Blood Rubber: The Effects of Labor Coercion on Development and Culture in the DRC (preliminary - please do not circulate or cite)*

Sara Lowes† Eduardo Montero‡

15 October 2014

Abstract: King Leopold's Congo Free State is infamous for its coercive and exploitative labor practices, particularly in the extraction of rubber. Between 1892 and 1906, Leopold divided parts of DRC into concessions that were granted to private companies to extract rubber. These companies used violence and coercion to force people to collect rubber. We examine the long term effects of exploitative institutions and labor coercion on development outcomes in DRC today. First, we use a geographic regression discontinuity design to demonstrate that those areas inside former rubber concessions have worse development outcomes today across a variety of measures. Second, we examine various potential channels for the persistence of these effects. We find evidence of worse public goods provision at the local level in former concessions, suggesting that local institutions were undermined during the rubber concession period.

1. Introduction

African countries continue to address the legacy of exploitative colonial institutions. European colonizers often used labor coercion to extract natural resources in the colony for the benefit of the colonizer. This paper intends to shed light on how historical exploitative institutions such as labor coercion affect current development by examining the case of the Congo Free State (CFS). The CFS, what is today the Democratic Republic of Congo (DRC), was the personal colony of King Leopold II of Belgium between 1895 and 1908. In one of the world's first human rights movements, Leopold became infamous for the use of coercive labor practices in the extraction

*We thank Nathan Nunn, Richard Hornbeck, Melissa Dell, James Robinson, Stelios Michalopoulos and participants at the Harvard Development Lunch, Brown Macro Lunch and the Harvard Economic History Lunch for excellent feedback.

†Harvard University. (email: slowes@fas.harvard.edu)

‡Harvard University. (email: emontero@fas.harvard.edu)

of natural resources such as rubber and ivory in the CFS. Historians Hothschild and Vansina estimate that 10 million people, approximately half of the population of Congo, died between 1880 and 1920 (Vansina, 2010, Hochschild, 1998). Today, the DRC remains one of Africa's most poor countries - it is ranked 186 (of 187) in the UNDP's 2012 Human Development Index.

The demand for natural rubber increased dramatically around 1890 with the invention of the pneumatic tire and remained high until 1910 when plantations in South East Asia and Brazil began producing rubber. During this period, Leopold designated parts of the present day DRC as concessions for private companies. These companies were given monopoly rights over natural resources within their boundaries. Additionally, they were given state resources – primarily soldiers from the *Force Publique*, the CFS armed forces – and a state mandate to use coercive means to reach their rubber extraction goals.

We exploit the well-defined boundaries of the historic rubber concessions to examine the long-term effects of labor coercion on economic development. Consistent with historical evidence on how the concession boundaries were determined, we demonstrate that those areas designated as concessions are geographically similar to the areas just outside of the concessions. We use DHS data from 2007 and 2014 to examine the effects of labor coercion on education, wealth and health outcomes. Additionally, we use satellite data to examine the effects of the concessions on population density. We find that the former concessions areas are significantly less educated, less wealthy, heave worse health outcomes and are less populated than areas just outside the former concessions. We analyze the DHS data by cohorts to demonstrate that there is little evidence of convergence on outcomes such as wealth and education. We explore multiple channels that might explain the persistent effects of the historical exposure to labor coercion on development today.

This paper is related to recent work on how colonial policies can have long-lasting effects on development. A growing body of empirical evidence finds that historical events are important determinants of economic development today (Nunn, 2009). In particular, many papers have focused on exploring the long-term impacts of colonial policies in Africa on modern development outcomes. For example, Nunn (2008) looks at the effect of the slave trade on economic development in Africa. Additionally, this paper proposes examining how historical events can have persistent impacts through social norms and culture. This is in line with work by Nunn and Wantchekon (2011) who examine the effect that the slave trade had on trust.

Our work also contributes to the literature on the long-term effects of labor coercion. Our paper is most related to Dell (2010), who uses a geographic regression discontinuity approach to examine the long-term impacts of the mining *mita* in Peru. On the theoretical front, Acemoglu and Wolitzky (2011) model conditions under which labor coercion will arise and persist. We hope to contribute to the literature on the interplay between culture and institutions building off theoretical work by Tabellini (2008), who models the co-evolution of culture and institutions. Finally, our paper is related to Blouin (2013) who examines how coffee quotas and historical exposure to coerced labor affect modern contracting institutions and trust between Tutsi and Hutus in Rwanda and Burundi.[1]

The paper is organized as follows. Section 2 provides historical background on the Congo Free State and the rubber concessions. Section 3 describes the data, the empirical strategy and presents the main empirical results. Section 4 presents evidence that there is not convergence and explores several different channels of persistence. Section 5 concludes.

2. The History of Rubber Concessions and Labor Coercion in DRC

By the mid-1870s, European powers had staked claims on most parts of Africa. However, the center of Africa remained largely unexplored. In a bid to make Belgium a colonial power, King Leopold II 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 Congo Free State (henceforth, 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 (Nzongola-Ntalaja, 2007).

Leopold needed to demonstrate continued state presence in the Congo in order to retain his rights over it. This proved a costly endeavor. By 1890, Leopold had invested 19 million francs in the Congo, nearly the entirety of his father's fortune (Van Reybrouck, 2014, p. 70). In an attempt to increase revenues, he declared all lands and any raw materials found on these lands to be the property of the CFS. In 1890, demand for natural rubber increased dramatically with the creation of the pneumatic tire. Congo had immense natural rubber resources, and Leopold finally saw an opportunity for profits.

[1] Alesina and Giuliano (2013) provide a nice survey of the literature on culture and institutions.

The state had limited manpower and capacity, so Leopold set up concessions to be given to private companies. Since most of the interior of DRC was unexplored at the time, the concession boundaries were defined using salient geographic characteristics such as major rivers (Harms, 1975). The Anglo-Belgian India Rubber Company (ABIR) was established in 1892 and given rights over the Maringa-Lopori basin.[2] In the same year Anversoise was created and given extraction rights near the Mongala river.[3] In return for this land, the state would collect 2 % of the concession companies' profits. Leopold himself was a majority stake holder in ABIR and Anversoise (Harms, 1975).

The concession companies forced individuals within their concessions to collect rubber as a form of paying taxes. Rubber was a unique commodity because it required little capital investment to be collected, in contrast to other natural resources such as diamonds or minerals. The intensity of rubber extraction in concession areas was thus linked to the productivity and supply of labor. Once the rubber concessions were allocated, the companies set up posts within the concession to collect rubber. One to 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. Concession companies set quotas for the collection of rubber. Male villagers were required to deliver a quota of about 4 kilos of dried rubber every 2 weeks. In addition, villages were required to provide food and supplies to maintain nearby posts. Rubber collection was both time intensive and physically demanding. Individuals would travel deep into the jungle, find a rubber vine, make incisions in the vine to let the sap trickle out, and then allow the sap to dry. This process could take days, particularly as rubber supplies dwindled and untapped rubber vines became more difficult to find. It became increasingly difficult for people to meet the rubber quotas (Harms, 1983).

Soldiers from the *Force Publique* ensured compliance with the rubber quotas.[4] Approximately 25 to 80 "post sentries" armed with rifles were assigned to each new post established. An additional 65 to 100 "village sentries" were stationed in villages surrounding the posts. Individuals were severely punished if they failed to meet their rubber quota. Punishment could take

[2]This concession area was defined by two rivers: the Maringa river and the Lopori river, plus a 25 km buffer area around them.

[3]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.

[4]These soldiers were generally Congolese recruited by the CFS from areas outside the concession areas.

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. The chief of the village could be imprisoned. Individuals could also be subjected to various forms of physical violence, including whipping by the *chicotte*, burning or death (Harms, 1983).

The soldiers from the *Force Publique* were primarily responsible for carrying out these violence tactics. 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. Testimony collected by Robert Casement, a British consul sent to Congo to investigate accusations of atrocities, documents the intensity of the labor coercion, and provided first-hand accounts of the violence from Africans directly:

> *"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."* (Casement, 1904)

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. The price of rubber went from 6.20 francs per kilo in 1894 to over 10 francs per kilo in 1898. The cost incurred by the concession companies to "purchase" a kilo of rubber in CFS and ship it to Antwerp was approximately 1.35 francs (Harms, 1983). The magnitude of profits earned by the concession companies led one contemporary observer to note "ABIR has in a single fiscal year made a net profit that represents more than twelve times the initial capital investment. Such a result is perhaps without precedent in the annals of our industrial companies" (Plas and Pourbaix, 1899).

By 1905, the 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 Anveroise left CFS in 1906. In 1908, the CFS became a Belgian colony and after 1910, competitive production of rubber from *hevea* plantations in Southeast Asia and South America, along with the invention of synthetic rubber, lead to a large decrease in rubber prices (Harms, 1975).

The regime of labor coercion had disastrous effects on the local population. Villages subjected to labor coercion 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). The brutality of the rubber collection tactics resulted in the deaths of an estimated 10 million people and earned the policy the nickname *"Red Rubber"* (Vangroenweghe, 1985)

3. Long-Term Effects of the Rubber Concessions

3.1. *Data*

To examine the long-run impact of rubber extraction by concession companies in DRC we combine survey data from DHS, satellite data from Landscan 2007, conflict data from PRIO and GIS data on geographic characteristics. The 2007 and 2014 DHS surveys from the DRC providesdetailed information on education, assets, and health outcomes for individuals in multiple villages. Data from Landscan 2007 has granular measures of population density and data from PRIO contains geo-referenced information on conflict incidence and intensity. Additionally, we use GIS data on soil suitability, precipitation and altitude to demonstrate balance on key geographic characteristics. These data sources and the variables used are described in detail in Appendix A.[5]

The maps of the rubber concessions are from Waltz (1918). This resource describes all of the concessions given by King Leopold. This includes details on the physical boundaries and the year when each concession was granted. Figure 2 is a map of the main rubber concessions of interest: ABIR and Anversoise. These were the largest concessions in the Upper Congo Basin. Figure 3 provides a map with the rubber concession borders and the DHS clusters from 2007 and 2014 that are within 200 kms of the borders of the rubber concessions.

3.2. *Summary Statistics*

Table 1 presents simple differences in means inside and outside the concession areas on variables from DHS 2007. We restrict our analysis for these differences in means to observations that are

[5]We also attempted to use nightlight data. However, as Figure 7b shows, the area of interest in DRC has little nightlight.

within 200 kms of the rubber concession borders in order to compare relatively similar areas.[6] Overall, the concession areas are less educated, less wealthy and have worse health outcomes than the areas just outside the concession borders.

Table 1: Summary Statistics

	Individuals Within 200 kms of Concession Borders			
	Mean Inside	Mean Outside	Clustered S.E.	(p-value)
Educational Attainment	1.013	1.415	(0.068)	0.0003
Obs	1,843	3,853	–	–
Years of Education	4.228	6.326	(0.364)	0.0006
Obs.	1,837	3,850	–	–
Literacy	0.884	1.289	(0.071)	0.0002
Obs.	1,836	3,829	–	–
Wealth Index	1.824	2.513	(0.157)	0.009
Obs.	1,843	3,853	–	–
Wealth Score	-54,511	-17,908	(9,560)	0.001
Obs.	1,843	3,853	–	–
Women Ht/Age Percentile	2,469	3,002	(205.5)	0.011
Obs.	545	1080	–	–
Child Ever Vaccinated	107.0	268.8	0.037	0.025
Obs.	599	1058	–	–
Child Ht/Age Percentile	2,314	2,623	182.6	0.093
Obs.	557	1046	–	–

Note: Data 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 on asset ownership. Wealth Index is a 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 1000, 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 1000. See Data Appendix for more details.

3.3. *Empirical Strategy*

A concern with the simple differences-in-means presented in Table 1 is that the rubber concession areas might be different along a number of dimensions. Specifically, the rubber concessions might have been chosen strategically for certain characteristics. For example, these areas might be more suitable for certain crops or have been populated by ethnic groups with different cultural norms. 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 concession were granted at a time when most of DRC had not been explored. The concession boundaries were defined by salient geographic characteristics such as rivers and

[6] We have also examined these differences in means – comparing areas inside the former concessions compared to areas outside the concessions – for (i) areas within 100 kms of the borders (ii) areas within 50 kms of the border and (iii) all DHS clusters in the DRC. The summary statistics are generally consistent with Table 1.

international boundaries. Thus, the boundaries are unlikely to have been selected based on local characteristics that also vary discontinuously at the border.

Following Dell (2010), we can estimate the causal effect of exposure to the rubber extraction and labor coercion on numerous outcomes by estimating the following regression discontinuity (RD) specification:

$$y_{i,v} = \alpha + \gamma RubberConcession_{i,v} + f(location_v) + X_i\beta + X_v\Gamma + \phi_{j(v)} + \epsilon_{i,v} \quad (1)$$

where $y_{i,v}$ is our outcome of interest for individual i in village v; $RubberConcession_{i,v}$ is an indicator equal to 1 if v is inside a rubber concession area and equal to 0 otherwise; X_i is a vector of covariates for individual i and X_v is a vector of geographic features for village v such as altitude;[7] $\phi_{j(v)}$ represent district fixed effects;[8] $f(DB_v)$ is the RD polynomial, which controls for smooth functions of geographic location for village v. We use various forms of the RD polynomial as robustness checks.[9] We limit our analysis to observations within 200 kms, 100 kms and 50 kms of the concession boundaries as this restricts the range in which unobservable parameters can vary. 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 concessions and others to be just outside concessions. These villages should be similar in most dimensions before the concessions were granted in terms of geography, culture, history and institutions, allowing us to identify the effect of rubber extraction on contemporaneous outcomes.

The RD approach presented in equation (1) requires two identifying assumptions. The first assumption is that all relevant factors before the concession areas were granted vary smoothly at the concession boundaries. This assumption is needed to ensure that individuals located just outside the concession are an appropriate counterfactual for those located just inside it. To assess the plausibility of this assumption, Table 2 estimates specification (1) for important geographic characteristics such as altitude and soil suitability and finds balance along these geographic

[7] In the future, we will incorporate a measure of how much "extraction" each village was exposed to by using historical records on how much rubber each village produced at the time.

[8] Specifically, for $\phi_{j(v)}$, $j(v)$ represents the function mapping each village v to district $j(v)$. In our area of interest, there are two provinces (Equateur and Orientale), eight districts and 48 territories.

[9] For our baseline results we follow Dell (2010) and present results using cubic polynomials in distance to the concession borders and cubic polynomial in latitude and longitude.

characteristics.[10] Table 3 presents results from estimating specification (1) for river characteristics such as navigable river density and access to rivers. Rivers are a particularly important geographic features for the area because they are one of the main forms of transportation. We find balance on these important geographic characteristics, suggesting that the areas inside and outside the concession are comparable along the border.[11]

A related concern is that Leopold selected the borders strategically, for example capturing only rubber-suitable areas or areas that were more dense in terms of population. However, historical evidence presented in Section 2 suggests that Leopold did not have very much information about the interior of Congo in 1890, as much of it was still unexplored. Additionally, most of the concession borders follow natural boundaries such as rivers and the Congo Free State borders.[12]

The second important assumption for this regression discontinuity approach is that there was no selective sorting across the RD threshold. While this is a concern, evidence from Harms (1975) suggests that the rubber companies greatly controlled migration (using village censuses they collected themselves) at the time and forced people to remain in their villages. Additionally,

Table 2: Balance on Geographic Characteristics

	Elevation			*Precipitation*			*Soil Suitability*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)	200 kms (7)	100 kms (8)	50 kms (9)
Inside Concession	4.569	2.466	5.267	0.523	0.470	-0.414	-0.004	-0.008	-0.025
	(3.106)	(3.071)	(3.634)	(0.459)	(0.448)	(0.516)	(0.014)	(0.014)	(0.017)
	[3.720]	[3.673]	[4.458]	[0.533]	[0.517]	[0.616]	[0.014]	[0.013]	[0.014]
Observations	1,339	843	610	1,339	843	610	156	104	76
Mean Dep. Var.	435	432	436	151	151	152	0.060	0.069	0.068

Note: The estimated regressions use a cubic polynomial on distance to the concession borders as the RD polynomial. We include district fixed effects. Elevation and precipitation come from the Global Climate Database created by Hijmans, Cameron, Parra, Jones and Jarvis (2005). This data provides monthly average rainfall in millimeters and elevation measures in meters. *Precipitation* is a measure the average yearly precipitation (in millimeters of rainfall per year) for each 20km by 20km grid cell. *Elevation* calculates the average elevation in meters for each 20km by 20km grid cell. *Soil Suitability* is from Ramankutty, Foley, Norman and McSweeney (2002) and Michalopoulos (2012). It is an index from 0-1, with higher values indicating higher soil suitability for agriculture. We present robust standard errors in ( ) and Conley standard errors in [ ] (assuming a cut-off window of 40 kms). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

[10] We use 20km by 20km grid cells to match the analysis in Dell (2010). See data appendix in Appendix A for detailed on data sources and variables used. Ideally we would also present balance on pre-colonial demographic characteristics. However, we have not been able to find detailed pre-colonial data for the DRC.

[11] These results are presented with robust standard errors and Conley standard errors. The Conley standard errors account for spatial auto-correlation (Conley, 1999). For the Conley standard errors, we use a cut-off window of 40 kms, but the results are robust to the use of different cut-offs and are available upon request.

[12] Additionally, the concession borders do not align with ethnic group borders. Figure 9a in the Appendix, a map of ethnic group boundaries from Murdock (1959) and the rubber concessions borders, shows that the rubber concession borders split many different ethnic groups.

Table 3: Balance on River Characteristics

	Navigable River Density			*Access to Navigable Rivers*			*Access to Any River*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)	200 kms (7)	100 kms (8)	50 kms (9)
Inside Concession	-0.427	-0.198	3.643*	-0.017	-0.019	0.058*	-0.041	-0.047	-0.026
	(1.799)	(1.810)	(2.136)	(0.029)	(0.030)	(0.035)	(0.036)	(0.036)	(0.042)
	[1.814]	[1.830]	[2.117]	[0.029]	[0.030]	[0.036]	[0.036]	[0.036]	[0.042]
Observations	1,339	843	610	1,339	843	610	1,339	843	610
Mean Dep. Var.	12.406	10.111	9.645	0.224	0.214	0.197	0.513	0.473	0.446

Note: We use a cubic polynomial on distance to the concession borders and include district fixed effects. *Navigable River Density* is defined as total length in meters of the respective navigable river in each grid divided by the grid's surface area in kilometers squared. *Access to Navigable Rivers* and *Access to Any River* are indicator variables equal to one if a grid cell contains a navigable river or any river, respectively. Data on navigable rivers and rivers in the DRC is from the *Referentiel Geographique Commun* (2010). We present robust standard errors in ( ) and Conley standard errors in [ ] (assuming a cut-off window of 40 kms). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Harms (1975) notes that local chiefs were held accountable when individuals that migrated did not meet their quotas, incentivizing chiefs to prevent migration. However, this assumption is unlikely to hold perfectly so we do not emphasize it and rather think of migration as potential channel for persistence.[13]

3.4. *Regression Discontinuity Results*

To examine the long run effects of exposure to extractive institutions, we analyze 2007 and 2014 DHS data and Landscan 2007 data.[14] We first focus on education and present results for observations within 200 kms, 100 kms and 50 kms of the concession border. We display results using both third-order polynomials in distance to the concession border and third-order polynomials in latitude and longitude following Dell (2010).[15] Table 4 estimates specification (1) for different education outcomes. The results are consistent with the summary statistics from Table 1: areas inside the concession have significantly lower levels of education. For example, individuals just inside the former rubber concessions are estimated to have approximately 1.5 fewer years of education than individuals just outside the concessions.

[13] We hope to eventually collect information on ancestry of current residents to examine whether or not we see large selective sorting along the border areas.

[14] The Data Appendix in Appendix A contains detailed information on the sources, the variables used and their definitions. Note that even though the DHS data has coordinates for each village, the DHS randomly displaces the exact coordinates by up to 5 km (and up to 10 km for 1% of rural clusters). Importantly, since the displacement is random, this simply induces classical measurement error. Additionally, we check the robustness of our results by dropping villages that are within 5 kms of the border and find that our results do not change significantly.

[15] Section 3.5 discusses additional polynomials and other robustness checks.

The results for years of education can be seen graphically in Figure 5 as in Dell (2010). The figure presents a geographic scatterplot of the DHS clusters shaded with the average years of education in each cluster. The background shows predicted values for a finely spaced grid of longitude-latitude coordinates from a regression using a cubic polynomial in latitude and longitude and the $RubberConcession$ dummy variable. The plot can thus be used to assess how well the RD fit is approximating the data across space.[16]

Table 5 estimates specification (1) for wealth measures in the DHS survey. Individuals in villages inside the former rubber concessions are approximately 15% less wealthy than similar individuals outside the rubber concessions. Finally, Table 6 estimate specification (1) for different health outcomes and finds evidence that areas inside the former concessions have worse health outcomes. Children inside the former concession areas have approximately 5 percentage points lower height-to-age percentile and have about 6.5 percentage points lower vaccination rates; similarly, women have approximately a 7 percentage points lower height-to-age percentile measure.[17] Overall, we find evidence that individuals residing in villages inside the former rubber concessions are less educated and less wealthy, and have worse health outcomes today than individuals in villages outside the former rubber concessions.

We next 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 approximately 1 km by 1 km resolution for the entire world. Figure 6b is a map of population data around the rubber concession areas. Table 8 presents our results from estimating specification (1) on 20 by 20 km grid cells constructed with GIS.[18] 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 1km by 1km grid

[16]These plots are available upon request for other variables but are not shown to conserve space. Likewise, standard "2D" plots are also available upon request.

[17]Note that the vaccination results are not as robust to some alternative specifications. As well, as discussed in Section 3.5 below, the health results tend to be a bit less robust than the education and wealth results. However, as described in the Data Appendix, many of these health questions are only posed to a small subset of the female respondents. This means that the sample size for these estimates is smaller and the estimates tend to be less precise.

[18]We use 20 by 20 km grid cells to match the analysis by Dell (2010). To conserve space, we present results only using 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 4: Rubber Concessions and Education RD Analysis

	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: Cubic Polynomial in Distance to Concession Border					
Inside Concession	-1.134***	-1.324***	-1.744***	-0.210***	-0.269***	-0.374***
	(0.351)	(0.355)	(0.381)	(0.075)	(0.073)	(0.078)
	Panel B: Cubic Polynomial in Latitude and Longitude					
Inside Concession	-1.552***	-1.710***	-1.640***	-0.294***	-0.365***	-0.394***
	(0.383)	(0.376)	(0.410)	(0.078)	(0.081)	(0.080)
Observations	5,629	4,233	2,623	5,633	4,237	2,626
Clusters	109	84	52	109	84	52
Mean Dep. Var.	5.648	5.130	5.209	1.173	1.068	1.077

Note: We cluster standard errors at the DHS cluster level. We include district fixed effects and control for age, age squared and gender. *Literacy* is a 0 to 2 categorical variable where 0 is cannot read at all and 2 is able to read a whole sentence. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 5: Rubber Concessions and Wealth RD Analysis

	Wealth Index			*Wealth Factor*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)
	Panel A: Cubic Polynomial in Distance to Concession Border					
Inside Concession	-0.471***	-0.556***	-0.616***	-11,353**	-16,744***	-20,415***
	(0.144)	(0.151)	(0.182)	(5,543)	(5,308)	(6,483)
	Panel B: Cubic Polynomial in Latitude and Longitude					
Inside Concession	-0.719***	-0.811***	-0.598***	-19,084***	-23,070***	-17,504**
	(0.171)	(0.191)	(0.199)	(6,716)	(6,502)	(6,746)
Observations	5,638	4,240	2,627	5,638	4,240	2,627
Clusters	109	84	52	109	84	52
Mean Dep. Var.	2.290	2.036	2.101	-29751	-46136	-43799

Note: We cluster standard errors at the DHS cluster level. We include district fixed effects and control for age, age squared and gender. *Wealth Factor* is an index generated by the DHS using principle component on asset ownership. *Wealth Index* is a is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Factor Score. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 6: Rubber Concessions and Health RD Analysis

	Child Ever Vaccinated			*Child Ht/Age Percentile*			*Respondent Ht/Age Percentile*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)	200 kms (7)	100 kms (8)	50 kms (9)
	Panel A: Cubic Polynomial in Distance to Concession Border								
Inside Concession	-0.083**	-0.084**	-0.090*	-410.0**	-451.2**	-576.4***	-710***	-791***	-900***
	(0.035)	(0.036)	(0.046)	(183.41)	(181.7)	(201.3)	(210.2)	(217.2)	(258.6)
	Panel B: Cubic Polynomial in Latitude and Longitude								
Inside Concession	-0.057	-0.050	-0.086	-444.0**	-590.5***	-395.9**	-780***	-902***	-844***
	(0.043)	(0.052)	(0.054)	(199.5)	(190.0)	(173.4)	(235.2)	(252.2)	(293.9)
Observations	3,161	2,533	1,627	1,596	1,305	822	1,578	1,207	758
Clusters	109	84	52	109	85	52	109	84	52
Mean Dep. Var.	0.816	0.799	0.793	2516	2453	2481	2693	2606	2628

Note: We cluster standard errors at the DHS cluster level. We include district fixed effects in all regressions. We control for age and age squared. The DHS health questions are only asked to a subset of female respondents. Respondent *Ht/Age Percentile* divides each respondent's height by their age and finds their percentile in the entire sample and normalizes this percentile to be within 0 and 10000. Similarly, Child *Ht/Age Percentile* divides each child's height by their age and finds their percentile in the entire sample and normalizes this percentile to be within 0 and 10000. *Child Ever Vaccinated* is an indicator variable equal to one if the respondent's child has ever been vaccinated. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

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.

Finally, we collected colonial data on population density from the *Atlas Général du Congo et du Ruanda-Urundi* (1954) in order to examine whether the population density differences are a common pattern throughout the history of this area or a more recent phenomenon. The *Atlas Général du Congo et du Ruanda-Urundi* (1954) contains extremely detailed population density maps for Equateur province for 1954. Table 8 presents our results from estimating specification (1) on 20 by 20 km grid cells constructed with GIS, where the dependent variable is the average

Table 7: Rubber Concession and Population Density RD Analysis

	Mean Population Density			*Mean Population Density (River Corrected)*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)
Inside Concession	-3.187***	-3.094***	-3.752***	-3.559***	-3.913***	-4.153***
	(1.097)	(1.080)	(1.264)	(1.258)	(1.293)	(1.510)
	[1.111]	[1.104]	[1.332]	[1.176]	[1.136]	[1.355]
Observations	1,325	835	602	1,325	835	602
Mean Dep. Var.	12.833	15.580	15.834	13.179	15.940	16.139

Note: We use a cubic polynomial on distance to the concession borders, include district fixed effects and control for elevation and presence of a river. River correction scales population density by percent of non-river land. Data is from Landscan 2007. We present robust standard errors in ( ) and Conley standard errors in [ ] (assuming a cut-off window of 40 kms). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

number of individuals per square km. We find that areas inside the former rubber concessions were also less populated in 1954 than areas outside. Areas inside the former concession borders have approximately 1 fewer person per 1km by 1km grid cell on average (this also corresponds to about 25 % fewer people per square kilometer). Thus, we find evidence that areas inside the former rubber concessions have been significantly less populated since the rubber extraction period.

3.5. *Robustness of DHS Results*

There are two main empirical concerns for the DHS results presented in Tables 4-5. The first concern is whether the results are robust to alternative specifications of the RD-polynomial. We find that our wealth and education results are robust to parsimonious polynomials (linear, quadratic, quartic, interacted-linear, interacted-quadratic) but our results begin to lose significance with higher-order polynomials in distance to the concession border (interacted third-order, interacted-quartic).[19] However, the coefficient magnitudes and sign all remain similar across most specifications, suggesting that we lose significance with higher-order polynomials due to over-fitting and not due to more precise estimation. The health results are slightly less robust to higher-order RD polynomials compared to the education and wealth results. However, as detailed in the Data Appendix, these questions are only asked to a subsample of the population (about a third of all females and children)

Table 8: Rubber Concession and Population Density in Equateur (1954)

	Population Density in 1954		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)
Inside Concession	-1.888*** (0.350) [0.454]	-2.027*** (0.362) [0.460]	-1.763*** (0.483) [0.617]
Observations	934	654	403
Mean Dep. Var.	3.368	3.871	4.059

Note: We use a cubic polynomial on distance to the concession borders, include district fixed effects and control for elevation and river density. Data is from 1954 and is from the *Atlas Général du Congo et du Ruanda-Urundi* (1954). *Population Density in 1954* measures the mean number of people per square km for Equateur province. We present robust standard errors in ( ) and Conley standard errors in [ ] (assuming a cut-off window of 40 kms). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

[19] We present the results for these specifications in the Appendix Section B.1.1. By "interacted" polynomial we mean that we interact the "Inside Concession" indicator with all terms in the polynomial.

Additionally, we use alternative latitude-longitude specifications where we modify $f(.)$ in Equation (1) to be a function of latitude, longitude, and their interactions.[20] Once again, we find that our education and wealth results results are robust to lower-order polynomials in latitude and longitude (linear, quadratic and cubic polynomials) but begin to lose significance with higher-order polynomials in latitude and longitude (fourth order polynomials and above). Nevertheless, the estimated coefficients from the latitude-longitude specifications generally have the same sign and are of similar magnitudes with the distance-to-the-border specifications from Tables 4-6.[21] Overall, we find that our results are robust to alternative RD polynomials.

4. Channels of Persistence

The results presented in Section 3.4 suggest that the effect of being a former rubber concession has large long-term consequences for development today. However, it is not clear why this effect persists. This is particularly puzzling for a few reasons. The rubber extraction period was relatively short (1892 to 1906) and the concession companies left these areas over 100 years ago. As Table 2 shows, the areas inside and outside former concession zones are similar on key geographic characteristics. Thus, in a region that relies primarily on agriculture as a source of income, we would expect these areas to converge over time.[22] Additionally, local populations do not use rubber plants for production or consumption; thus, differences in rubber resource endowments after the concession companies left are unlikely to be relevant for present day development.

In this section, we explore several potential channels that might explain why the former rubber concession areas are less developed today. There are several potential channels through which the rubber extraction policies could have affected development today. The rubber extraction could have (i) affected subsequent Belgian colonial policies and colonial infrastructure investments, (ii) increased inter-group hostility, (iii) undermined local institutions and public goods provision, and

[20]These results are presented in the Appendix in Section B.1.2.

[21]The health results are less robust to these alternative specifications as the sample size for these estimates is smaller.

[22]In particular, given the population density results, 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, 2013) and the 1609 Spanish expulsion of the Moriscos (Chaney and Hornbeck, 2013) – suggest that the concession areas would have converged to a similar level of development by now.

(iv) affected cultural norms. We test some these channels using available data and suggest future work to test additional channels.

4.1. On a Convergence Path?

Before examining channels for persistence, it is important to establish that the results presented in Section 3.4 are indeed persistent. That is, we need to check that it is not the case that areas inside the former rubber concessions are in fact on a path to convergence with areas outside the former concession zones 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 Tables 4-6.[23] To do this, we compare 5-year cohorts inside and outside the concessions by estimating a regression that includes cohort fixed effects along with the interactions between the $InsideConcession$ indicator and the cohort fixed effects.[24] Formally, we estimate the following specification:

$$y_{i,v} = \gamma InsideConcession_{i,v} + \alpha_y C_y + \gamma_y C_y \times InsideConcession_{i,v} + X_i\beta + X_v\Gamma + \phi_{j(v)} + \epsilon_{i,v} \quad (2)$$

where C_y are 5-year cohort fixed effects and the other variables are defined as in equation (1).[25]

Figure 1 plots the estimated cohort coefficients for years of education, wealth, women's height-to-weight, and literacy. 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.

4.2. Differences in Subsequent Colonial Policies

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 colonial data from the *Atlas Général du Congo et du Ruanda-Urundi* (1954) and Rouck (1945) to assess whether colonial policies and investments were different in the former concession areas relative to areas just outside the border.[26]

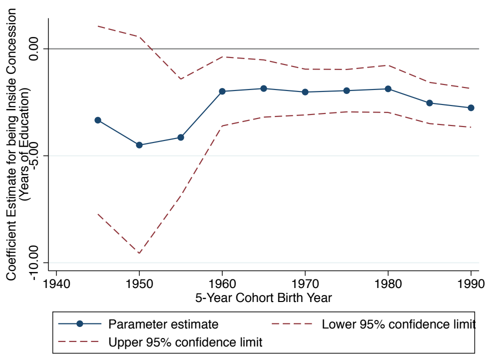
[23]As well, this allows us to test that our results are not driven by older cohorts.

[24]We do this for observations within 200 kms of the concession borders.

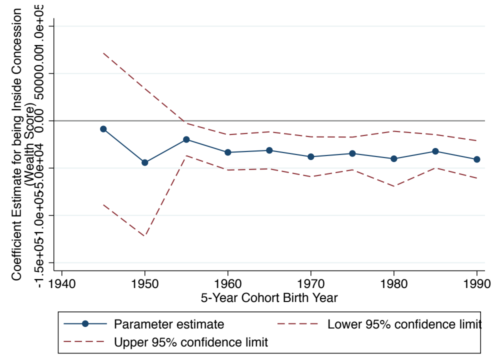
[25]Note that we are not estimating a distinct RD polynomial for each cohort for each cohort as that would be too demanding of the data given our sample size.

[26]The data appendix describes our data sources and variable definitions in detail.

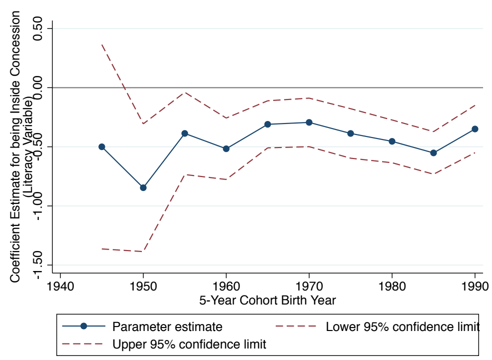
Figure 1: Estimated Cohort Coefficients for Individuals within 200 kms of the Rubber Concessions



(a) Years of Education



(b) Wealth Factor Score



(c) Literacy

Note: These figures plot the estimated coefficient for each 5 year cohort fixed effect 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 for years of education, wealth and literacy 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.

4.2.1. Missionary Presence

During the colonial period, Catholic and Protestant missions were the primary providers of schooling (Hochschild, 1998). Given the differences in education found in Table 4, differences in missionary presence might be responsible for the results if missionaries or colonial officials decided not to engage as much with the former concession areas. Using data from Nunn (2010) on missionary posts in 1924 and colonial maps from 1897 and 1953, Table 9 presents results from estimating equation (1) to test whether areas inside the concessions had fewer missionary posts in 1897, 1924 and 1953.[27] We find no evidence that areas inside and outside the concessions had significantly different missionary presence during the colonial period.

[27]Data from 1897 are from Rouck (1945) and data from 1953 is from the *Atlas Général du Congo et du Ruanda-Urundi* (1954).

Table 9: Rubber Concessions and Missionary Stations

	Number of Missionary Stations in:								
	1897			*1924*			*1953*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)	200 kms (7)	100 kms (8)	50 kms (9)
Inside Concession	0.0027	0.0020	0.0039	0.016	0.016	0.027*	-0.047	-0.037	-0.089**
	(0.0020)	(0.0020)	(0.0040)	(0.012)	(0.013)	(0.016)	(0.030)	(0.028)	(0.035)
	[0.0020]	[0.0019]	[0.0038]	[0.013]	[0.013]	[0.017]	[0.033]	[0.029]	[0.035]
Observations	1,339	843	610	1,339	843	610	1,339	843	610
Mean Dep. Var.	0.0022	0.0012	0.0016	0.025	0.026	0.026	0.181	0.254	0.254

Note: We use a cubic polynomial on distance to the concession borders and include district fixed effects. Data from 1924 are from Nunn (2010) and is available on Nathan Nunn's website. Data from 1897 are from Rouck (1945) and data from 1953 is from the *Atlas Général du Congo et du Ruanda-Urundi* (1954). Number of Missionary Stations in each year is a measure of the number of missions in each 20 by 20 km grid cell for each year that we have a map with the exact locations of missions. We present robust standard errors in ( ) and Conley standard errors in [ ] (assuming a cut-off window of 40 kms). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

4.2.2. Colonial Infrastructure Investments

Even though the Belgian colonial government was not primarily responsible for the provision of schooling due to the large missionary presence in the DRC, the government did provide infrastructure investment and other public good provisions (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 *Atlas Général du Congo et du Ruanda-Urundi* (1954), we test whether areas inside the former concession borders had fewer telecommunication stations and health centers in 1953 and lower road network density in 1968.[28] We find little evidence that colonial investments in these goods were different inside and outside the concessions: areas inside the concessions had similar numbers of telecommunication stations and health centers in 1953, and similar (but sightly lower) road network density in 1968.

*4.3. **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.[29]

[28]The DRC achieved independence in 1960 and dealt with political instability in the following years (Van Reybrouck, 2014); thus, road network density in 1968 serves as an appropriate proxy for colonial road investments even though the data is from after independence.

[29]Exposure to violence could desensitize individuals and makes violent reactions less costly and seem more normal.

Table 10: Rubber Concessions and Colonial Infrastructure Investments

	Number of Telecomm Stations in 1953			*Number of Health Centers in 1953*			*Road Network Density in 1968*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)	200 kms (7)	100 kms (8)	50 kms (9)
Inside Concession	-0.014	-0.014	-0.014	-0.047	-0.045	-0.060	-4.567*	-3.924	-2.708
	(0.013)	(0.014)	(0.016)	(0.035)	(0.036)	(0.042)	(2.612)	(2.678)	(3.091)
	[0.013]	[0.013]	[0.016]	[0.036]	[0.038]	[0.044]	[3.154]	[3.161]	[3.548]
Observations	1,339	843	607	1,339	843	610	1,339	843	607
Mean Dep. Var.	0.030	0.033	0.033	0.199	0.246	0.246	31.14	35.23	35.23

Note: We use a cubic polynomial on distance to the concession borders. We include district fixed effects. 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 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. *Road Density per Capita* is *Road Density* divided by mean population density from Landscan 2007 used in Table 8. We present robust standard errors in ( ) and Conley standard errors in [ ] (assuming a cut-off window of 40 kms). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

These differences could have led to areas inside the former concession to experience more violence and conflict, and this could have hindered economic development.

We test for differences in violence use data from PRIO that documents the location and intensity of major conflict events in the DRC since 1989. Table 11 examines whether areas inside former concessions experience more violent conflict.[30] We find no evidence that these areas experience more conflict. Note however 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. We find weak evidence for this channel and can not conclude that this explains the persistence.

Table 11: Rubber Concession and Conflict RD Analysis

	Number of Conflict Events			*No. of Conflict Events (Controlling for Ethnic Boundary)*			*Number of Civilian Deaths*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)	200 kms (7)	100 kms (8)	50 kms (9)
Inside Concession	0.046	0.040	0.025	0.048*	0.037	0.020	-3.395	-3.753	-6.320
	(0.028)	(0.027)	(0.022)	(0.028)	(0.026)	(0.021)	(2.636)	(2.947)	(5.166)
	[0.0273]	[0.0263]	[0.0211]	[0.0274]	[0.0255]	[0.0213]	[2.629]	[3.047]	[4.113]
Observations	1,337	843	610	1,337	843	610	1,337	843	610
Mean Dep. Var.	0.088	0.070	0.067	0.088	0.070	0.067	1.907	2.090	2.481

Note: We use a cubic polynomial on distance to the concession borders. Conflict data is from the Uppsala Conflict Data Program, ethnicity boundary data comes from Murdock (1959) geo-referenced by Nathan Nunn, and population density data is from LandScan 2007. We include district fixed effects and control for percentage river and population density. The specification with ethnic boundaries includes an indicator for whether a cell is on the border of two ethnic groups. We present robust standard errors in ( ) and Conley standard errors in [ ]. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

[30] The dependent variable is total amount of conflict in 20 km by 20 km grid cells.

4.4. *Market Access and Public Good Provision Since Independence*

One channel of persistence could be due to differences in investment in market access and public goods post-independence. These investment levels could have been different for two reasons: (i) local governments did not invest in public goods and infrastructure maintenance in former concession areas,[31] and (ii) individuals in areas that experienced the rubber concessions became substantially less trusting of outsiders and chose to invest less in public goods and infrastructure.

Using data from the *Referentiel Geographique Commun* on current road networks and bridges in DRC today, Table 12 examines whether areas inside the former concession have lower market access today.[32] We find strong evidence that areas inside the former rubber concessions have fewer roads and bridges today.[33] Thus, we find evidence that differences in public good and infrastructure investments are a plausible channel of persistence in this setting but we cannot distinguish between hypotheses (i) and (ii). As we outline in Section 4.5.1, in the future we will distinguish between (i) and (ii) by incorporating more post-independence data and potentially conducting lab experiments in the field for levels of trust.

Table 12: Rubber Concessions and Market Access RD Analysis

	Number of Bridges in 2010			*Road Density in 2010*			*Road Density Per Capita in 2010*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)	200 kms (7)	100 kms (8)	50 kms (9)
Inside Concession	-0.048**	-0.047**	-0.046*	-17.01***	-16.22***	-11.07***	-1.434***	-0.615**	-0.111
	(0.019)	(0.020)	(0.024)	(3.293)	(3.357)	(3.893)	(0.338)	(0.288)	(0.342)
	[0.020]	[0.020]	[0.023]	[4.348]	[4.411]	[4.881]	[0.458]	(0.361]]	[0.424]
Observations	1,339	843	610	1,339	843	610	1,336	843	610
Mean Dep. Var.	0.062	0.056	0.056	55.97	56.71	56.71	5.736	4.414	4.414

Note: We follow Dell (2010) and use a cubic polynomial on distance to the concession borders. We include district fixed effects. 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 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. *Road Density per Capita* is *Road Density* divided by *Mean Population Density* from Landscan 2007 used in Table 8. We present robust standard errors in ( ) and Conley standard errors in [ ] (assuming a window of 40 kms). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

[31] After independence, the central government suffered from constant political instability (Van Reybrouck, 2014); thus, much of the subsequent maintenance of roads and the provision of public goods was undertaken by local governments instead.

[32] Figures 4 and 6 provide maps with the GIS data on road networks and bridges.

[33] The differences in bridges is especially interesting because, as Table 3 shows, areas inside and outside the former concessions have similar geography and river densities.

4.5. *Differences in Culture as a Result of the Rubber Extraction*

As mentioned above in Section 4.4 hypothesis (ii), cultural norms may account for the persistence of the negative impacts of the coercive rubber extraction. There are many ways in which the extreme violence suffered inside the concession areas could have affected cultural norms that were then passed down through generations. In particular, building off of the historical evidence presented in Section 2, we posit that social norms could have changed as a result of the rubber extraction in a few specific ways:

1. 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. Additionally, as the intensity of rubber extraction increased, villages were competing over increasingly scarce resources, leading to historical conflict across villages and ethnic groups.[34]

2. The rubber extraction could have decreased social norms of trust.[35] Specifically, the rubber extraction could have (i) decreased trust in outsiders (since the labor coercion was carried out by outsiders – both Europeans and Congolese outsiders), (ii) decreased trust in the state (as the state granted *Force Publique* soldiers to the concession companies to enforce the labor coercion), or (iii) decreased trust in local institutions (as these institutions were unable to stop the labor coercion and were complicit in enforcing rubber quotas).

3. The rubber extraction and the intense competition for rubber amongst villagers might have changed social norms of fairness within villages and across local villages.

4. Similarly, competition for rubber might have reduced social norms of cooperation within and across villagers.

Currently available data does not allow us to test for these various channels. The next section outlines our next steps to be able to test these channels.

[34]The idea here would be that exposure to violence desensitizes individuals and makes violent reactions less costly and seem more normal.

[35]For historical anecdotes of this decreased trust by villagers due to the effects of the rubber trade see Conan Doyle (1909)

4.5.1. Steps for Future Work: Using Behavioral Games to Identify Cultural Channels of Persistence

The results presented above in Section 4.5 suggest that culture could be one of the channels of persistence. However, with the available data, we have not been able to show this definitively. For instance, we have not been able to test whether areas inside the concessions have lower levels of trust. For future work, we propose exploiting the well-defined concession boundaries to examine the persistent long-term effects of forced labor on beliefs and cultural attitudes, as measured by behavioral games.

Building off of recent work by Lowes, Nunn, Robinson and Weigel (2014) who use behavioral games to examine the long-term effects of the Kuba Kingdom, we hope to conduct a series of behavioral games results to examine differences in social norms between former concession areas and the areas outside the concessions. This would allow us to test how labor coercion affected both social norms and local institutions in these areas.

5. Conclusion

This paper exploits the well-defined boundaries of concessions given to private companies to extract rubber during the Congo Free State era to the examine the long-term effects of extractive institutions on development. Previous research has shown that historical events of long duration matter for long run development. We demonstrate that the 12 year exposure to extractive institutions has affected the development of this region of Congo. Former rubber concession areas have lower levels of education, worse health outcomes and are poorer than areas outside of the concessions. We find no evidence that areas inside the former concessions are converging to the development levels of areas outside the former concessions.

We explore different channels to explain the persistent effects of rubber extraction in DRC. We find little evidence that the differences in development are driven by subsequent differential colonial treatment or missionary presence, and find no evidence for differences in conflict. However, we find evidence of worse public goods provision, suggesting an effect on local institutions and culture. For future work, we plan to use behavioral games and additional post-independence data to distinguish between these two channels. Overall, we hope to provide an understanding of why coercive labor and colonial extraction can often have persistent long-term effects on development.

References

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

Alesina, Alberto and Paola Giuliano, "Culture and Institutions," December 2013. NBER Working Paper No. 19750.

Angrist, Joshua D. and Jorn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.
Atlas Général du Congo et du Ruanda-Urundi

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

Blouin, Arthur, *"Culture and Contracts: The Historical Legacy of Forced Labour," November 2013. Job Market Paper.*

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.*

Casement, Roger, *Casement Report 1904.*

Chaney, Eric and Richard Hornbeck, *"Economic Dynamics in the Malthusian Era: Evidence from the 1609 Spanish Expulsion of the Moriscos," 2013. Working Paper.*

Conan Doyle, Arthur, *The Crime of the Congo, London, United Kingdom: Hutchinson & Co., 1909.*

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.*

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

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and HŇvard Strand, *"Armed Conflict 1946-2001: A New Dataset," Journal of Peace Research, 2002, 39 (5).*

Harms, Robert, *"The End of Red Rubber: A Reassessment," The 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.*

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.*

Lowes, Sara, Nathan Nunn, James Robinson, and Jon Weigel, *"The Evolution of Culture and Institutions: Evidence from the Kuba Kingdom," 2014. Working Paper.*

MacKinnon, James G. and Matthew D. Webb, *"Wild Bootstrap Inference for Wildly Different Cluster Sizes," 2013. Working Paper.*

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

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

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

_ , *"The Importance of History for Economic Development," Annual Review of Economics, 2009, 1 (1), 65–92.*

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

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

Nzongola-Ntalaja, Georges, *The Congo from Leopold to Kabila: a People's History, New York, NY: Zed Books, 2007.*

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," 2013. Working Paper.*

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

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

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

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

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

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

Maps and Figures

Figure 2: ABIR and Anversoise Rubber Concession Areas

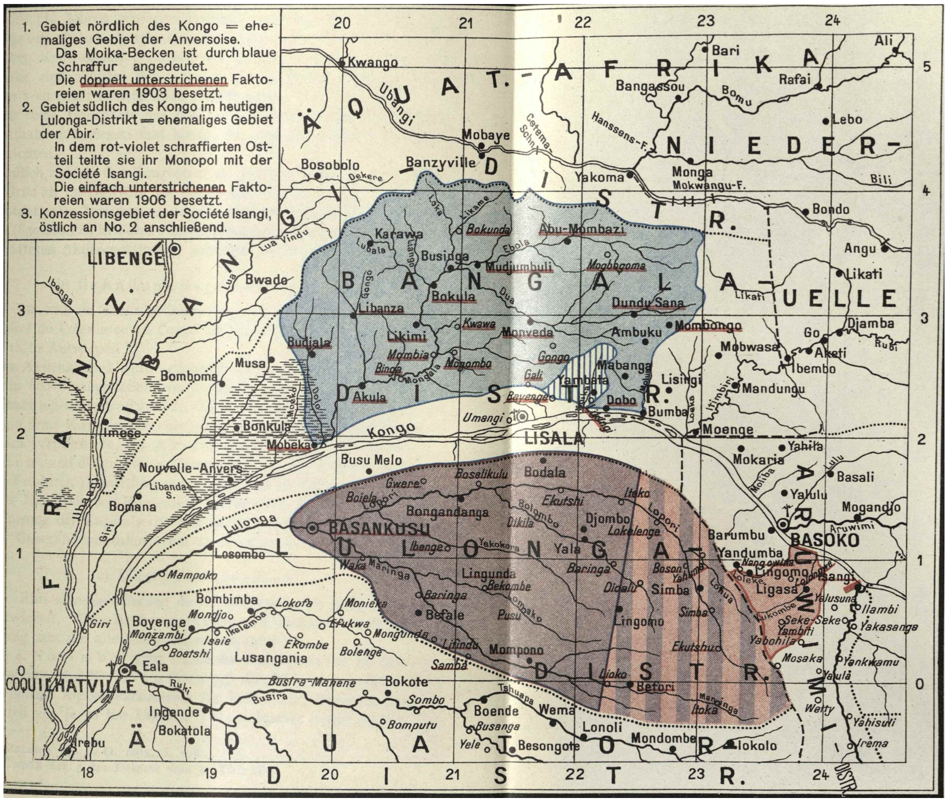
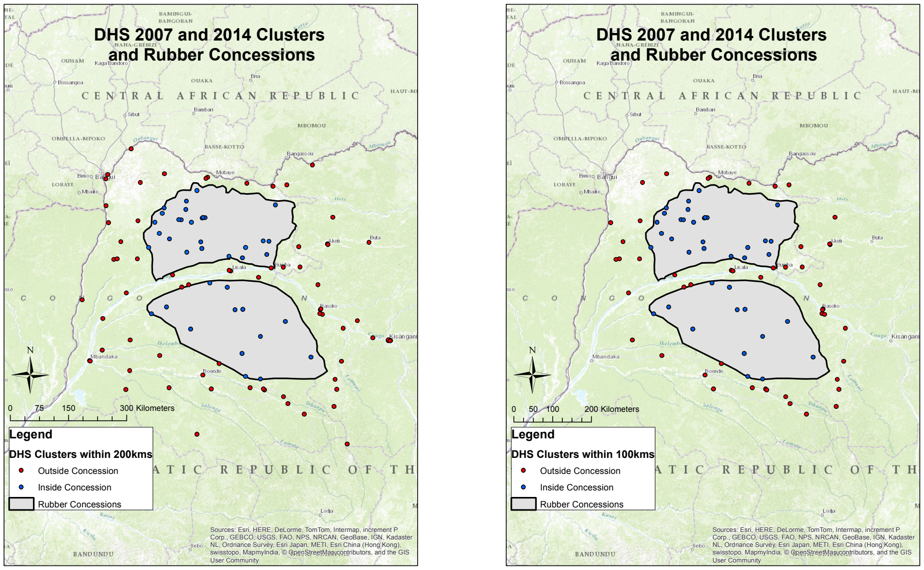


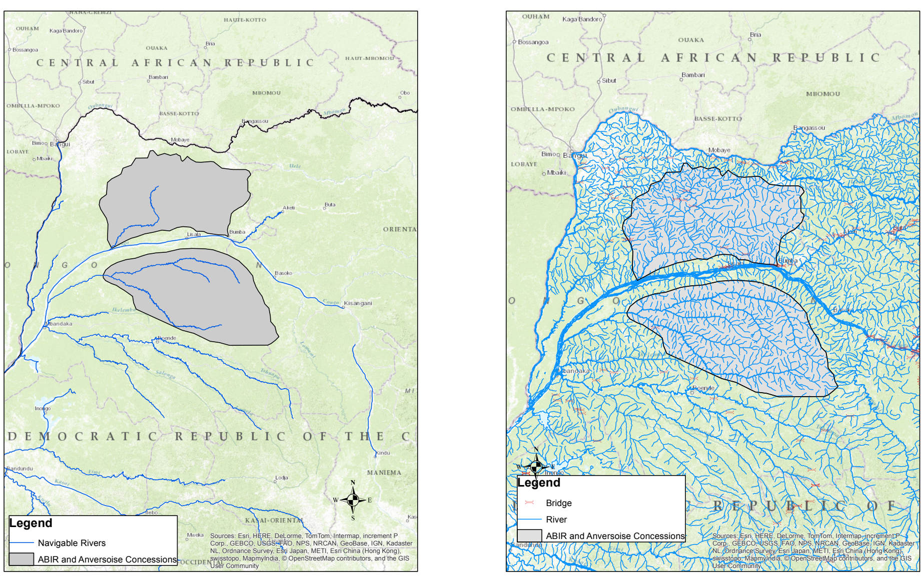
Figure 3: Map of Sample Used from the DHS 2007 and 2014 DRC Surveys - Within 200 km and 100 km



(a) Within 200km

(b) Within 100km

Figure 4: Rivers, Bridges and Rubber Concessions



(a) Navigable Rivers

(b) All Rivers

Figure 5: RD Plot for Years of Education

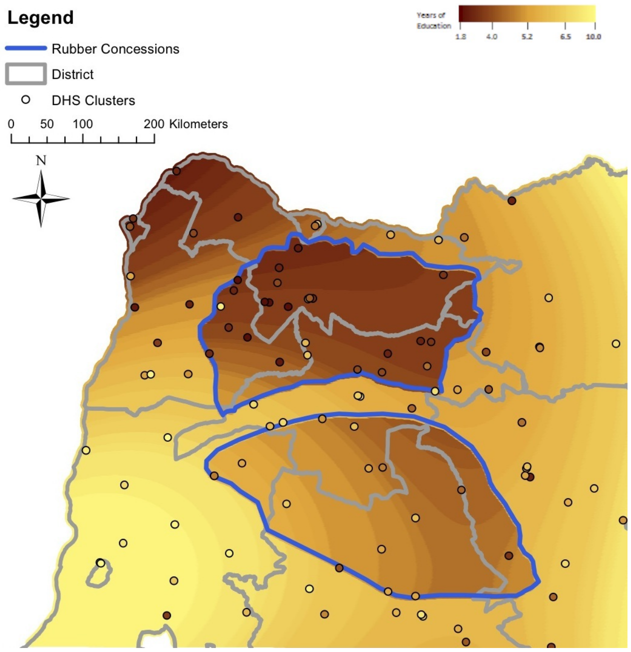
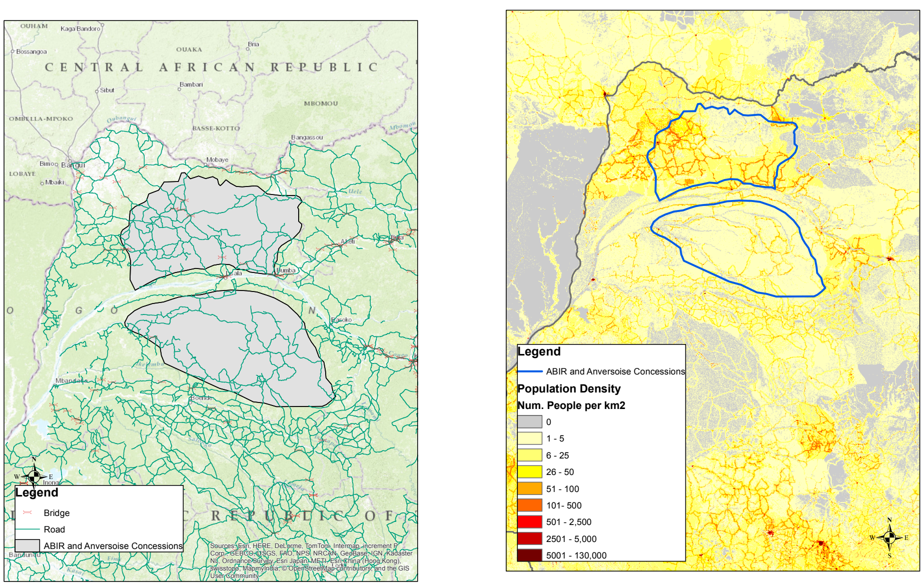


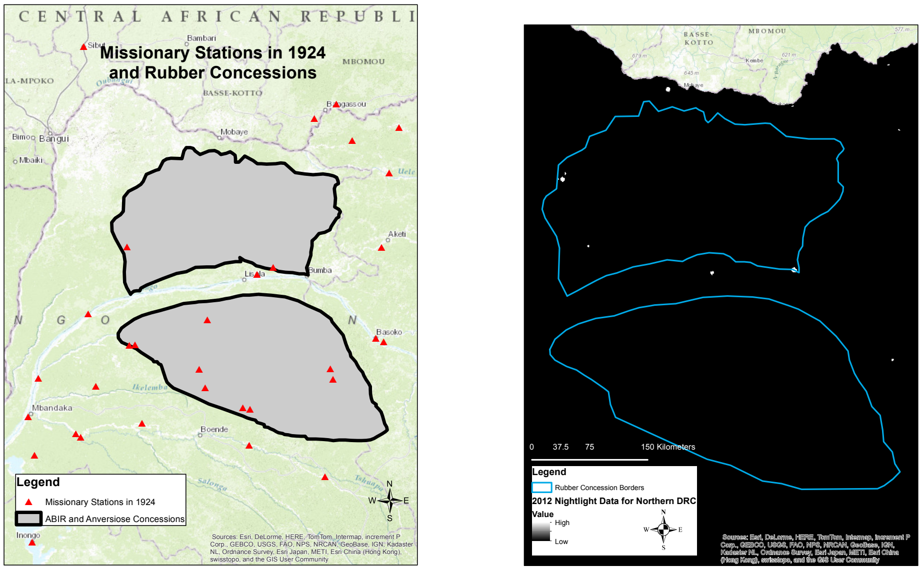
Figure 6: Road Networks, Population Density and Rubber Concessions



(a) Road Network

(b) Population Density Measure from Landscan 2007

Figure 7: Missions in 1924, Nightlights today and Rubber Concessions



(a) Missionary Stations in 1924 and Rubber Concessions

(b) Nightlights in Northern DRC

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 et al. (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.[36]

- **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.

- **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 et al. (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

[36] See Figure 10 for a Map of the Grid Cells.

higher values indicate higher soil suitability for agriculture. Our Soil Suitability variable measures the average soil suitability in each 20km by 20km grid cel to provide a measure of soil suitability that also ranges between 0 and 1, with higher values indicate higher soil suitability for agriculture.

- **Crop Suitability:** Crop suitability reefers to the average suitability for rain-fed, low-input crops provided by the FAO's Global Ecological Zones website: `http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm`. FAO crop suitability model uses data on elevation, precipitation, soil and slope constraints to construct estimates of crop suitability at the 1 km^2 level for different crops. This measure is normalized to be between 0 and 1, where higher values indicate higher crop suitability.

- **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.

A.2. 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.

- **Educational Attainment:** Educational Attainment is a 0 to 3 categorical variable that measures the highest education level attained, where 0 is no education, 1 is primary education, 2 is secondary and 3 is higher education.

- **Literacy:** Literacy is a 0 to 2 categorical variable for each individual where 0 is "cannot read at all", 1 is "able to read only parts of a sentence" and 2 is "able to read a whole sentence".

- **Wealth Factor:** Wealth Factor is an index generated by the DHS using principle component analysis on asset ownership for each individual.

- **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.

- **Number of Deceased Children:** Number of Deceased Children is the total number of male and/or female children that have died (at any age) for each respondent.

- **Number of Children:** Number of children measures the total number of children ever born for each respondent.

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:

- **Ever Vaccinated:** The DHS female-subsample includes question whether or not an individual has ever had a vaccination in her life. Thus, *Ever Vaccinated* is an indicator variable equal to 1 if the respondent has ever received a vaccination in their lifetime.

- **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 1000.

- **Body Mass Index:** Body Mass Index measures the body mass index, or Quetelet's index, for each respondent. BMI is is defined as her weight in kilograms divided by the square of her height in meters (kgs/m^2). Note that in the DHS variable there are two implied decimal place in the BMI, and so to produce the exact BMI one would divide by 100. In this paper we left the variable as-is in the DHS file and did not divide by 100 for ease of exposition in the tables. As well, the DHS does not adjust BMI for pregnant women; thus, we drop all women in our sample who are pregnant for regressions that include respondent BMI.

- **Respondent Wt/Ht Std Deviations:** This variable measures the standard deviation of the ratio of height-to-weight (in kgs/cms) and calculates the z-score for each respondent's weight (kgs) to height (cms) as compared to the entire sample and normalizes this z-score by 1000 (so that it falls within -1000 and 1000).

- **Wife Beating is Justified if ....:** The aforementioned subsample also provides questions on when wife beating is justified. Importantly, this question is only asked to females in the sample. The respondent is asked to answer Yes or No to whether wife beating is justified under different scenarios. Thus, the variables "Wife Beating is Justified if Goes Without Saying", "Wife Beating is Justified if Wife Argues" and "Wife Beating is Justified if Wife Refuses Sex" are indicator variables equal to one if the respondent agrees with the respective statement.

A.3. Population Density Data and Variables:

We use data from Landscan 2007 provided by the Oak Ridge National Laboratory to get a measure of population density for each 20 km by 20 km grid cell. The Landscan 2007 algorithm uses detailed satellite imagery and imagery analysis technologies to disaggregate census counts within administrative boundaries in order to construct measures of population density at approximately 1 km by 1 km resolution for the entire world. The units for the Landscan data is therefore the total number of people per square kilometer. The Landscan 2007 data and detailed information on the data construction is available at `http://www.ornl.gov/sci/landscan`. Our variable for *Mean Population Density* measures the average number of individuals per square kilometer for each 20 km by 20 km grid cell. Our variable for *Mean Population Density (River Corrected)* scales this *Mean Population Density* by the percent of non-river land for each 20 km by 20 km grid cell.

A.4. Market Access 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.

A.5. 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 *Atlas Général du Congo et du Ruanda-Urundi* (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 *Atlas Général du Congo et du Ruanda-Urundi* (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 *Atlas Général du Congo et du Ruanda-Urundi* (1954). The *Atlas Général du Congo et du Ruanda-Urundi* (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-CastaŰeda 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.

- ***Force Publique* Stations**: *Force Publique* station location for 1914, 1932 and 1953 for the DRC are from the *Atlas Général du Congo et du Ruanda-Urundi* (1954). The *Atlas Général du Congo et du Ruanda-Urundi* (1954) includes a map with *Force Publique* station locations for these years that was digitized in ArcGIS. Our variable *Number of Force Publique Stations* for a given year – 1914, 1932 or 1953 – is defined as the total number of *Force Publique* station in that year located in each 20 km by 20 km grid cell.

A.6. Conflict Data and Variables:

- **Conflict Events**: Conflict data is provided by Gleditsch, Wallensteen, Eriksson, Sollenberg and Strand (2002) and available online at `http://www.prio.org/Data/Armed-Conflict/UCDP-PRIO/`. This data provides geo-referenced data on major conflict events since 1948 for the entire world as best as possible. However, conflict data for the DRC is only available starting 1989. The dataset provides a shapefile with geo-referenced points for each major conflict event along with data on the scale of each event. Our variable *Number of Conflict Events* measures the total number of conflict events since 1989 for each 20 km by 20 km grid cell. *Number of Civilian Deaths* measures the total number of civilian deaths as a result of major conflict since 1989 for each 20 km by 20 km grid cell.

Table A1: Rubber Concessions and Education RD Analysis

Alternative RD 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 Distance to Concession Border					
Inside Concession	-1.343***	-1.476***	-1.798***	-0.226***	-0.279***	-0.352***
	(0.348)	(0.357)	(0.394)	(0.073)	(0.072)	(0.078)
	Panel B: Quadratic Polynomial in Distance to Concession Border					
Inside Concession	-1.344***	-1.464***	-1.834***	-0.226***	-0.273***	-0.358***
	(0.347)	(0.355)	(0.385)	(0.072)	(0.071)	(0.077)
	Panel C: Quartic Polynomial in Distance to Concession Border					
Inside Concession	-1.347***	-1.419***	-1.828***	-0.234***	-0.263***	-0.356***
	(0.351)	(0.360)	(0.366)	(0.073)	(0.071)	(0.079)
	Panel D: Interacted Linear Polynomial in Distance to Concession Border					
Inside Concession	-1.762***	-1.883***	-1.332*	-0.304***	-0.323***	-0.233
	(0.489)	(0.520)	(0.765)	(0.108)	(0.110)	(0.152)
	Panel E: Interacted Quadratic Polynomial in Distance to Concession Border					
Inside Concession	-1.725***	-2.039**	-2.039**	-0.282**	-0.340**	-0.218
	(0.656)	(0.830)	(0.924)	(0.133)	(0.166)	(0.190)
Observations	5,629	4,233	2,623	5,633	4,237	2,626
Clusters	109	84	52	109	84	52
Mean Dep. Var.	5.648	5.130	5.209	1.173	1.068	1.077

Note: We cluster standard errors at the DHS cluster level. We include district fixed effects and control for age, age squared and gender. *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. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A2: Rubber Concessions and Wealth RD Analysis

Alternative RD Polynomials

	Wealth Index			*Wealth Factor*		
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 Distance to Concession Border					
Inside Concession	-0.427*** (0.148)	-0.478*** (0.149)	-0.540** (0.202)	-16,420*** (5,390)	-18,058*** (5,477)	-21,314*** (7,228)
	Panel B: Quadratic Polynomial in Distance to Concession Border					
Inside Concession	-0.430*** (0.146)	-0.486*** (0.149)	-0.547** (0.206)	-16,465*** (5,351)	-18,285*** (5,485)	-22,166*** (7,287)
	Panel C: Quartic Polynomial in Distance to Concession Border					
Inside Concession	-0.424*** (0.145)	-0.474*** (0.153)	-0.533*** (0.186)	-16,216*** (5,336)	-18,072*** (5,599)	-20,722*** (6,241)
	Panel D: Interacted Linear Polynomial in Distance to Concession Border					
Inside Concession	-0.283 (0.225)	-0.421 (0.270)	-0.125 (0.439)	-14,978* (7,639)	-18,249** (8,962)	-7,084 (17,174)
	Panel E: Interacted Quadratic Polynomial in Distance to Concession Border					
Inside Concession	-0.470 (0.326)	-0.349 (0.462)	-0.487 (0.595)	-18,686* (10,968)	-17,783 (16,395)	-28,074 (20,010)
Observations	5,638	4,240	2,627	5,638	4,240	2,627
Clusters	109	84	52	109	84	52
Mean Dep. Var.	2.290	2.036	2.101	-29751	-46136	-43799

Note: We cluster standard errors at the DHS cluster level. We include district fixed effects and control for age, age squared and gender. *Wealth Factor* is an index generated by the DHS using principle component on asset ownership. *Wealth Index* is a is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Factor Score. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A3: Rubber Concessions and Health RD Analysis

Alternative RD Polynomials

	Child Ever Vaccinated			*Child Ht/Age Percentile*			*Respondent Ht/Age Percentile*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)	200 kms (7)	100 kms (8)	50 kms (9)
	Panel A: Linear Polynomial in Distance to Concession Border								
Inside Concession	-0.095***	-0.097***	-0.091*	-329.9*	-316.8	-456.2*	-694.198***	-783.215***	-718.914**
	(0.035)	(0.036)	(0.047)	(192.3)	(197.9)	(227.9)	(238.6)	(243.8)	(294.9)
	Panel B: Quadratic Polynomial in Distance to Concession Border								
Inside Concession	-0.095***	-0.094**	-0.091*	-329.7*	-311.2	-451.3*	-697.4***	-799.8***	-742.6**
	(0.035)	(0.036)	(0.046)	(192.3)	(197.6)	(226.7)	(240.7)	(247.2)	(288.0)
	Panel C: Quartic Polynomial in Distance to Concession Border								
Inside Concession	-0.093***	-0.094**	-0.093*	-348.8*	-345.3*	-409.8*	-748.6***	-836.6***	-739.3**
	(0.035)	(0.036)	(0.049)	(193.4)	(191.3)	(217.6)	(239.7)	(244.3)	(287.5)
	Panel D: Interacted Linear Polynomial in Distance to Concession Border								
Inside Concession	-0.101	-0.093	-0.146	-452.2*	-558.6*	60.5	-446.6	-572.1	264.9
	(0.062)	(0.075)	(0.095)	(260.6)	(294.0)	(382.9)	(346.2)	(382.9)	(443.4)
	Panel E: Interacted Quadratic Polynomial in Distance to Concession Border								
Inside Concession	-0.092	-0.164	-0.166	-58.0	-401.9	-568.8	-568.8	95.97	425.5
	(0.091)	(0.103)	(0.115)	(357.2)	(384.0)	(468.0)	(465.9)	(499.3)	(435.3)
Observations	3,161	2,533	1,627	1,596	1,305	822	1,578	1,207	758
Clusters	109	84	52	109	85	52	109	84	52
Mean Dep. Var.	0.816	0.799	0.793	2516	2453	2481	2693	2606	2628

Note: We cluster standard errors at the DHS cluster level. We include district fixed effects in all regressions. We control for age and age squared. The DHS health questions are only asked to a subset of female respondents. Respondent *Ht/Age Percentile* divides each respondent's height by their age and finds their percentile in the entire sample and normalizes this percentile to be within 0 and 10000. Similarly, Child *Ht/Age Percentile* divides each child's height by their age and finds their percentile in the entire sample and normalizes this percentile to be within 0 and 10000. *Child Ever Vaccinated* is an indicator variable equal to one if the respondent's child has ever been vaccinated. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A4: 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	-1.308***	-1.417***	-1.677***	-0.239***	-0.272***	-0.322***
	(0.340)	(0.364)	(0.420)	(0.070)	(0.072)	(0.083)
	Panel B: Quadratic Polynomial in Latitude and Longitude					
Inside Concession	-1.700***	-1.617***	-1.843***	-0.292***	-0.311***	-0.375***
	(0.400)	(0.416)	(0.448)	(0.084)	(0.094)	(0.096)
	Panel C: Cubic Polynomial in Latitude and Longitude					
Inside Concession	-1.743***	-1.747***	-1.706***	-0.306***	-0.352***	-0.385***
	(0.375)	(0.379)	(0.412)	(0.077)	(0.084)	(0.081)
Observations	5,629	4,233	2,623	5,633	4,237	2,626
Clusters	109	84	52	109	84	52
Mean Dep. Var.	5.648	5.130	5.209	1.173	1.068	1.077

Note: We cluster standard errors at the DHS cluster level. We include district fixed effects and control for age, age squared and gender. *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. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A5: Rubber Concessions and Wealth RD Analysis

Alternative RD Specification: Latitude and Longitude Polynomials

	Wealth Index			*Wealth Factor*		
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	-0.455***	-0.489***	-0.612***	-16,888***	-18,397***	-24,391***
	(0.143)	(0.148)	(0.214)	(5,139)	(5,303)	(7,497)
	Panel B: Quadratic Polynomial in Latitude and Longitude					
Inside Concession	-0.578***	-0.703***	-0.712***	-21,318***	-22,793***	-26,368***
	(0.177)	(0.196)	(0.218)	(6,341)	(6,784)	(8,258)
	Panel C: Cubic Polynomial in Latitude and Longitude					
Inside Concession	-0.676***	-0.736***	-0.530**	-24,209***	-23,808***	-18,366**
	(0.177)	(0.200)	(0.207)	(6,353)	(6,696)	(7,158)
Observations	5,638	4,240	2,627	5,638	4,240	2,627
Clusters	109	84	52	109	84	52
Mean Dep. Var.	2.290	2.036	2.101	-29751	-46136	-43799

Note: We cluster standard errors at the DHS cluster level. We include district fixed effects and control for age, age squared and gender. *Wealth Factor* is an index generated by the DHS using principle component on asset ownership. *Wealth Index* is a is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Factor Score. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A6: Rubber Concessions and Health RD Analysis

Alternative RD Specification: Latitude and Longitude Polynomials

	Child Ever Vaccinated			*Child Ht/Age Percentile*			*Respondent Ht/Age Percentile*		
Sample Within:	200 kms (1)	100 kms (2)	50 kms (3)	200 kms (4)	100 kms (5)	50 kms (6)	200 kms (7)	100 kms (8)	50 kms (9)
	Panel A: Linear Polynomial in Latitude and Longitude								
Inside Concession	-0.093***	-0.089**	-0.070	-167.2	-219.9	-308.2	-618.6**	-715.5***	-619.3*
	(0.035)	(0.036)	(0.045)	(181.3)	(188.3)	(256.4)	(260.3)	(265.8)	(332.2)
	Panel B: Quadratic Polynomial in Latitude and Longitude								
Inside Concession	-0.069	-0.044	-0.061	-430.2**	-570.2**	-447.5	-748.7***	-863.2***	-709.3**
	(0.047)	(0.056)	(0.062)	(209.2)	(221.9)	(277.7)	(280.9)	(297.6)	(351.1)
	Panel C: Cubic Polynomial in Latitude and Longitude								
Inside Concession	-0.067	-0.057	-0.091*	-378.5*	-501.4**	-276.7	-751.9***	-885.1***	-624.3*
	(0.043)	(0.052)	(0.053)	(209.1)	(205.3)	(210.1)	(282.5)	(293.7)	(322.8)
Observations	3,161	2,533	1,627	1,596	1,305	822	1,578	1,207	758
Clusters	109	84	52	109	85	52	109	84	52
Mean Dep. Var.	0.816	0.799	0.793	2516	2453	2481	2693	2606	2628

Note: We cluster standard errors at the DHS cluster level. We include district fixed effects in all regressions. We control for age and age squared. The DHS health questions are only asked to a subset of female respondents. Respondent *Ht/Age Percentile* divides each respondent's height by their age and finds their percentile in the entire sample and normalizes this percentile to be within 0 and 10000. Similarly, Child *Ht/Age Percentile* divides each child's height by their age and finds their percentile in the entire sample and normalizes this percentile to be within 0 and 10000. *Child Ever Vaccinated* is an indicator variable equal to one if the respondent's child has ever been vaccinated. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Appendix B. Robustness Tables

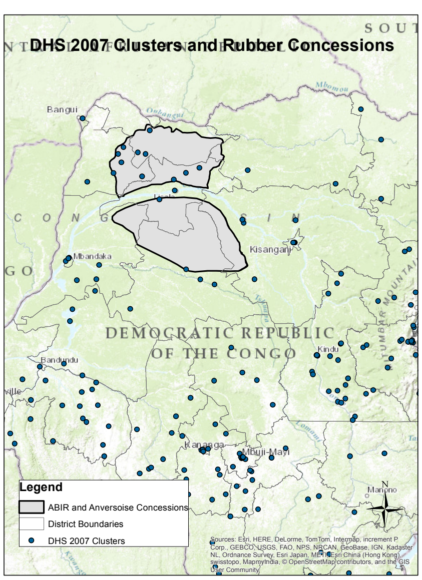
B.1. DHS Results - Varying the RD Polynomial and Specifications

B.1.1. Alternative RD Polynomials

B.1.2. Alternative RD Specifications: Latitude and Longitude Specification

Appendix C. Additional Maps and Tables

Figure 8: District and Territories Boundaries near the Rubber Concessions



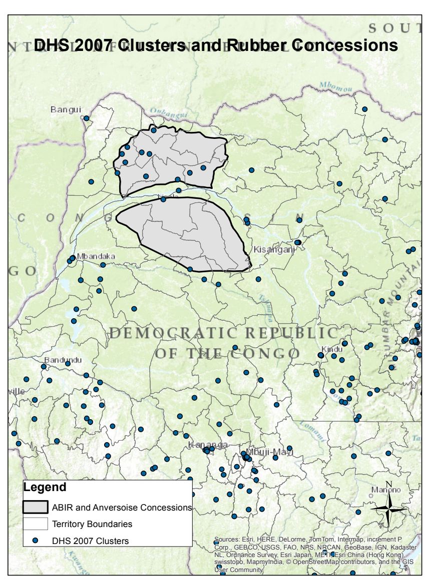
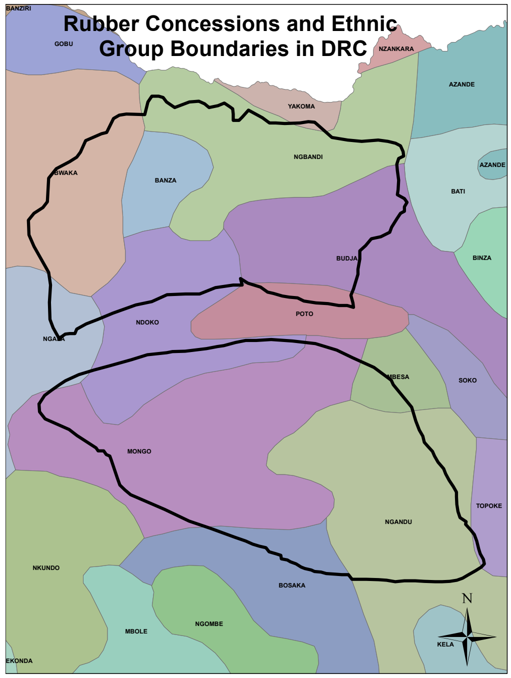
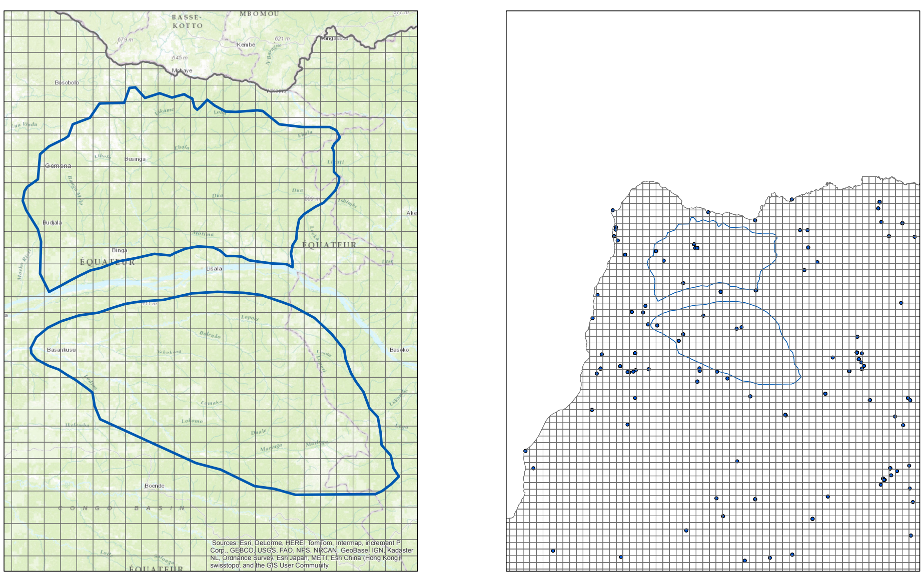


Figure 9: Murdock Map, Territories and Rubber Concessions



(a) Murdock Ethnic Group Borders and Rubber Concessions

Figure 10: Example of the 20km by 20km Grid



(a) 20 km by 20 km Grid Cell Example (Zoomed in)

(b) 20 km by 20 km Grid Cells and PRIO Conflict Events in Northern DRC

Table A7: Rubber Concessions and FAO Cereals Suitability RD Analysis

	Cereals Suitability		
Sample Within:	200 kms	100 kms	50 kms
	(1)	(2)	(3)
Inside Concession	0.196	0.115	-0.916
	(0.866)	(0.903)	(1.081)
	[1.204]	[1.221]	[1.456]
Observations	1,304	832	602
Mean Dep. Var.	35.234	35.901	36.311

Note: We use a cubic polynomial on distance to the concession borders. We include district fixed effects. The cereals suitability measure is the low-input, rain-fed cereal suitability from the FAO GAEZ database and ranges from 0 to 100, with higher values indicating higher cereal suitability. The regressions control for the percentage of each grid cell that is a river. We present robust standard errors in ( ) and Conley standard errors in [ ] (assuming a window of 40 kms). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$