You are going to hear words about chief authority and hear words about happiness or sadness one at a time. If you hear a word related to chief authority in the middle of the screen I want you to press the red button on the left side of the screen as quickly as you can like you were doing earlier. There is a picture of the word "chief" on the right side to remind you to press the red button on the right when you hear sounds about chief authority. If you hear a good word I also want you to press the left button as fast as you can like you were doing before. And if you hear a bad word I also want you to press the red button on the right side of the screen as fast as you can like you were doing before.

Remember: please try to play the game as fast as possible, and don't worry about making mistakes.

[When they get to the break in the middle, say:]

That was great.

Third Block:

Now they've changed sides on you. This time, if you hear a word related to chief authority please press the red button on the right side of the screen as quickly as you can.

[Point out that the category reminders – i.e. the word for chief - has switched sides when you are reciting the instructions].

But, as before if you hear a good word I also want you to press the left button as fast as you can like you were doing before. And if you hear a bad word I also want you to press the right button as fast as you can like you were doing before.

Are you ready to play?

G.5. Summary Statistics

Table G2: Fieldwork: Summary Statistics for the Full Sample

	Individuals Within 200 kms of Concession Borders			
	Mean Inside	Mean Outside	Clustered S.E.	(p-value)
Educational Attainment	2.594	2.626	0.085	0.708
Obs.	254	257		
Years of Education	7.165	7.798	0.351	0.072
Obs.	231	233		
Student	0.091	0.093	0.024	0.907
Obs.	254	257		
Income: Last Week	11,770	13,312	2,907	0.596
Obs.	254	257		
Primary Earner	0.626	0.607	0.047	0.687
Obs.	254	257		
Married	0.839	0.794	0.032	0.167
Obs.	254	257		
Female	0.374	0.424	0.046	0.273
Obs.	254	257		
Age	39.260	39.988	1.269	0.566
Obs.	254	257		

Notes: Data collected in Gemena, DRC during the summer of 2015. Standard errors are clustered at the village of origin level. Educational Attainment is a 0 to 4 categorical variable where 0 is no education and 4 is higher education. Student is an indicator variable equal to 1 if the respondent is currently a student. Income: Last Week is self-reported income level by respondents. Primary Earner is an indicator variable equal to 1 if the respondent is currently the primary earner for his/her household. Married is an indicator variable equal to 1 if the respondent is currently married. Female is an indicator variable equal to 1 if the respondent is a female.

G.6. Migrant Characteristics

Table G3: Differences in Migrant Characteristics and Reasons for Migration

	Individuals From Within 200 kms of Concession Borders			
	Mean Inside	Mean Outside	Clustered S.E.	(p-value)
First-Generation Migrant	0.142	0.171	0.032	0.353
Obs.	254	257		
Father Migrant	0.642	0.650	0.045	0.857
Obs.	254	257		
Mother Migrant	0.685	0.638	0.041	0.255
Obs.	254	257		
Father: Migrated to Find Better Economic Opportunities	0.393	0.425	0.060	0.587
Obs.	163	167		
Father: Migrated to Find Better Educational Opportunities	0.074	0.072	0.028	0.951
Obs.	163	167		
Mother: Migrated to Find Better Economic Opportunities	0.034	0.073	0.025	0.123
Obs.	174	164		
Mother: Migrated to Find Better Educational Opportunities	0.023	0.018	0.015	0.757
Obs.	174	164		
Migrant Father Educational Attainment	2.410	2.181	0.136	0.094
Obs.	139	144		
Migrant Mother Educational Attainment	0.962	1.185	0.145	0.125
Obs.	157	151		

Note: Data collected in Gemena, DRC during the summer of 2015. Standard errors are clustered at the village of origin level. First-Generation Migrant is an indicator variable equal to 1 if the respondent is a first generation migrant. Father Migrant is an indicator variable equal to 1 if the respondent's father is a migrant to Gemena. Mother Migrant is an indicator variable equal to 1 if the respondent's mother is a migrant to Gemena. Father: Migrated to Find Better Economic Opportunities is an indicator variable equal to 1 if the respondent's father migrated to Gemena in search of better economic opportunities. Father: Migrated to Find Better Educational Opportunities is an indicator variable equal to 1 if the respondent's father migrated to Gemena in search of better educational opportunities. Mother: Migrated to Find Better Economic Opportunities is an indicator variable equal to 1 if the respondent's mother migrated to Gemena in search of better economic opportunities. Mother: Migrated to Find Better Educational Opportunities is an indicator variable equal to 1 if the respondent's mother migrated to Gemena in search of better educational opportunities. Migrant Father Educational Attainment and Migrant Mother Educational Attainment are 0 to 4 categorical variables where 0 is no education and 4 is higher education for migrant parents.

Table G4: Fieldwork: Summary Statistics for First-Generation Migrants

	First-Generation Migrants From Within 200 kms of Concession Borders			
	Mean Inside	Mean Outside	Difference	(p-value)
Migrated to Find Better Economic Opportunities	0.222	0.205	0.018	0.849
Obs.	36	44		
Migrated to Find Better Educational Opportunities	0.167	0.205	-0.038	0.647
Obs.	36	44		
Migrated due to Disagreement with Family or Villagers	0.056	0.068	-0.013	0.816
Obs.	36	44		
Migrated with Parents (as a Child)	0.167	0.341	-0.174	0.090
Obs.	36	44		
Outcast from Village	0.028	0.023	0.005	0.889
Obs.	36	44		
Years of Education	7.562	9.353	-1.790	0.057
Obs.	32	34		
Student	0.111	0.227	-0.116	0.168
Obs.	36	44		
Income: Last Week	4,689	20,629	-15,940	0.010
Obs.	36	44		

Notes: Data collected in Gemena, DRC during the summer of 2015. Standard errors are clustered at the village of origin level. Student is an indicator variable equal to 1 if the respondent is currently a student. Income: Last Week is self-reported income level by respondents. Migrated to Find Better Economic Opportunities is an indicator variable equal to 1 if the respondent migrated to Gemena in search of better economic opportunities. Migrated to Find Better Educational Opportunities is an indicator variable equal to 1 if the respondent migrated to Gemena in search of better educational opportunities.

Table G5: Fieldwork: Summary Statistics for Second-Generation or Higher Migrants

	Second-Generation Migrants From Within 200 kms of Concession Borders			
	Mean Inside	Mean Outside	Difference	(p-value)
Educational Attainment	2.578	2.554	0.024	0.798
Obs.	218	213		
Years of Education	7.101	7.533	-0.432	0.255
Obs.	199	199		
Student	0.087	0.066	0.021	0.363
Obs.	218	213		
Income: Last Week	12,940	11,800	1,140	0.720
Obs.	218	213		
Primary Earner	0.651	0.620	0.032	0.540
Obs.	218	213		
Married	0.844	0.789	0.055	0.120
Obs.	218	213		
Female	0.374	0.424	0.046	0.273
Obs.	254	257		
Age	39.48	39.78	-0.298	0.826
Obs.	218	213		

Note: Data collected in Gemena, DRC during the summer of 2015. Standard errors are clustered at the village of origin level. Educational Attainment is a 0 to 4 categorical variable where 0 is no education and 4 is higher education. Student is an indicator variable equal to 1 if the respondent is currently a student. Income: Last Week and Income: Last Month are self-reported income levels by respondents. Primary Earner is an indicator variable equal to 1 if the respondent is currently the primary earner for his/her household. Married is an indicator variable equal to 1 if the respondent is currently married. Female is an indicator variable equal to 1 if the respondent is a female.

Table G6: Survey and Experimental Measures of Trust and Sharing Beliefs
Second-Generation Migrants

	Panel A: Trust and Closeness			
	Trust Index		*Closeness Index*	
	(1)	(2)	(3)	(4)
Inside Concession	0.548***	0.362***	0.669***	0.564***
	(0.165)	(0.108)	(0.229)	(0.138)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	281	393	269	393
Clusters	153	272	144	272
Bandwidth	28.66	100.00	25.81	100.00
Mean Dep. Var.	0.177	0.111	0.207	0.105
SD Dep. Var.	0.635	0.616	0.693	0.696
	Panel B: Survey Measures of Sharing Norms			
	Respondent		*Village of Origin*	
	(1)	(2)	(3)	(4)
Inside Concession	0.458	0.318	0.582**	0.263*
	(0.324)	(0.211)	(0.267)	(0.153)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	243	382	227	306
Clusters	123	263	120	219
Bandwidth	21.43	100.00	25.02	100.00
Mean Dep. Var.	0.060	0.005	0.036	0.008
SD Dep. Var.	0.870	0.851	0.801	0.785
	Panel C: Experimental Measures of Sharing Norms			
	Dictator Game: Share Sent		*Effort Task: Share Redistributed*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.024	−0.005	0.053	0.048**
	(0.026)	(0.021)	(0.033)	(0.024)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	224	374	309	373
Clusters	117	262	162	261
Bandwidth	20.09	100.00	34.08	100.00
Mean Dep. Var.	0.449	0.444	0.405	0.403
SD Dep. Var.	0.125	0.126	0.129	0.127

Notes: Standard errors clustered at the origin village level. Sample includes only individuals born in Gemena. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions include nearest concession fixed effects and control for age, age squared and sex. *Trust Index* is a summary index for the following questions: How much do you trust (1) people from your village of origin, (2) people of another tribe, (3) people of your own tribe, (4) people you meet for the first time, (5) your family, (6) your neighbors, (7) people of another nationality, and (8) people of your sub-tribe; all questions answered on a 0 (Not at All) to 4 (Completely) scale. *Closeness to Others Index* is a summary index for the following questions: (1) How close to you feel to people from your village of origin?, (2) How close to do you feel to people of Gemena?, (3) How close do you feel to people of your own tribe?, (4) How close do you feel to people of your age set from your origin village?, and (5) How close do you feel to people of your age set in Gemena?; all questions answered in a scale from 0 (Not Close at All) to 5 (Very Close). *Sharing Norms Index* is a summary index for the following questions: (1) If you get money from luck you should share it, (2) If you earn money from hard work you should share it, (3) If someone else earns money from luck they should share it, (4) If someone else earns money from hard work they should share it; all questions answered in a scale from 1 (Strongly Disagree) to 5 (Strongly Agree). *Sharing Norms Index Village of Origin* is a summary index for the following questions, where all questions start with "How much would someone from your village of origin agree with the following statements", for the same statements listed above. *Dictator Game: Amount Shared* measures the amount sent to an anonymous player 2 in the standard Dictator Game. *Effort Task: Share Redistributed* is the total share taken (weighted by the maximum budget amount possible to take) in the effort task from the anonymous player 1's earned income. It represents an experimental measure of respect for earned income property rights. Two individuals declined participating in the Dictator Game, and one additional individual declined participating in the Reverse Dictator Game. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table G7: Survey and Experimental Measures of Trust and Sharing Beliefs
First-Generation Migrants

	Panel A: Trust and Closeness			
	Trust Index		*Closeness Index*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.451	−0.041	0.056	0.377
	(0.660)	(0.484)	(0.468)	(0.352)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	50	72	56	72
Clusters	30	67	30	67
Bandwidth	29.64	100.00	33.27	100.00
Mean Dep. Var.	-0.322	-0.343	-0.305	-0.284
SD Dep. Var.	1.048	0.998	0.757	0.746
	Panel B: Survey Measures of Sharing Norms			
	Respondent		*Village of Origin*	
	(1)	(2)	(3)	(4)
Inside Concession	0.647	0.538***	0.546	0.383
	(0.424)	(0.208)	(0.427)	(0.347)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	40	71	40	57
Clusters	24	66	23	54
Bandwidth	25.16	100.00	33.02	100.00
Mean Dep. Var.	0.093	0.172	-0.164	-0.001
SD Dep. Var.	0.536	0.652	0.732	0.780
	Panel C: Experimental Measures of Sharing Norms			
	Dictator Game: Share Sent		*Effort Task: Share Redistributed*	
	(1)	(2)	(3)	(4)
Inside Concession	0.104***	0.037	0.124	0.078
	(0.028)	(0.050)	(0.095)	(0.071)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	30	64	44	64
Clusters	22	62	33	62
Bandwidth	23.13	100.00	35.91	100.00
Mean Dep. Var.	0.450	0.450	0.446	0.417
SD Dep. Var.	0.101	0.102	0.168	0.169

Notes: Standard errors clustered at the origin village level. Sample includes only first-generation migrants to Gemena. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions include nearest concession fixed effects and control for age, age squared and sex. *Trust Index* is a summary index for the following questions: How much do you trust (1) people from your village of origin, (2) people of another tribe, (3) people of your own tribe, (4) people you meet for the first time, (5) your family, (6) your neighbors, (7) people of another nationality, and (8) people of your sub-tribe; all questions answered on a 0 (Not at All) to 4 (Completely) scale. *Closeness to Others Index* is a summary index for the following questions: (1) How close to you feel to people from your village of origin?, (2) How close to do you feel to people of Gemena?, (3) How close do you feel to people of your own tribe?, (4) How close do you feel to people of your age set from your origin village?, and (5) How close do you feel to people of your age set in Gemena?; all questions answered in a scale from 0 (Not Close at All) to 5 (Very Close). *Sharing Norms Index* is a summary index for the following questions: (1) If you get money from luck you should share it, (2) If you earn money from hard work you should share it, (3) If someone else earns money from luck they should share it, (4) If someone else earns money from hard work they should share it; all questions answered in a scale from 1 (Strongly Disagree) to 5 (Strongly Agree). *Sharing Norms Index Village of Origin* is a summary index for the following questions, where all questions start with "How much would someone from your village of origin agree with the following statements", for the same statements listed above. *Dictator Game: Amount Shared* measures the amount sent to an anonymous player 2 in the standard Dictator Game. *Effort Task: Share Redistributed* is the total share taken (weighted by the maximum budget amount possible to take) in the effort task from the anonymous player 1's earned income. It represents an experimental measure of respect for earned income property rights. Two individuals declined participating in the Dictator Game, and one additional individual declined participating in the Reverse Dictator Game. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G.7. Index Components Coefficient Plots

This section presents the coefficient plots for the ICW indexes presented in the main text. For each index, we plot the estimated index coefficient and then plot the standardized individual component coefficient for each component that is included in the index. The coefficient plots are presented for regression discontinuity estimates using the MSE optimal bandwidth from Cattaneo et al. (2020) in kms. All estimates use a local linear polynomial in distance to the concession border estimated separately on each side of the border. Regression control for age, age squared, gender, and include a nearest concession fixed effect. Standard errors are clustered at the village of origin level. The plots include 95% confidence intervals for each coefficient.

Figure G9: Index Components Coefficient Plots - Village Public Goods – Objective

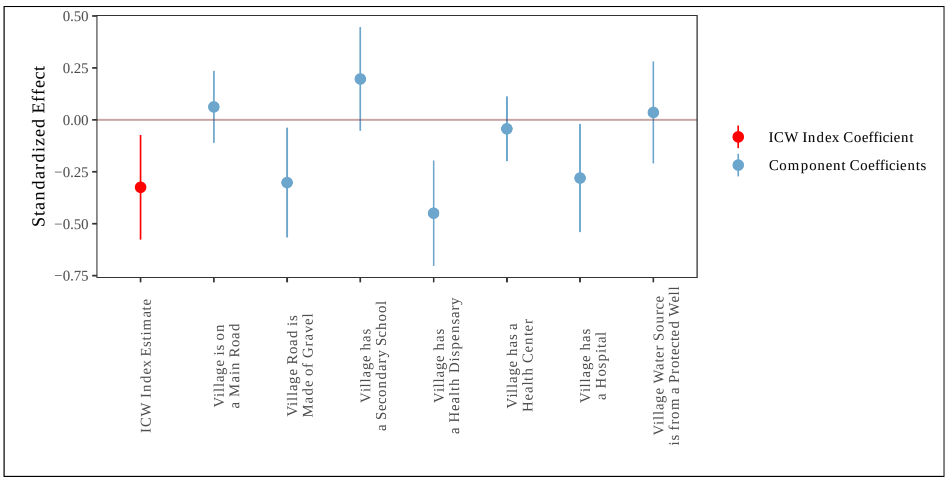


Figure G10: Index Components Coefficient Plots - Village Public Goods – Subjective

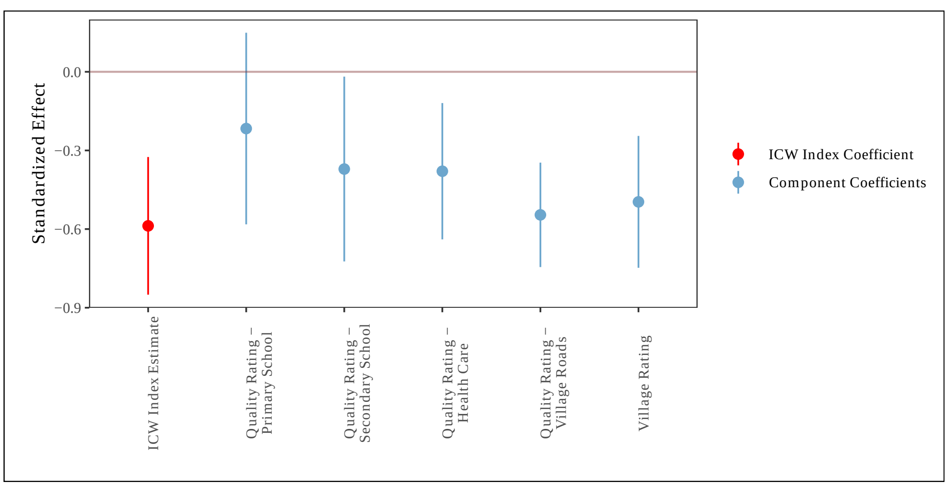


Figure G11: Index Components Coefficient Plots - Chief Public Good Provision

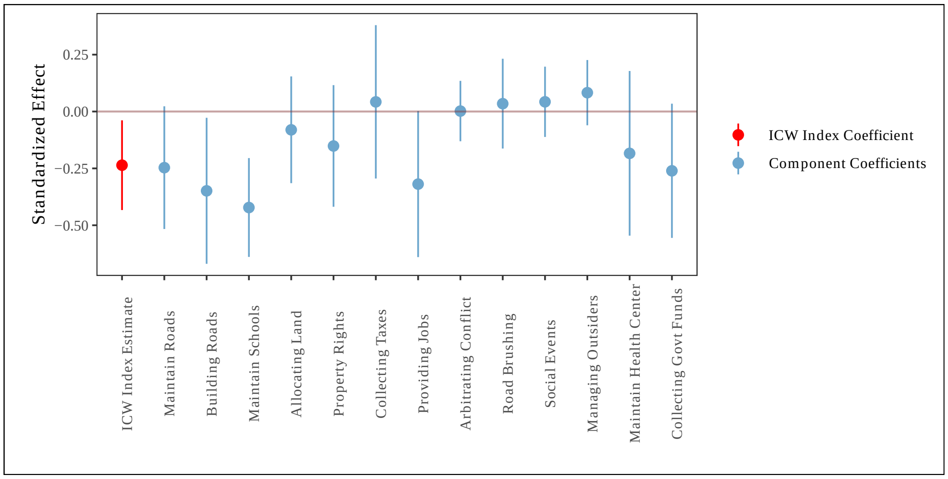


Figure G12: Index Components Coefficient Plots - Respect for Authority

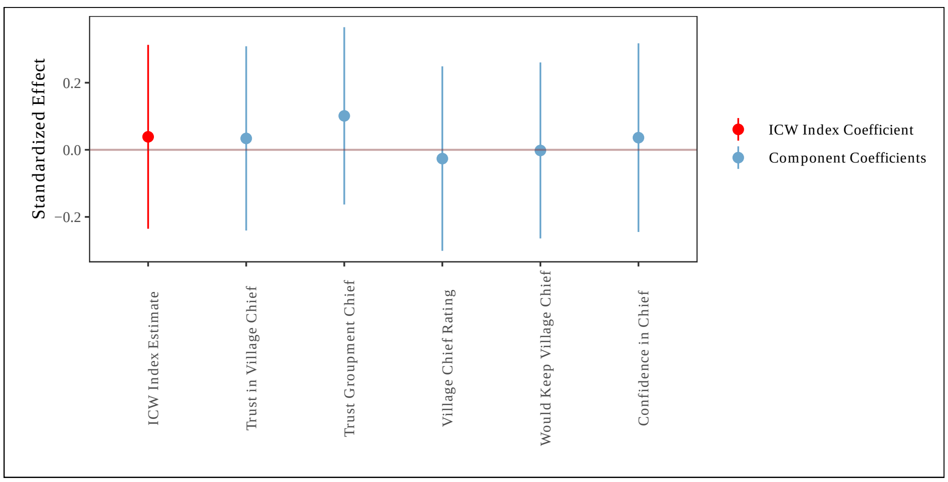


Figure G13: Index Components Coefficient Plots - Trust in Others

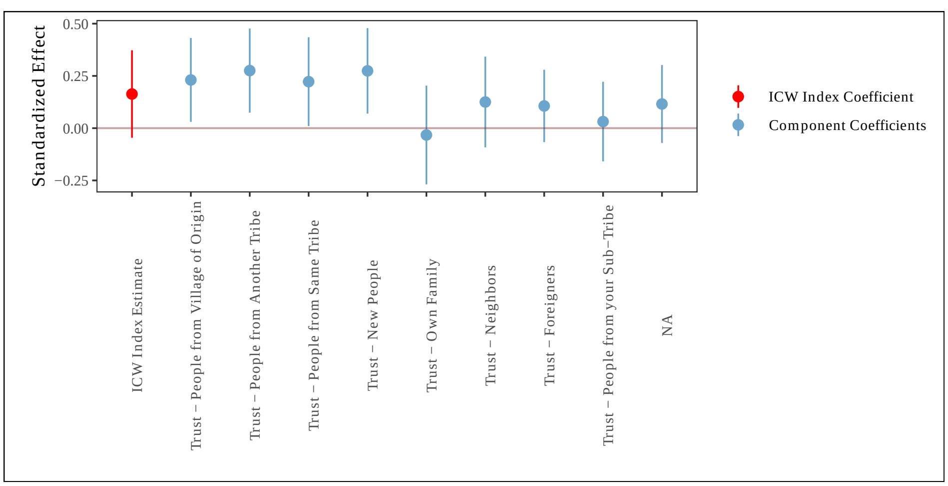


Figure G14: Index Components Coefficient Plots - Closeness

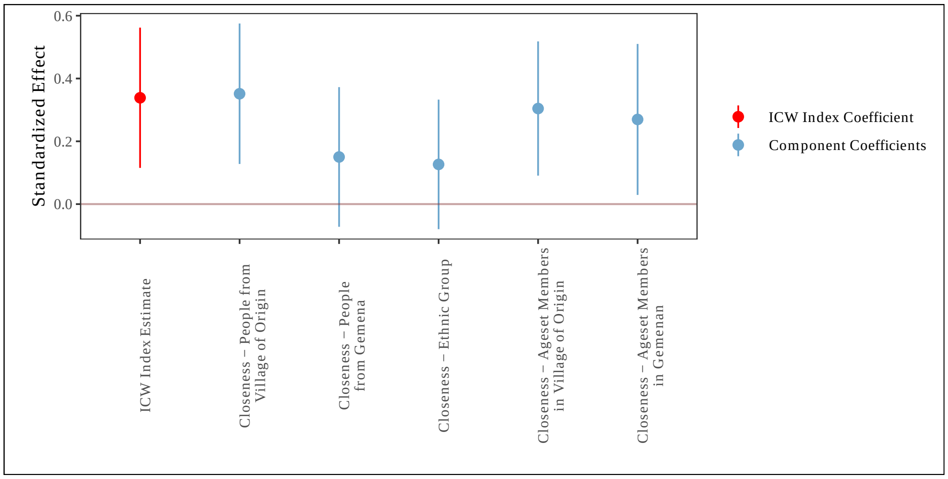


Figure G15: Index Components Coefficient Plots - Individual Sharing Norms

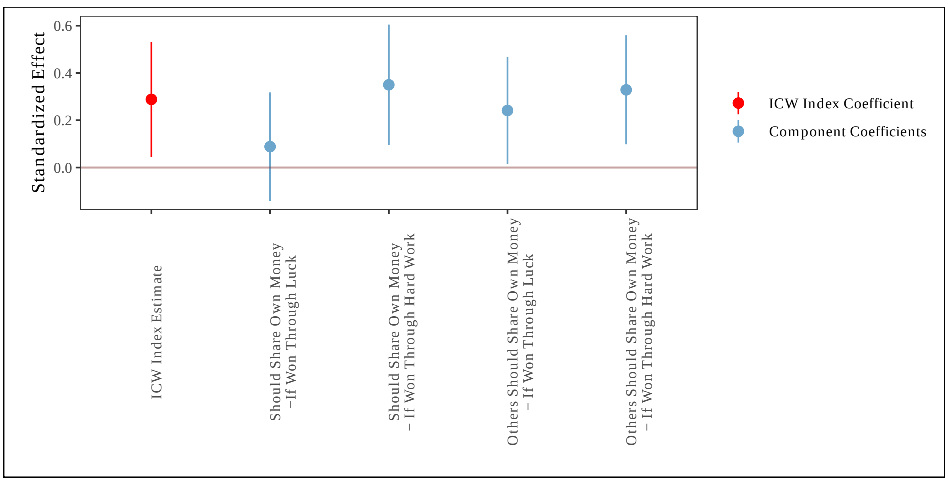
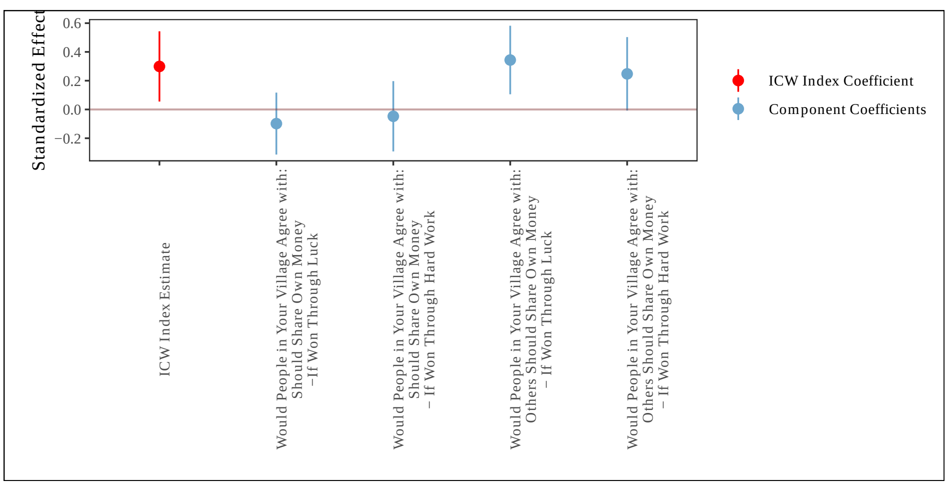


Figure G16: Index Components Coefficient Plots - Village Sharing Norms



Appendix H. Results using Baseline Data from an Ongoing RCT in Northern DRC

This section replicates the regression discontinuity results from Section 4 using baseline data conducted in 302 villages in northern DRC from Lowes et al. (2016). We first describe the data collection process. We then discuss our results using this data, and show that the results confirm the findings in Section 4: we find that villages inside the former concession have less accountable chiefs, but respondents have higher sharing norms.

H.1. Data

Before implementing the proposed intervention described in Lowes et al. (2016), we undertook a round of baseline data collection in northern DRC. Our sample comprises 302 randomly-selected villages from the territories of Gemena (102), Kungu (100) and Lisala (100), which are located in the Northwest of the DRC.

To select the 302 villages we first compiled a Census of villages in each territory from the territory governor's office of each territory (where a village is defined as a collection of households managed by a village chief according to records on village chiefs compiled by territory administrators). We then randomly selected 100 villages from each territory. A map of the villages in the sample is shown in Figure H1.

The first round of surveys was undertaken during the summer of 2016 and summer 2017. We collected basic information on various political, social, and economic characteristics of the village, as well as information about its history. We attempted to ask a similar set of questions as in the data collection described in Appendix G. However, we did not conduct experimental games to measure respect in authority (chief IAT), sharing norms, or altruism. We discuss these differences more in the following sub-section

The surveys comprised surveys of the Village Chief, the Sage, the Secretary, one notable from each clan of the village, and 12 randomly-chosen citizens from the village (4 old men, 4 young men, 2 old women and 2 young women). In choosing the citizens we stratified by gender and age. On average 18 villagers were surveyed per village. In total, we surveyed over 5,400 individuals from these 302 villages.

H.2. Variables Used & Differences with the Gemena Sample

In this sub-section, we describe the similarities and differences between the baseline data from Lowes et al. (2016) and the Gemena data described in Section 4. Many of the questions from the baselline data from Lowes et al. (2016) were taken from our surveys conducted in Gemena. For questions regarding village institutions and public goods, we used the answers provided by the village secretary.[43] For questions regarding social norms, we used the answers provided by 12 randomly-chosen citizens from the villages.[44]

Because the data Lowes et al. (2016) was not collected with the intention to replicate the analysis in this paper (and instead inform an ongoing RCT), an important difference relative to the Gemena sample is that we did not conduct any experimental games in the villages. That means we no longer have measures for the chief IAT score, the dictator game, or the reverse dictator game.

[43] The village secretary is a position in the village usually taken by a younger individual to help the village chief and notables, and to keep records. Focus groups we conducted in villages suggested that the secretaries tended to be the best informed (and less biased) individuals for these matters.

[44] We do this to avoid differences in answers by the composition/size of the local authorities. In particular, villages have differing numbers of notables (due to different numbers of clans). For example, some villages have two notables while others have over twelve. Therefore, to keep the sample comparable in terms of size and composition by village, we use the answers from the 12 randomly-selected individuals.

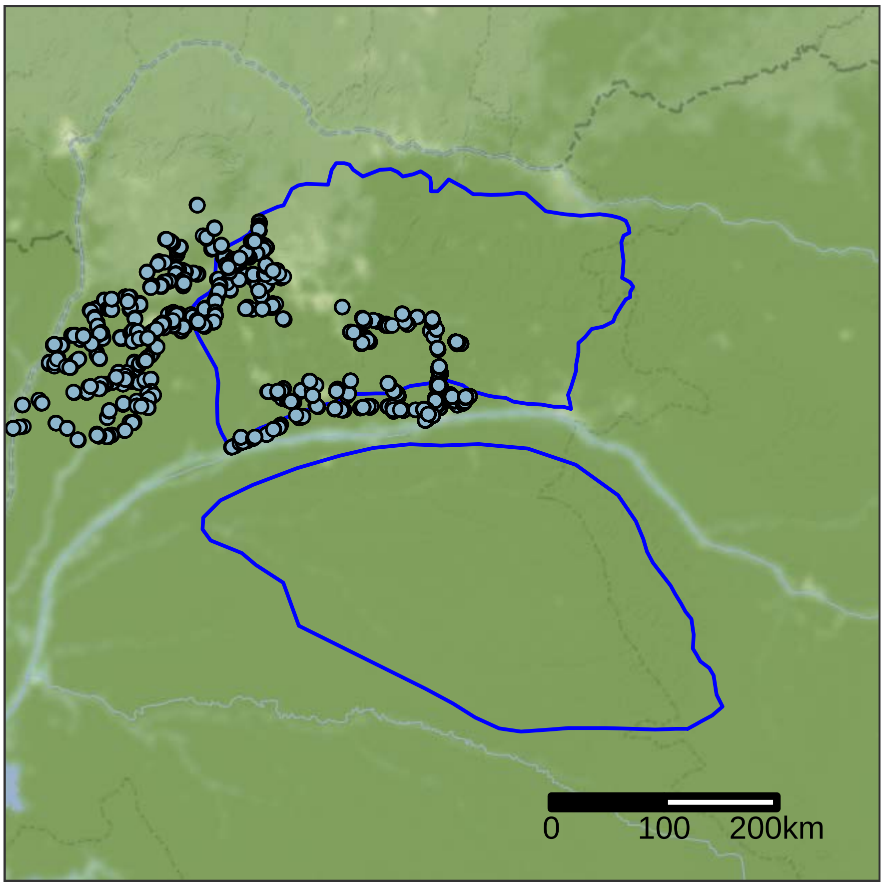


Figure H1: Villages in the Lowes, et al, 2016 Baseline Data

Below, we describe the questions and indexes definitions from the Lowes et al. (2016) sample. For each question/index, if there are slight differences in the definitions (or response options), we highlight these differences compared to what we used in Gemena survey:

- *Chief Public Good Index* is a summary index for the following questions (with the number of components for each question in brackets): Is the chief in your village of origin responsible for providing (1) road maintenance, (2) new roads, (3) school maintenance, (4) land allocation, (5) protection of property rights, (6) tax collection, (7) jobs, (8) conflict arbitration (adultery, theft, or witchcraft), and (9) road brushing; all questions answered as a 0 to 2 categorical variable where 0 is No, 1 is Partially, and 2 is Yes.

  These questions and associated response options matched the Gemena survey. The one difference is with (8), where we broke up the "conflict arbitration" task to be three separate types of conflict arbitration (adultery, theft or witchcraft conflict). Therefore, the index from the Lowes et al. (2016) sample includes three questions regarding conflict arbitration.

- *Chief Elected* is an indicator variable equal to 1 if the village chief of a village is selected by elections according to the secretary.

  This question and response option matches the Gemena survey.

- *Village Public Goods Index* is a summary index for the following questions asked to the village secretary (with the number of components for each question in brackets): (1) What material is the road in your village of origin made of? [2: 0=Sand, 1=Gravel or Pavement] (2) Is your village of origin on a main road? 3) Does your village of origin have a primary school? [2] (4) Does your village of origin have a secondary school? [2] (5) Does your village of origin have a Health Dispensary? [2] (6) Does your village of origin have a Health Center? [2] (7) Does the water in your village of origin come from a well? [2: 0=Spring water, 1=Well].

  This question and response options match the Gemena survey. The index is expanded however to include questions on having a primary school and a health dispensary. The results are very similar when we exclude these questions.[45] Note that the question regarding road material was answered by our enumerators and not the village secretary, as we wanted consistent answers for this question in particular (to understand how the next phase of the RCT could occur).

- *Village Subjective Ratings Index* is a summary index for the following questions asked to the village secretary (with the number of components for each question in brackets): (1) How would you rate the condition of the primary school in your village? [5] (2) How would you rate the condition of the secondary school in your village? [5] (3) How would you rate the quality of the health center in your village relative to other roads in the area? [5] (4) Relative to other villages in the area you have visited, how would your rate your village of origin overall? [5]

  These question, response options, and index construction matches the Gemena survey.

- *Trust in Chief Index* is a summary index for the following questions to villagers (with the number of components for each question in brackets): (1) How much do you trust your village chief? [4], (2) How much does the village trust your village chief? [4].

  This index is different as it only includes two questions on trust in the chief, and no longer includes questions we deemed too sensitive in the field (e.g. would you keep your village

[45] we had effectively zero variation in the Gemena sample for this question, but there was more variation for some of the very rural villages we visited, so we decided to include. The village surveys also improved the health center question to try to break up "health center" into more components relevant in the field.

chief, how would you rate your chief, how much confidence do you have in your chief) and no longer included questions regarding trust or confidence in the sub-tribe chief (as this was not of interest in the baseline survey and also potentially sensitive). The index also includes a new question on how much trust *the village* has in the chief, to try to avoid the question being too sensitive for an individual respondent.

- *Satisfaction w/ Chief Index* is a summary index for the following questions to villagers : (1) How satisfied are you with your village of origin chief? [4], (2) How satisfied is the village with your village chief? [2].

  This index is new and added to the table to proxy for the IAT scores (which we did not conduct in the field). It only includes two questions on satisfaction in the chief (note that the term satisfaction was deemed less sensitive in the field than questions on confidence and rating). Similar to the *Trust in Chief Index*, the index also includes a question on how satisfied *the village* is with the chief, to try to avoid the question being too sensitive for an individual respondent.

- *Trust Index* is a summary index for the following questions asked to the randomly selected villagers: How much do you trust (1) people from your village of origin, (2) people of another tribe, (3) people of your own tribe, (4) your family, (5) your neighbors, (6) people of another nationality, and (7) people of your sub-tribe; all questions answered on a 0 to 4 scale where 0 is Not at All and 4 is Completely.

  These questions, response options, and index construction matches the Gemena sample. The one minor difference was that we replaced the term "sub-tribe" with "clan" to more accurately capture the relevant social structure in the villages (as the term "sub-tribe" can refer either to your clan in your village, or your clan across villages).

- *Closeness to Others Index* is a summary index for the following questions: (1) How close to you feel to people from your village of origin?, (2) How close do you feel to people of your age set from your origin village?; questions answered in a scale from 0 (Not Close at All) to 5 (Very Close).

  These two questions and response options match the Gemena sample. However, in the villages, we no longer asked individuals about closeness with people in Gemena, people in your age set in Gemena, and people in your ethnic group (as villages are often homogenous, so it would proxy for closeness with other villagers).

- *Survey Measures of Sharing Norms: Respondent* is a summary index for the following questions: (1) If you get money from luck you should share it, (2) If you earn money from hard work you should share it, (3) If someone else earns money from luck they should share it, (4) If someone else earns money from hard work they should share it; all questions answered in a scale from 1 (Strongly Disagree) to 5 (Strongly Agree).

  These four questions, the response options, and the index construction match the Gemena sample.

- *Survey Measures of Sharing Norms: Village* measures a respondent's answer to "How often do people in this village share with others", answered in a scale from 1 (Never) to 5 (Always). We did not conduct experimental games in this baseline data collection.

  This is a question that we include in the subsequent analysis to proxy for the *Sharing Norms – Village of Origin* index from the Gemena sample. We asked individuals one question on how often *villagers* (rather than the respondent themselves) shared with others to attempt to capture general village-level sharing norms to proxy for *Sharing Norms – Village of Origin*.

H.3. Results

We reproduce the results from Table 3 and Table 4 as closely as possible and present them in Table H1 and H2, respectively.

We find that the results are generally similar to the main results from Section 4: villages inside the former concession have less accountable chiefs, but respondents have higher sharing norms. This provides an additional robustness test to our results and assuages concerns that selective migration may be driving the results in Section 4.

Table H1: Rubber Concessions and Village Institutions

	Panel A: Village Development			
	Village Public Goods Index		*Village Subjective Ratings Index*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.148^{**}	-0.069	-0.671^{**}	-0.083
	(0.070)	(0.053)	(0.297)	(0.200)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	92	135	81	159
Clusters	92	135	81	159
Bandwidth	30.16	100.00	15.08	100.00
Mean Dep. Var.	0.057	0.013	0.177	0.135
SD Dep. Var.	0.197	0.241	0.702	0.682
	Panel B: Chief Quality and Accountability			
	Chief Elected		*Chief Public Good Index*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.361^{**}	-0.082	-0.154	-0.218^{***}
	(0.152)	(0.109)	(0.121)	(0.084)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	165	260	161	231
Clusters	165	260	161	231
Bandwidth	31.70	100.00	33.87	100.00
Mean Dep. Var.	0.519	0.527	-0.062	-0.035
SD Dep. Var.	0.502	0.500	0.362	0.378
	Panel C: Respect for Authority			
	Trust in Chief Index		*Satisfaction w/ Chief Index*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.008	0.320^{***}	0.040	-0.102
	(0.201)	(0.099)	(0.218)	(0.100)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	1,617	3,279	2,383	4,818
Clusters	100	277	98	277
Bandwidth	17.78	100.00	17.32	100.00
Mean Dep. Var.	0.087	0.065	-0.014	-0.012
SD Dep. Var.	0.885	0.936	1.020	1.033

Notes: Standard errors clustered at the village level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. *Village Public Goods Index* is a summary index for the following questions asked to the village secretary (with the number of components for each question in brackets): (1) What material is the road in your village of origin made of? [2: 0=Sand, 1=Gravel or Pavement] (2) Is your village of origin on a main road? 3) Does your village of origin have a primary school? [2] (4) Does your village of origin have a secondary school? [2] (5) Does your village of origin have a Health Dispensary? [2] (6) Does your village of origin have a Health Center? [2] (7) Does the water in your village of origin come from a well? [2: 0=Spring water, 1=Well]. *Village Subjective Ratings Index* is a summary index for the following questions asked to the village secretary: (1) How would you rate the condition of the primary school in your village? [5] (2) How would you rate the condition of the secondary school in your village? [5] (3) How would you rate the quality of the health center in your village relative to other roads in the area? [5] (4) Relative to other villages in the area you have visited, how would your rate your village of origin overall? [5] *Chief Public Good Index* is a summary index for the following questions: Is the chief in your village of origin responsible for providing (1) road maintenance, (2) new roads, (3) school maintenance, (4) land allocation, (5) protection of property rights, (6) tax collection, (7) jobs, (8) conflict arbitration, and (9) road brushing; all questions answered as a 0 to 2 categorical variable where 0 is No, 1 is Partially, and 2 is Yes. *Chief Elected* is an indicator variable equal to 1 if the village chief of a village is selected by elections according to the secretary. *Trust in Chief Index* is a summary index for the following questions to villagers: (1) How much do you trust your village chief? [4], (2) How much does the village trust your village chief? [4]. *Satisfaction w/ Chief Index* is a summary index for the following questions to villagers : (1) How satisfied are you with your village of origin chief? [4], (2) How satisfied is the village with your village chief? [2]. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table H2: Survey Measures of Trust and Sharing Beliefs

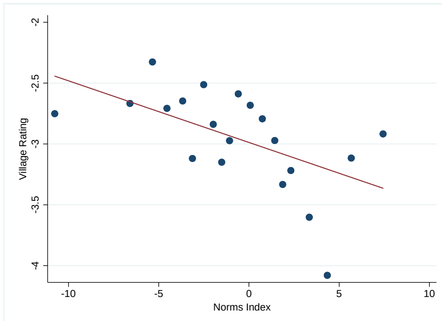
	Panel A: Trust and Closeness			
	Trust Index		*Closeness Index*	
	(1)	(2)	(3)	(4)
Inside Concession	0.021	0.059***	0.107	0.076
	(0.030)	(0.021)	(0.125)	(0.059)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	2,405	4,860	2,217	4,799
Clusters	89	277	84	277
Bandwidth	15.22	100.00	14.69	100.00
Mean Dep. Var.	0.044	0.018	-0.016	0.018
SD Dep. Var.	0.376	0.475	0.934	0.987
	Panel B: Survey Measures of Sharing Norms			
	Respondent		*Village*	
	(1)	(2)	(3)	(4)
Inside Concession	0.235	0.211**	0.282*	0.261***
	(0.218)	(0.106)	(0.156)	(0.094)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	1,795	3,950	2,779	4,849
Clusters	84	277	123	277
Bandwidth	14.76	100.00	22.24	100.00
Mean Dep. Var.	0.002	-0.011	0.020	0.010
SD Dep. Var.	0.880	0.854	1.088	1.059

Notes: Standard errors clustered at the village level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions include nearest concession fixed effects. *Trust Index* is a summary index for the following questions: How much do you trust (1) people from your village of origin, (2) people of another tribe, (3) people of your own tribe, (4) your family, (5) your neighbors, (6) people of another nationality, and (7) people of your sub-tribe; all questions answered on a 0 to 4 scale where 0 is Not at All and 4 is Completely. *Closeness to Others Index* is a summary index for the following questions: (1) How close to you feel to people from your village of origin?, (2) How close do you feel to people of your age set from your origin village?; questions answered in a scale from 0 (Not Close at All) to 5 (Very Close). *Survey Measures of Sharing Norms: Respondent* is a summary index for the following questions: (1) If you get money from luck you should share it, (2) If you earn money from hard work you should share it, (3) If someone else earns money from luck they should share it, (4) If someone else earns money from hard work they should share it; all questions answered in a scale from 1 (Strongly Disagree) to 5 (Strongly Agree). *Survey Measures of Sharing Norms: Village* measures a respondent's answer to "How often do people in this village share with others", answered in a scale from 1 (Never) to 5 (Always). We did not conduct experimental games in this baseline data collection. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

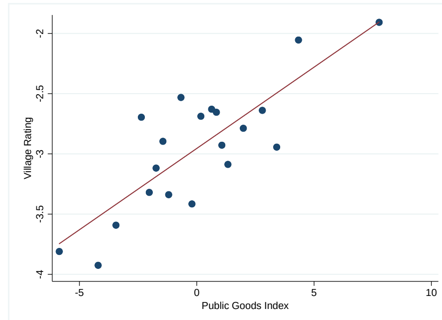
Appendix I. Additional Results

I.1. Relationship Between Development, Institutions, and Culture in the Fieldwork Data

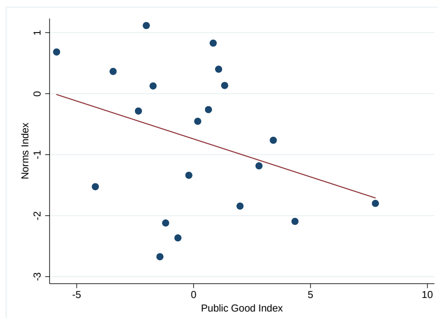
Figure I1: Correlation Plots: Development, Institutions, and Culture



(a) Village Rating and Strength of Pro-Social Norms



(b) Village Rating and Public Goods Provision by Chief



(c) Strength of Pro-Social Norms and Public Goods Provision by Chief

Notes: *Village Rating* measures how a respondent answers the following question "Relative to other villages in the area you have visited, how would you rate your village of origin overall?" on a scale from 1 to 5, where 1=a lot worse than other communities and 5=much better off than other communities. *Public Good Index* is the sum of the following public good questions (with the number of components for each question in brackets), where each question is standardized before being included in the summation: (1) What material is the road in your village of origin made of? [2: 0=Sand, 1=Gravel or Pavement] (2) Is your village of origin on a main road? (3) Does your village of origin have mobile coverage? [5] (4) Does your village of origin have a secondary school? [2] (5) Does your village of origin have a Health Dispensary? [2] (6) Does your village of origin have a Hospital? [2] (7) Does the water in your village of origin come from a well? [2: 0=Spring water, 1=Well]. More positive values of the index indicates higher amounts of public goods in a village of origin. *Norms Index* is the sum of the following sharing norms questions (with the number of components for each question in brackets), where each question is standardized before being included in the summation: (1) If you get money from luck you should share it [5], (2) If you earn money from hard work you should share it [5], (3) If someone else earns money from luck they should share it [5], (4) If someone else earns money from hard work they should share it [5], (5) How close to you feel to people from your village of origin? [5], (6) How close to do you feel to people of Gemena? [5], (7) How close do you feel to people of your own tribe? [5], (8) How close do you feel to people of your age set from your origin village? [5], and (9) How close do you feel to people of your age set in Gemena? [5]. More positive values of the index indicate stronger pro-social norms for an individual. The figures are binscatters. All binscatters include district fixed effects, a linear polynomial in latitude and longitude, and baseline controls for individuals (age, age squared and a male indicator).

I.2. Establishing a First-Stage Discontinuity - Commercial Posts in 1897 and 1905

While it is not required to show a first stage for an RD analysis, we can examine whether the probability of having a "commercial post" is higher within the concession boundaries. A commercial post corresponds to places where rubber is collected and traded. In Figure I2 we present digitized maps of commercial posts and show that the former concession areas are much more likely to have had commercial posts. If there were no "first-stage" in the sense that the concession areas were not more likely to be exposed to the rubber extraction, then it is unlikely we would find effects of being inside a former concession. Additionally, if the RD were "fuzzy" such that the concession boundaries were not perfectly respected, this would bias our coefficients toward zero. Ideally, we would have detailed granular data of exposure to violence or rubber production. We have been unable to find such data, though in Section I.4 we examine the correlation between post level rubber production for a six month period of 1904 for which we were able to find data and wealth today.

Table I1 presents the regression discontinuity estimates for indicator variable for having a "commercial post" is higher within the concession boundaries. The indicator is constructed at the grid cell level as in Table 1. A commercial post corresponds to places where rubber is collected and traded. Columns (1)-(3) present results using Figure I2 on commercial post locations in 1897 from Rouck (1945), while Columns (4)-(6) present results using Figure I3 on commercial post locations in 1905 from Goffart (1908). The results demonstrate that the former concession areas are much more likely to have had commercial posts compared to areas just outside the former concession boundaries.

Table I1: Establishing a First-Stage Discontinuity - Differences in "Commercial" Post Presence

	Commercial Post in 1897		*Commercial Post in 1905*	
	(1)	(2)	(3)	(4)
Inside Concession	0.012	0.030	0.037*	0.071***
	(0.030)	(0.027)	(0.020)	(0.021)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	787	850	783	850
Clusters	26	29	26	29
Bandwidth	68.93	100.00	64.89	100.00
Mean Dep. Var.	0.023	0.024	0.057	0.047
SD Dep. Var.	0.151	0.152	0.232	0.212

Notes: We present standard errors clustered at the territory level in ( ). All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions include nearest concession fixed effects. *Commercial Posts in 1897* is an indicator variable equal to one if a 20 by 20 km grid cell had at least one commercial posts in 1897 in Rouck (1945). *Commercial Posts in 1905* is an indicator variable equal to one if a 20 by 20 km grid cell had at least one commercial posts in 1905 in Goffart (1908). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

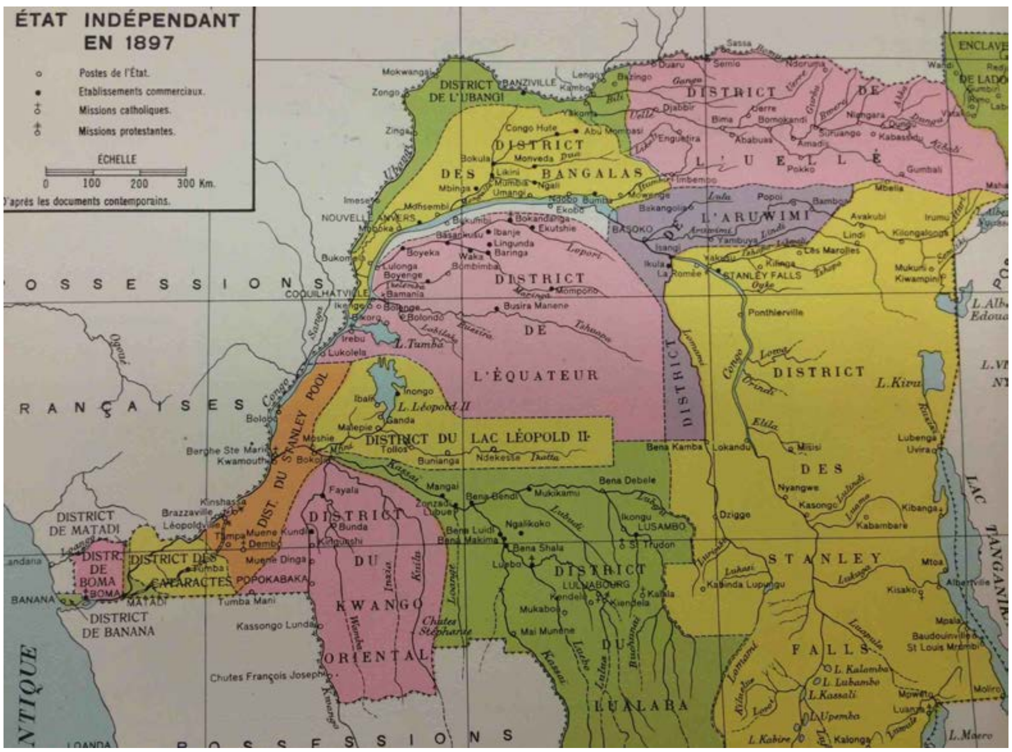


Figure I2: Map with 1897 Commercial Post Locations

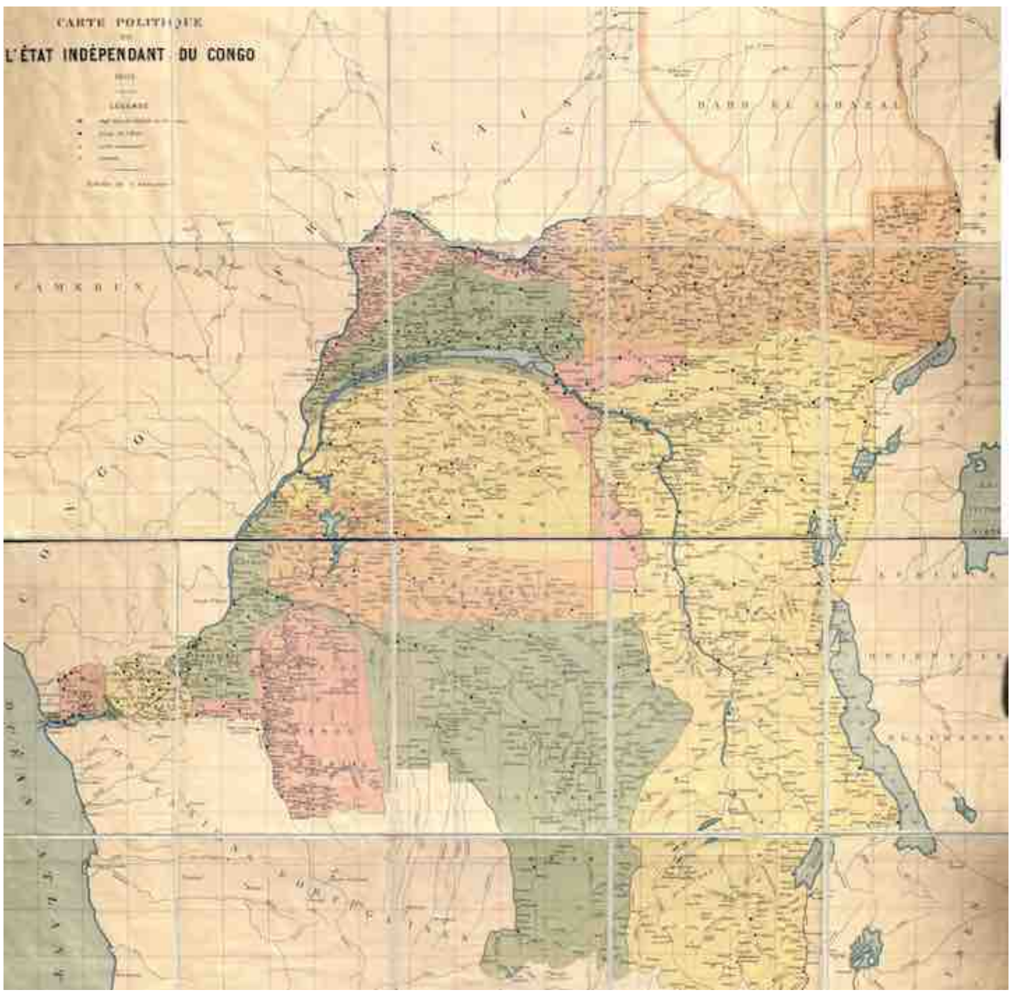


Figure I3: Map with 1905 Commercial Post Locations

I.3. Wealth Inequality at the DHS Cluster Level

Table I2 presents the regression discontinuity estimates for wealth inequality for DHS clusters. Columns (1)-(3) present results using the standard deviation in the DHS wealth factor score for each DHS cluster as the dependent variable, while Columns (4)-(6) present results using the inter-quartile range in the DHS wealth factor score for each DHS cluster as the dependent variable. The results demonstrate that villages inside the former concessions have lower levels of wealth inequality compared to villages just outside the former concession boundaries.

Table I2: Rubber Concessions and Wealth Inequality

	St. Dev. of Wealth Score		*IQR of Wealth Score*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.589^{**}	-0.670^{**}	-0.811^{*}	-0.910^{**}
	(0.291)	(0.284)	(0.456)	(0.439)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	133	155	133	155
Clusters	29	34	29	34
Bandwidth	115.87	200.00	111.93	200.00
Mean Dep. Var.	2.419	2.222	1.968	1.746
SD Dep. Var.	0.920	0.961	1.253	1.351

Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

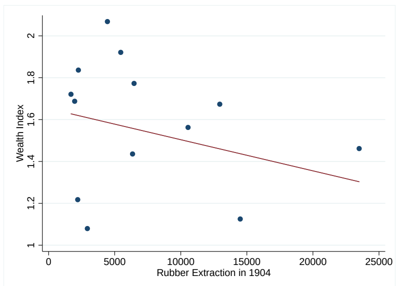
I.4. Analysis Using Historical Post Level Data

As a complement to the RD analysis, we analyze post-level rubber production data from 1904 for ABIR. We combined data on rubber production from the Belgian Foreign Public Service Foreign Affairs archives with data from the De Ryck Collection, a collection of Congo colonial manuscripts at the University of Wisconsin library. We were able to compile data on rubber production for 19 posts within the ABIR concession between July and December 1904 (see Figure 2b for map of post locations) (de Ryck, 1885-1954). We use these measures of production as a proxy for intensity of exposure to extractive institutions. We match DHS clusters to rubber posts within 50 kilometers. Even though we are limited by the small number of DHS clusters near former rubber posts, we find that individuals within DHS clusters close to posts that produced more rubber during these 6 months of 1904 are less wealthy today, as seen in Figure I4.[46] To the extent that rubber production captures the intensive margin of exposure to colonial extraction, these results suggest that greater exposure indeed leads to worse development outcomes, with the caveat that these are merely correlations. As an alternative measure of intensity of exposure, we use year of post establishment. Posts within ABIR were established between 1892 and 1903. We find that individuals close to posts that were operating for more years are also worse off. These results are presented in Table I3 and in Figure I4, and they suggest that some of the heterogeneity in

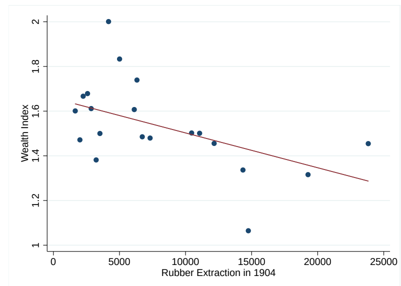
[46] Note that once controls are added in Figure I4, there is more variation within a bin, which is why there appear to be more observations in the binscatters.

development outcomes near the former concessions can be explained by the intensity of extraction during the Congo Free State period.

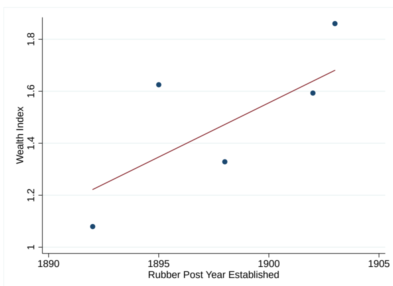
Figure I4: Analysis Using Historical Post Level Data



(a) Wealth and rubber production in 1904



(b) Wealth and rubber production in 1904 (controls)



(c) Wealth and year of post establishment

(d) Wealth and year of post establishment (controls)

Notes: We use data on the amount of rubber produced in 19 posts within the ABIR concession between July and December 1904 and match posts to DHS clusters within 50 km of the former posts. Figures (b) and (d) include controls for age, age squared, gender, survey year, latitude and longitude. Rubber Production in 1904 is measured in tons.

I.5. On a Convergence Path?

It is important to understand whether areas inside the former rubber concessions are actually on a path to convergence with areas outside the former concessions but have simply not caught up yet. We test for convergence in our setting by examining whether younger cohorts inside the former concessions are "catching up" to similar cohorts outside the former concessions in terms of the development outcomes examined in Table 2. Effectively, we are examining how the effect of being inside a concession varies over time.

To do this, we compare cohorts inside and outside the concessions born within five years of each other by estimating a regression that includes fixed effects for each 5-year cohort along with the interactions between the $RubberConcession$ indicator and cohort fixed effects.[47] Figure I5 plots the estimated cohort coefficients for years of education, literacy, height-to-age and wealth.

[47] Formally, we estimate the following specification for DHS clusters within 200 kms of the concession borders:

$$y_{i,v} = \gamma RubberConcession_{i,v} + \alpha_y C_y + \gamma_y C_y \times RubberConcession_{i,v} + \mathbf{X}_i\beta + \phi + \varepsilon_{i,v} \qquad \text{(A1)}$$

where C_y are 5-year cohort fixed effects and the other variables are defined as in equation (1). Note that we are not estimating a distinct RD polynomial for each cohort as that would be too demanding of the data given our sample size.

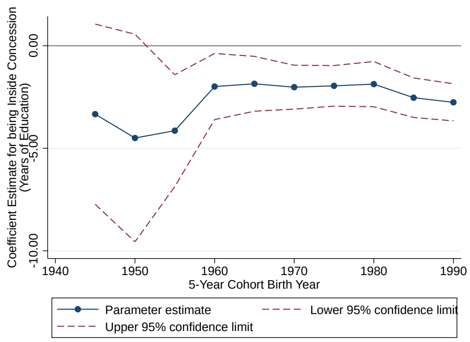
Table I3: Post Level Rubber Production in 1904, Year of Post Establishment, and Development Outcomes

	Wealth Index		*Wealth Index*	
	(1)	(2)	(3)	(4)
Rubber Production in 1904	−0.025***	−0.062***		
	(0.008)	(0.009)		
Year Post was Established			0.038*	0.045***
			(0.021)	(0.014)
Observations	704	704	704	704
Clusters	16	16	16	16
Controls	N	Y	N	Y
Mean Dep. Var.	1.534	1.534	1.534	1.534
SD Dep. Var.	0.789	0.789	0.789	0.789
	Log(Wealth Score)		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Rubber Production in 1904	−0.014**	−0.033***		
	(0.005)	(0.003)		
Year Post was Established			0.001	0.021***
			(0.015)	(0.006)
Observations	704	704	704	704
Clusters	16	16	16	16
Controls	N	Y	N	Y
Mean Dep. Var.	10.516	10.516	10.516	10.516
SD Dep. Var.	0.392	0.392	0.392	0.392
	Years of Education		*Years of Education*	
	(1)	(2)	(3)	(4)
Rubber Production in 1904	−0.017	−0.105		
	(0.036)	(0.083)		
Year Post was Established			−0.066	0.080
			(0.084)	(0.072)
Observations	703	703	703	703
Clusters	16	16	16	16
Controls	N	Y	N	Y
Mean Dep. Var.	5.579	5.579	5.579	5.579
SD Dep. Var.	3.683	3.683	3.683	3.683
	Literacy		*Literacy*	
	(1)	(2)	(3)	(4)
Rubber Production in 1904	−0.002	−0.006		
	(0.003)	(0.009)		
Year Post was Established			−0.013*	0.017*
			(0.007)	(0.009)
Observations	700	700	700	700
Clusters	16	16	16	16
Controls	N	Y	N	Y
Mean Dep. Var.	0.494	0.494	0.494	0.494
SD Dep. Var.	0.500	0.500	0.500	0.500
	Respondent Ht./Age		*Respondent Ht./Age*	
	(1)	(2)	(3)	(4)
Rubber Production in 1904	−0.010**	−0.011		
	(0.004)	(0.010)		
Year Post was Established			0.012	0.032***
			(0.009)	(0.005)
Observations	265	265	265	265
Clusters	16	16	16	16
Controls	N	Y	N	Y
Mean Dep. Var.	0.269	0.269	0.269	0.269
SD Dep. Var.	0.279	0.279	0.279	0.279

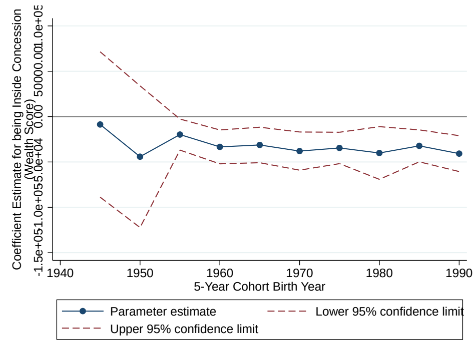
Notes: Rubber Production in 1904 measures production in tons for the last six months of 1904 for ABIR posts. We match DHS clusters to the closest ABIR post and limit the sample to clusters within 50 kms of the former ABIR posts. We cluster standard errors at the DHS cluster level. In columns (2) and (4) we include district fixed effects and control for age, age squared, gender, survey year, latitude, longitude, malaria and tsetse suitability, and distance to concession border. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. Standard errors clustered at the DHS cluster level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We see no evidence for convergence across cohorts: the estimated coefficients for each cohort are similar, stable and do not get closer to zero for younger cohorts.[48]

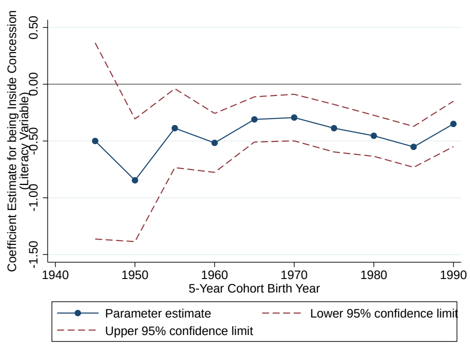
Figure I5: Estimated Cohort Coefficients for Individuals within 200 kms of the Rubber Concessions



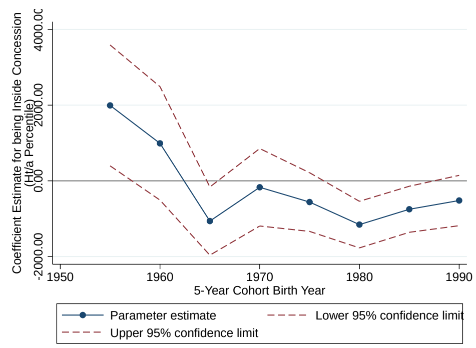
(a) Years of Education



(b) Wealth Factor Score



(c) Literacy



(d) Height/Age Percentile

Notes: These figures plot the estimated coefficient for each 5 year cohort indicator interacted with the indicator for being inside a former concession area for observations within 200 kms of the concession borders. The regression also includes cohort fixed effects. Standard errors are clustered at the DHS cluster level. The figures also plot 95% confidence intervals for the coefficients. All outcome variables are from the DHS 2007 and 2014 surveys. The regressions all have 1496 observations. Wealth Factor Score is an index generated by the DHS using principle component on asset ownership. Literacy is a 0 to 2 categorical variable where 0 is cannot read at all and 2 is able to read a whole sentence. Ht/Age Percentile divides each respondent's height by her age and finds her percentile in the entire sample and normalizes this percentile to be within 0 and 10000.

I.6. Market Access and Public Good Provision Since Independence

An additional potential channel of interest is differences in investment in market access and public goods post-independence. After independence, the central government suffered from political instability (Van Reybrouck, 2014); thus, much of the subsequent maintenance of roads and the provision of public goods was not provided by the central government. If investment levels in roads are significantly lower within the former concessions, this would suggest local failure of collective action as a potential mechanism for persistence. This lack of local maintenance could arise for many reasons. For example, local governments may not have the capacity to invest in public goods or in infrastructure maintenance in former concession areas, or individuals in

[48] The one exception are the estimates for the health outcome, where older cohorts appear to have slightly higher height-to-age percentiles inside the former concessions. This could potentially be explained by selective survival – e.g. for the older individuals we only observe those healthy enough to survive inside the former concessions.

the former concessions are less trusting of outsiders and therefore choose not to invest in public goods and infrastructure.

Using data from the *Referentiel Geographique Commun* on current road networks and bridges in DRC today, Panel C of Table E1 examines whether areas inside the former concession have lower market access today (Appendix B presents maps of the road networks and bridges). We find that areas inside the former rubber concessions have fewer roads and bridges today relative to areas outside of the former concessions. The results in Panel C of Table E1 combined with the results in Panel B – in which we find no evidence of differences in road network investments by the Belgian colonial government – suggest that differences in public good and infrastructure provision *since independence* are a plausible channel of persistence in this setting. Because the Belgians did not invest differentially in road infrastructure inside and outside the concessions, road network density was similar at independence. Yet, today we find that road networks are less dense inside the former concessions. Given that there have not been any substantial investments in new roads in the area since independence, these results suggest that the observed differences in road network density today are driven by a failure by local chiefs and their constituents to maintain roads that existed at the time of independence.

I.7. Population Density

This section examines differences in population density. The results below show that the former concession areas have lower population density. A Maluthusian model would predict higher income per capita inside the concession areas and a simple Solow model would predict convergence. Empirical evidence from other settings that experienced intense violence – such as Rwanda in the 1990s (Rogall and Yanagizawa-Drott, 2014) and the 1609 Spanish expulsion of the Moriscos (Chaney and Hornbeck, 2016) – suggest that the concessions would have converged to a similar level of development by now. This suggests that differences in population density directly to the violence are unlikely to explain the results.

We use data from Landscan 2007 to get a measure of population density as an additional indicator of development. Landscan 2007 data uses detailed satellite imagery to construct measures of population density at a resolution of approximately 1 km by 1 km for the entire world. Figure I6 is a map of population data around the rubber concession areas. Table I4 presents our results from estimating specification (1) on 20 km by 20 km grid cells constructed with GIS.[49] We find that areas inside the former rubber concessions are less populated today than areas outside. Areas inside the former concession borders have approximately three fewer people per 1 km by 1 km grid cell on average (this corresponds to about 25 % fewer people per square kilometer). Thus, even though the rubber extraction and violence occurred over 100 years before the population density measure, the areas inside the former rubber concessions continue to be less populated today than areas outside the former concessions.

[49] We use 20 km by 20 km grid cells to match the analysis by Dell (2010). To conserve space, we present results using only a cubic polynomial in distance to the concession border; the results are very similar for the cubic polynomial in latitude and longitude and are available upon request. For the Landscan analysis, we drop outlier grid cells before running our analysis; specifically, we drop any observation above the 99th percentile.

Table I4: Rubber Concessions and Population Density

	Log(Population Density) *Landscan*		*Log(Population Density)* *African Pop. Database*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.589^{**}	-0.670^{**}	-0.811^{*}	-0.910^{**}
	(0.291)	(0.284)	(0.456)	(0.439)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	133	155	133	155
Clusters	29	34	29	34
Bandwidth	115.87	200.00	111.93	200.00
Mean Dep. Var.	2.419	2.222	1.968	1.746
SD Dep. Var.	0.920	0.961	1.253	1.351

Notes: We present standard errors clustered at the territory level in ( ). All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions include nearest concession fixed effects and control for elevation, rainfall, land suitability, ruggedness, malaria suitability, and tsetse suitability. Data in columns 1 and 2 is from Landscan for 2007 and in columns 3 and 4 from the African Population database for 2000. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure I6: Population Density Measure from Landscan 2007

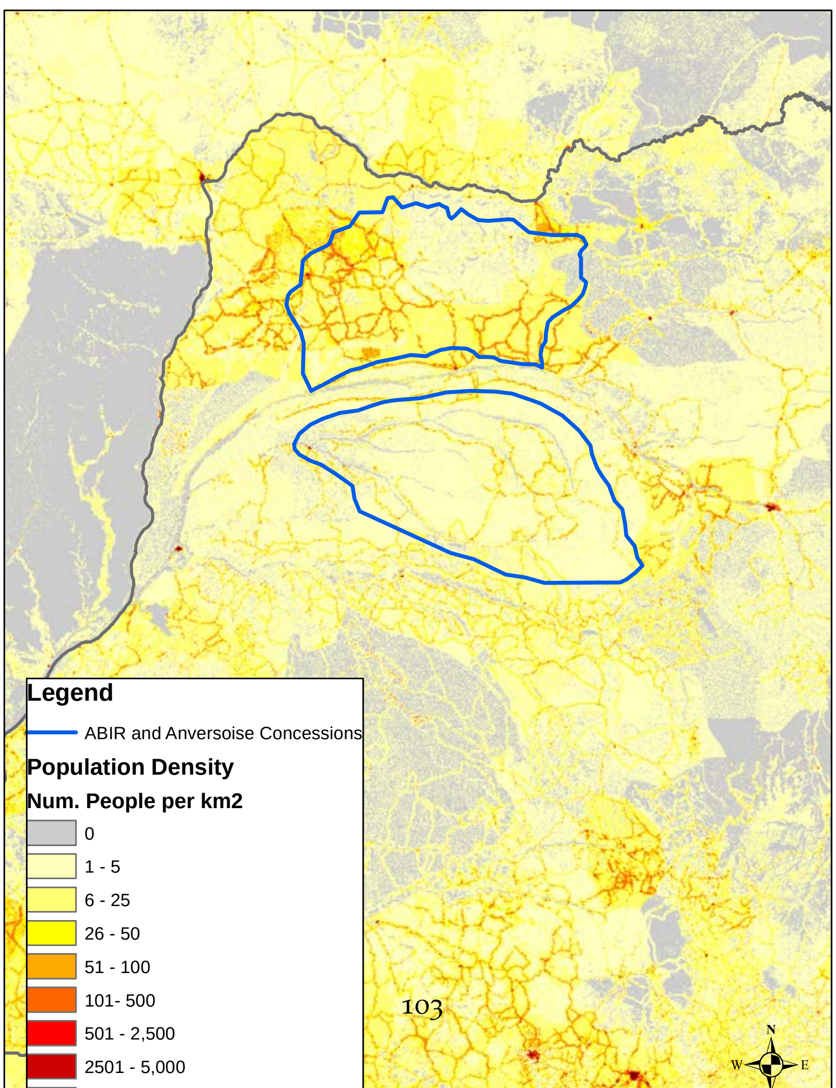
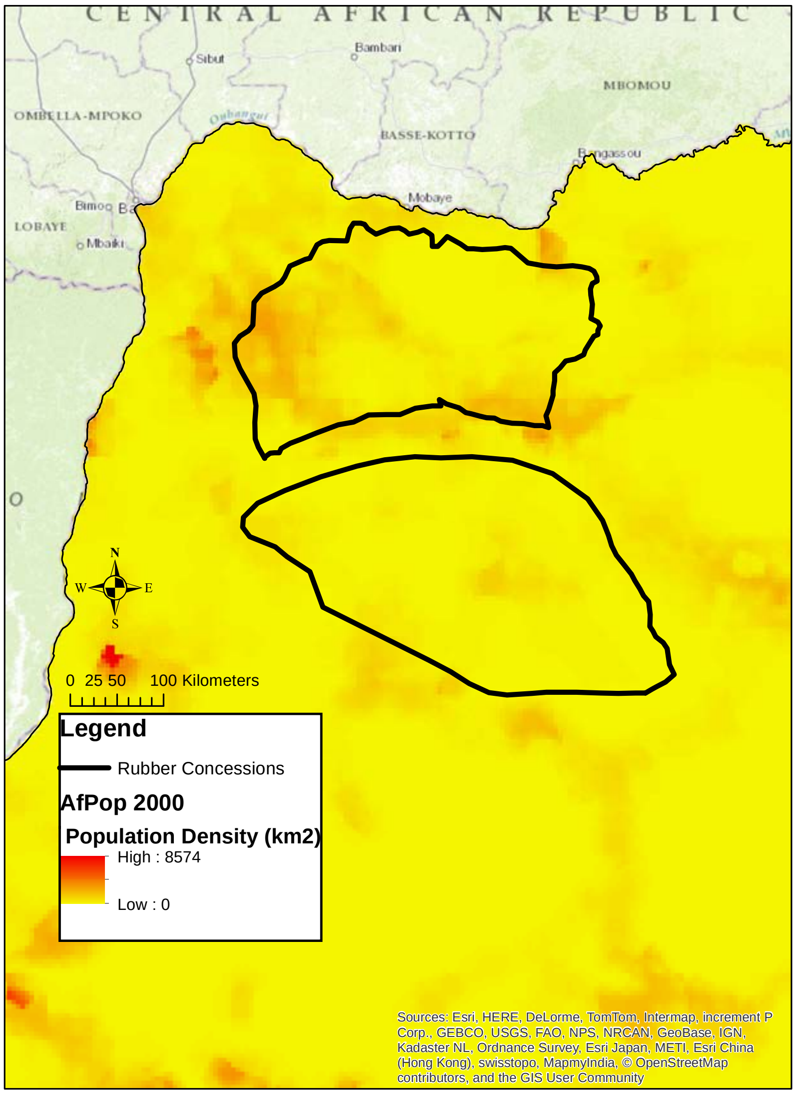


Figure I7: Population Density Measure from the African Population Database



I.8. *Nightlights*

To explore the impacts of the rubber concessions on development outcomes available across the entire border, we use satellite data on nightlight intensity from NOAA (in particular, the VIIRS NDB Nighttime Lights for 2016) as a measure of development (Henderson, Storeygard and Weil, 2012). Table I5 presents estimates on differences in nightlight intensity at the former rubber concession border. We find evidence that the former rubber concession areas have lower levels of nightlight intensity today. These results are consistent with the DHS results and show that the former rubber concessions areas continue to have lower development levels today.

Table I5: Rubber Concessions and Nightlight Intensity

	Nightlights		*Log(Nightlights)*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.852	−1.134**	−0.127*	−0.158**
	(0.577)	(0.495)	(0.071)	(0.067)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	136	158	136	158
Clusters	29	34	29	34
Bandwidth	99.59	200.00	108.48	200.00
Mean Dep. Var.	1.632	1.542	1.932	1.900
SD Dep. Var.	2.160	2.909	0.274	0.342

Notes: We present standard errors clustered at the territory level in ( ). All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions include nearest concession fixed effects and control for elevation, rainfall, land suitability, ruggedness, malaria suitability, and tsetse suitability. Data is from the VIIRS NDB Nighttime Lights data from the NOAA. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I.9. *Violence and Conflict*

The intense violence experienced during the rubber extraction period might have changed social norms for violence by making people more prone to resort to violence for conflict resolution. These differences could have led areas inside the former concessions to experience more violence and conflict, and this could have hindered economic development.

We test for differences in violence using data from PRIO that documents the location and intensity of major conflict events in the DRC since 1989. Table I6 presents estimates on differences in violent conflict. The dependent variable is total amount of conflict in 20 km by 20 km grid cells. We find some weak evidence that these areas experience less conflict. However, note that this is not the ideal test of differences in social norms for violence, since most of the PRIO data for Congo captures large-scale conflicts that were a consequence of the Congo Wars. Thus, we cannot conclude that differences in conflict and violence explain the main results.

Table I6: Rubber Concessions and Conflict

	Conflict Event		*Num. Casualties*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.133	−0.110	−16.752	−23.044
	(0.091)	(0.101)	(28.387)	(20.020)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	118	158	108	158
Clusters	29	34	25	34
Bandwidth	96.24	200.00	55.29	200.00
Mean Dep. Var.	0.071	0.088	2.804	1.885
SD Dep. Var.	0.431	0.667	48.423	31.821

Notes: We present standard errors clustered at the territory level in ( ). All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions include nearest concession fixed effects and control for elevation, rainfall, land suitability, ruggedness, malaria suitability, and tsetse suitability. Conflict data is from the Uppsala Conflict Data Program. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix J. All CFS Concession Results

This section presents the regression discontinuity results using all the CFS concession borders presented in Figure 1 as the discontinuity and the DHS data as the outcome data. Section J.1 presents the results pooling all the concessions together, and Section J.2 presents the results excluding ABIR and Anversoise.

J.1. All Concessions

Table J1: All Concession and Economic Development

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	-2.624^{***}	-2.031^{***}	-0.268^{***}	-0.212^{***}
	(0.711)	(0.624)	(0.077)	(0.069)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	13,462	11,561	13,308	11,509
Clusters	175	234	178	234
Bandwidth	33.23	50.00	34.11	50.00
Mean Dep. Var.	6.221	6.211	0.559	0.562
SD Dep. Var.	4.048	4.052	0.497	0.496
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.956^{**}	-0.669^{*}	-0.599^{***}	-0.396^{**}
	(0.394)	(0.344)	(0.220)	(0.190)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	13,388	11,574	13,159	11,574
Clusters	180	234	174	234
Bandwidth	35.09	50.00	32.91	50.00
Mean Dep. Var.	2.776	2.731	11.285	11.239
SD Dep. Var.	1.401	1.370	0.736	0.715
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	-0.068^{*}	-0.065^{*}	-0.030	-0.030
	(0.038)	(0.037)	(0.035)	(0.035)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	4,975	3,884	5,079	3,700
Clusters	225	234	234	234
Bandwidth	47.34	50.00	49.94	50.00
Mean Dep. Var.	0.258	0.257	0.313	0.313
SD Dep. Var.	0.263	0.263	0.376	0.376

Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J.2. Excluding ABIR and Anversoise

Table J2: All Concession and Economic Development
Excluding ABIR and Anversoise

	Panel A: Education			
	Years of Education		*Literacy*	
	(1)	(2)	(3)	(4)
Inside Concession	-3.058^{***}	-2.267^{***}	-0.329^{***}	-0.241^{**}
	(1.093)	(0.872)	(0.118)	(0.095)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	13,462	11,561	13,308	11,509
Clusters	175	234	178	234
Bandwidth	33.23	50.00	34.11	50.00
Mean Dep. Var.	6.221	6.211	0.559	0.562
SD Dep. Var.	4.048	4.052	0.497	0.496
	Panel B: Asset Wealth			
	Wealth Index		*Log(Wealth Score)*	
	(1)	(2)	(3)	(4)
Inside Concession	-1.211^{**}	-0.853^{*}	-0.791^{**}	-0.529^{*}
	(0.565)	(0.455)	(0.342)	(0.271)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	10,490	9,105	10,490	9,105
Clusters	136	185	134	185
Bandwidth	31.24	50.00	30.89	50.00
Mean Dep. Var.	2.929	2.896	11.347	11.319
SD Dep. Var.	1.424	1.390	0.765	0.750
	Panel C: Health			
	Respondent Ht./Age Percentile		*Child Ht./Age Percentile*	
	(1)	(2)	(3)	(4)
Inside Concession	−0.074	−0.050	−0.056	−0.050
	(0.054)	(0.050)	(0.037)	(0.035)
Bandwidth Choice	Optimal	Wide	Optimal	Wide
Observations	3,853	3,051	3,370	2,520
Clusters	161	185	171	185
Bandwidth	41.10	50.00	43.97	50.00
Mean Dep. Var.	0.255	0.255	0.222	0.222
SD Dep. Var.	0.264	0.265	0.300	0.300

Notes: Standard errors clustered at the DHS cluster level. All regressions include a local linear specification estimated separately on each side of the concession boundary and use a triangular kernel. Optimal bandwidths are chosen using the MSE-minimizing procedure suggested by Cattaneo et al. (2020) and are reported in kms. Regressions control for age, age squared, gender, survey year, and nearest concession fixed effects. *Literacy* is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. *Wealth Score* is an index generated by the DHS using principle component of asset ownership. *Wealth Index* is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the *Wealth Score*. *Respondent Ht/Age Percentile* is measured for a subset of female respondents divides each respondent's height by her age and finds her percentile in the entire sample. Similarly, *Child Ht/Age Percentile* was asked to a subset of children and divides each child's height by his or her age and finds his or her percentile in the entire sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix References

Académie Royale des Sciences d'Outre-Mer, *Atlas Général du Congo et du Ruanda-Urundi*, Brussels: Académie Royale des Sciences d'Outre-Mer, 1954.

Alsan, Marcella, "The Effect of the TseTse Fly on African Development," *American Economic Review*, January 2015, *105* (1), 382–410.

Bazzi, Samuel, Arya Gaduh, Alexander D. Rothenberg, and Maisy Wong, "Skill Transferability, Migration, and Development: Evidence from Population Resettlement in Indonesia," *American Economic Review*, 2016, *106* (9), 2658–2698.

Bluemke, Matthias and Malte Friese, "Reliability and Validity of the Single-Target IAT (STIAT): Assessing Automatic Affect Toward Multiple Attitude Objects," *European Journal of Social Psychology*, 2008, *38*, 977–997.

Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik, *A Practical Introduction to Regression Discontinuity Designs: Foundations*, Cambridge, U.K.: Cambridge University Press, 2020.

Chaney, Eric and Richard Hornbeck, "Economic Dynamics in the Malthusian Era: Evidence from the 1609 Spanish Expulsion of the Moriscos," *Economic Journal*, 2016, *126* (594), 1404–1440.

Cherry, Todd L., Peter Frykblom, and Jason F. Shogren, "Hardnose the Dictator," *American Economic Review*, 2002, *94* (4), 1218–1221.

Conley, Timothy, "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, *92*, 1–45.

de Rouck, René, *Atlas Géographique et Historique du Congo Belge et Des Territoires sous Mandat du Ruanda-Urundi* 1945.

de Ryck, Maurice Martin, "Zaire Colonial Documents: De Ryck Collection of Materials on General Administration, Equateur, Kivu, and Ruanda-Urundi," Microfilm Collection 1885-1954. University of Wisconsin.

Dell, Melissa, "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 2010, *78* (6), 1863–1903.

Farh, René and Bernd Irlenbusch, "Fairness as a Constraint on Trust in Reciprocity: Earned Property Rights in a Reciprocal Exchange Experiment," *Economic Letters*, 2000, *66* (3), 275–282.

Goffart, Ferdinand, *Le Congo: Géographie, Physique, Politique et Économique* 1908.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil, "Measuring Economic Growth from Outer Space," *American Economic Review*, April 2012, *102* (2), 994–1028.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis, "Very High Resolution Interpolated Climate Surfaces for Global Land Areas," *International Journal of Climatology*, 2005, *25*, 1965–1978.

Hochschild, Adam, *King Leopold's Ghost: A Story of Greed, Terror, and Heroism in Colonial Africa*, Houghton Mifflin, 1998.

Hoffman, Elizabeth, Kevin McCabe, Kevin Shachat, and Vernon Smith, "Preferences, Property Rights, and Anonymity in Bargaining Games," *Games and Economic Behavior*, 1994, *7* (3), 346–380.

Jakiela, Pamela, "Social Preferences and Fairness Norms as Informal Institutions: Experimental Evidence," *American Economic Review: Papers & Proceedings*, 2011, *101* (3), 509–513.

Kiszewski, Anthony, Andrew Mellinger, Andrew Spielman, Pia Malaney, Sonia Ehrlich Sachs, and Jeffrey Sachs, "A Global Index Representing the Stability of Malaria Transmission," *The American Journal of Tropical Medicine and Hygiene*, 2004, *70* (5), 486–498.

Lehner, Bernhard and Gunther Grill, "Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the World's Large River Systems," *Hydrological Processes*, 2013, *27* (15), 2171–2186.

List, John, "On the Interpretation of Giving in Dictator Games," *Journal of Political Economy*, 2007, *115* (3), 482–493.

Lowes, Sara, Eduardo Montero, Nathan Nunn, and James A. Robinson, "Age Organization & Accountability: Evidence from the Democratic Republic of Congo," 2016. AEA RCT registration number: *AEARCTR-0005370*. J-PAL GI Funding Request. See https://www.povertyactionlab.org/evaluation/age-organization-accountability-evidence-democratic-republic-congo.

_ , Nathan Nunn, James A. Robinson, and Jonathan Weigel, "Understanding Ethnic Identity in Africa: Evidence from the Implicit Association Test (IAT)," *American Economic Review Papers and Proceedings*, 2015, *105* (5), 340–345.

_ , _ , James A. Robinson, and Jonathan Weigel, "The Evolution of Culture and Institutions: Evidence from the Kuba Kingdom," *Econometrica*, 2017, *85* (4), 1065–1091.

Michalopoulos, Stelios, "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 2012, *102* (4), 1508–39.

Murdock, George Peter, *Africa: It's Peoples and Their Cultural History*, New York, NY: McGraw-Hill Book Company, 1959.

_ , *Ethnographic Atlas*, Pittsburgh, PA: University of Pittsburgh Press, 1967.

Nunn, Nathan and Diego Puga, "Ruggedness: The Blessing of Bad Geography in Africa," *Review of Economics and Statistics*, 2012, *94* (1), 20–36.

_ and Leonard Wantchekon, "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, 2011, *101* (7), 3221–3252.

Ramankutty, Navin, Jonathan A. Foley, John Norman, and Kevin McSweeney, "The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Possible Climate Change," *Global Ecology and Biogeography*, 2002, *11*, 377–392.

Rogall, Thorsten and David Yanagizawa-Drott, "The Legacy of Political Mass Killings: Evidence from the Rwandan Genocide Evidence from the Rwandan Genocide," March 2014.

Roome, William, *Ethnographic Survey of Africa: Showing the Tribes and Languages; Also the Stations of Missionary Societies [map]*, Vol. 1:5 1924.

Van Reybrouck, David, *Congo: The Epic History of a People*, New York: Ecco, 2014.