The Long-Run Effects of the Scramble for Africa*

Stelios Michalopoulos†, Elias Papaioannou‡

January 2011

Abstract

We examine the economic consequences of the partitioning of Africa among European powers in the late 19th century; a process historically known as the scramble for Africa. First, using information on the spatial distribution of African ethnicities before colonization we establish that border drawing was largely arbitrary. Apart from the land mass and water area of an ethnicity's historical homeland, no other geographic, ecological, historical, and ethnic-specific trait predicts which ethnic groups have been partitioned by the national borders. Second, employing data on the location of civil conflicts after independence we show that compared to ethnicities that have not been impacted by the border design, partitioned ethnic groups have suffered significantly more, longer, and more devastating civil wars. Third, we find that economic development –as reflected by satellite data on light density at night- is systematically lower in the historical homeland of partitioned ethnicities. These results are robust to a rich set of controls at a fine level and the inclusion of country and ethnic-family fixed-effects. Our regressions thus identify a sizable causal negative effect of the scramble for Africa on comparative regional development.

Keywords: Africa, Borders, Ethnicities, Conflict, Development.

JEL classification Numbers: O10, O40, O43, N17, D74, Z10.

*We thank Gregorios Siourounis, Heiwai Tang and participants at the Econometric Society Meetings in Denver for useful comments. All errors are our sole responsibility.

†Tufts University, Braker Hall, 8 Upper Campus Rd., Medford, MA 02155. E-mail: stelios.michalopoulos@tufts.edu. The author was a Deutsche Bank Member at the Institute for Advanced Study, Princeton and gratefully acknowledges their hospitality.

‡Dartmouth College and CEPR and Harvard University. Address: Harvard University, Economics Department, Littauer 324, 1805 Cambridge Street, Cambridge, MA 02138, USA. E-mail: elias.papaioannou@dartmouth.edu.

1 Introduction

The predominant explanations on the deep roots of contemporary African underdevelopment are centered around the influence of Europeans during the colonial period (Acemoglu *et al.* (2001, 2002, 2005)), but also in the centuries before colonization when close to 20 million slaves were exported from Africa (Nunn (2008)). Yet in the period between the ending of the slave trades and the commencement of the colonial period, another major event took place in the European capitals that according to the African historiography had malicious long-lasting consequences. The "scramble for Africa" starts with the Berlin Conference of $1884-1885$ and is completed by the turn of the 20th century. In this brief period, Europeans partitioned Africa into spheres of influence, protectorates, colonies, and free-trade-areas. The borders designed in European capitals at a time when Europeans had barely settled in Africa and had little knowledge of the geography and ethnic composition of the areas whose borders were designing. These borders endured after the African independence in the 1960's leading to the partitioning of numerous ethnic groups across the newly created African states.[1] A considerable body of research in African history (e.g. Asiwaju (1985); Dowden (2008)) argues that the main impact of Europeans' influence in Africa was not colonization per se, but the improper border design. Partitioning, the argument goes, led to ethnic struggles, patronage politics, and spurred civil conflict, leading to poverty and underdevelopment. Yet there is very little work that formally examines the impact of ethnic partitioning.[2]

This paper is a first-step to empirically assess the long-run effects of the scramble for Africa on civil conflict and economic development. We identify partitioned groups projecting the current national borders on George Peter Murdock's Ethnolinguistic Atlas (1959) that portrays the spatial distribution of ethnicities before colonization (Figure 1*a*). We find that out of a total of 834 ethnicities, for 231 ethnic groups at least 10% of their historical homeland falls into more than one contemporary state (Figure 1*b*).[3] While there is little disagreement that colonial borders were arbitrarily drawn, we start our empirical analysis by establishing formally their randomness. We estimate probabilistic models trying to identify significant factors that predict whether a group was partitioned. With the sole exceptions of the size of

[1] Asiwaju (1985) identifies 177 partitioned ethnic groups that span all African borders. Moreover, Englebert et al. (2002) estimates that the population of partitioned ethnic groups is on average more than 40% of the total population. Likewise Alesina et al. (2011) estimate that in many African countries the percent of the population that belongs to a partitioned group exceeds 80% (e.g. Guinea-Bissau (80%); Guinea (88.4%); Eritrea (83%); Burundi (97.4%); Malawi (89%); Senegal (91%); Rwanda (100%); Zimbabwe (99%)).

[2] Notable exceptions are the cross-country studies of Alesina et al. (2011) and Englebert et al. (2002) whose relationship to our work we discuss below.

[3] When we consider all split ethnic groups irrespective of the degree of partitioning, we find that 358 groups have been affected by the border design. When we use a more restrictive threshold of 20%, we find 164 ethnicities to have been partitioned across the national borders.

the historical homeland and water bodies, we are unable to detect any significant differences between partitioned and non-partitioned ethnicities with respect to other geographical features, the disease environment, natural resources and the magnitude of slave raids. We further show that there are no major differences between split and non-partitioned groups across many precolonial ethnic-specific institutional, cultural, and economic features (Murdock (1967)).

Figure 1*a*

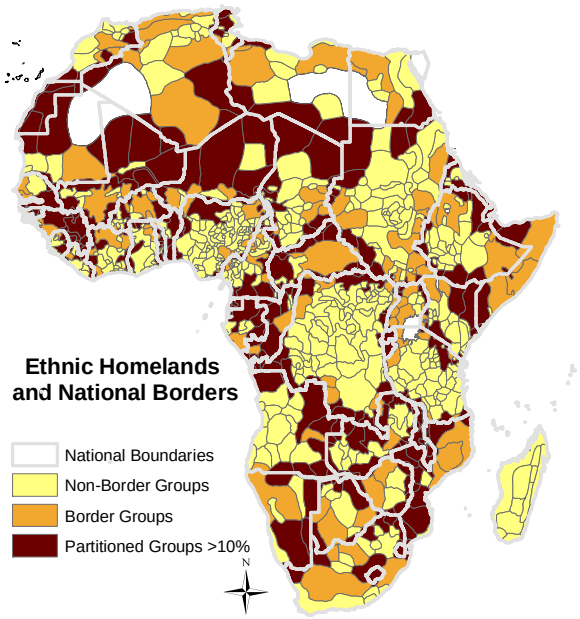


Figure 1*b*

We then employ the scramble for Africa as a quasi-natural experiment to assess the impact of partitioning on development and civil conflict, as this has been theorized to be the main channel of influence. Using regional data on the incidence, duration, and total casualties of all civil wars in Africa in the post-independence period ($1970-2005$), we find that civil conflict is concentrated in the historical homelands of partitioned ethnicities. We also find that border areas populated by ethnic groups only modestly affected by the artificial border design also experience more conflict, though this effect is of smaller magnitude and not always statistically significant. The positive effect of partitioning on all aspects of civil war retains its economic and statistical significance when we condition on country fixed-effects and ethnic-family fixed-effects, so as to control for national and broad ethnic characteristics respectively. Our most conservative estimates suggest that civil conflict intensity, as reflected in war casualties and duration, is approximately 35% higher in areas where partitioned ethnicities reside. We then examine the effect of partitioning on regional development. Due to data unavailability on income at our ethnic-level of analysis, we follow Henderson et al. (2009) and proxy regional development with satellite image data on light density at night. The cross-ethnicity estimates suggest that development in the historical homeland of partitioned ethnic groups is lower by

almost a half (−46%), compared to non-partitioned ethnic areas.

Historical Background

The scramble for Africa started in the late 19th century and was fully completed by the turn of the 20th century.[4] The event that stands for the partitioning of Africa is the conference that Otto von Bismarck organized in Berlin from November 1884 till February 1885. While the Berlin conference mainly discussed the boundaries of Central Africa (the Congo Free State), it came to symbolize the partitioning, because it laid down the principles that would be used among Europeans to divide the continent.[5] The key consideration was to prevent conflict among Europeans over the African territories (as the memories of the European wars of the 18th-19th century were still alive). In light of this, Europeans divided territories and drew borders in maps, without taking into account local geographic conditions and/or the ethnic composition. African leaders were not invited and had no say. Europeans were in such a rush that they didn't wait for the new information arriving from explorers, geographers, and missionaries.[6]

The anecdotal evidence suggests that the scramble for the continent was arbitrary. As the British prime minister at the time Lord Salisbury put it in a famous sally, "*we have been engaged in drawing lines upon maps where no white man's feet have ever tord; we have been giving away mountains and rivers and lakes to each other, only hindered by the small impediment that we never knew exactly where the mountains and rivers and lakes were.*" There is little disagreement among African scholars on the artificial border design. For example, Asiwaju (1985) argues that the "*study of European archives supports the accidental rather than a conspiratorial theory of the marking of African boundaries.*"[7]

[4] The scramble for Africa begins in 1860s-1870s when the French and the British start exploring systematically Western Africa and sign bilateral agreements assigning spheres of influence. The scramble was completed with the Franco-Spanish partition of Morocco and the annexation of Libya by Italy in 1912.

[5] The three major principles that emerged from the Berlin Conference were: First, the hinterland doctrine, according to which a power claiming the coast had also a right to its interior. Yet, the applicability of this principle became problematic, as it was not clear what exactly constitutes the hinterland. For example, at some point France demanded Nigeria claiming that it was the hinterland of Algeria. Second, the principle of effective possession required that Europeans need to base their claim on treaties with local tribe leaders. Yet, it was hard to assign zones of influence based on such treaties, because as Bismarck pointed out "*it was too easy to come by a piece of paper with a lot of Negro crosses at the bottom*" (Wesseling (1996)). Third, the effective occupation doctrine required that European powers exert significant control of the territory they were claiming. Yet, with the insistence of the British this principle was soon diminished to apply mostly in the coastline.

[6] For example, Asiwaju (1985) notes that "*the Berlin conference, despite its importance for the subsequent history of Africa, was essentially a European affair: there was no African representation, and African concerns were, if they mattered at all, completely marginal to the basic economic, strategic, and political interests of the negotiating European powers*".

[7] Likewise Hargreaves (1985) writes "*rather than attempting to follow the boundaries of states whose rulers might not be able to describe them accurately, the French preferred to allocate territory along some natural feature like a watershed. Yet, the problem was that the Europeans had a rather imperfect idea of where the water streams exactly where. A prominent example is the Anglo-German agreement on the Nigeria-Cameroon boundary that*

There are many reasons explaining the arbitrary border design: First, at the time Europeans had limited knowledge of local geographic conditions and were unwilling to wait. Second, Europeans were not drawing borders of prospective states or in many cases even colonies. For example, the Democratic Republic of Congo that corresponds to Congo Free State is so large simply because it was meant to be a free trade area rather than King Leopold's property or Belgian colony, not to say an independent state. Third, there was a constant imperialist back and forth with European powers swapping pieces of land with no idea what they were worth of. An illustrative case is the annexation of Katanga by the Congo Free State which turned out to be the richest and most important province. Leopold demanded and eventually got Katanga in exchange for the Niari-Kwilu area that the French insisted of having themselves (Wesseling (1996)). Fourth, while in most cases the treaties indicated that the exact boundaries would be set and demarcated by special commissions, demarcation was poor and the commissions did not alter much. Fifth, Europeans were not willing to sacrifice their commitment to not go to war for any part of Africa.[8] In many cases London and Paris turned down requests from local administrators to redraw the border because it did not coincide with a physical boundary or because an ethnic group was split. Sixth, there was an implicit agreement between Europeans that ethnicities could freely move across colonial borders. Asiwaju (1985) cites the Ketu king, saying that "*we regard the boundary (between Benin-Dahomey and Nigeria) separating the English and the French, not the Yoruba.*"

The other major event in African history, the wave of independence, was also rapid. The independence of Northern African countries in the 1950's was followed by Ghana's and Guinea' independence in 1957 and in 1958, respectively. By the end of 1966, 40 countries had gained independence. While at the time, many proposed changing the colonial borders, African leaders and leaving Europeans avoided this issue. The leaders of African independence believed that nation building would sideline ethnic divisions; moreover building new states and national institutions seemed more important than massive border realignment. Likewise Europeans' main objective was to maintain their special rights and corporate deals with their former colonies, and as such, they were reluctant to open the border issue. Thus almost all African countries accepted the colonial borders when signing the Charter of the Organization of African Union in 1964.[9] Moreover, the treaty for the formation of African Union explicitly mentioned that nations would allow their citizens moving across the border, so as to mitigate

was supposed to be Rio del Rey. The latter proved to be an estuary receiving several small streams."

[8] For example Wesseling (1996) writes "*in later years, Katanga was to become a most desirable possession in the eyes of British imperialists such as Cecil Rhodes and Harry Johnston. When they approached the British government on the subject, it stuck to its guns. Anderson let them know that Leopold's map had been recognized in 1885 and that his territory unmistakably comprised the mining region of Katanga. What was done, was done.*"

[9] Only Somalia and Morocco did not accept formally the colonial borders. Ghana and Togo raised also objections on their boundary that splits the Ewe.

the effect of ethnic partitioning.

Channels

The African historiography has put forward many explanations on how the partitioning of ethnicities and the creation of artificial states has contributed to underdevelopment.

First, in many instances partitioning has generated irredentist demands, as ethnicities that are minority groups in a country want to unify with their peers across the border. For example, Somali tribes were split between three different European colonies, while Ethiopia also got a slice. As a result, nowadays besides Somalia a large portion of Somalis occupy Northern Kenya, the Ogaden region in Ethiopia, as well as Eritrea and Djibouti. Three long-lasting wars in our sample have (partly at least) been driven by the desire of Somalis in Ethiopia, Djibouti, and Kenya to become part of Somalia.

Second, partitioned ethnicities may fight to gain independence or obtain autonomy.[10] An illustrative example is the recurring civil conflict in the Casamance region in Southern Senegal, where the partitioned ethnic groups Diola and the Malinke reside. As Gambia effectively splits Senegal into a Northern and Southern part, the Southern province of Casamance is disconnected from the central government in Dakar and has demanded independence.[11]

Third, partitioned ethnicities have reacted to their marginalization by participating in coups and rebellions to overthrow or capture the government. For example, the Ewe in Togo helped Flt.-Lt. Jerry Rawlings (half Ewe) in his coup in 1979 and 1981 to overthrow the government in Ghana. This escalated ethnic tensions between the Ewe, the Ashanti and the Akan in Ghana leading to civil warfare in the subsequent years.

Fourth, African borders are poorly demarcated and not well delineated due to the imprecise colonial treaties. This has resulted in border disputes, especially when such poorly demarcated borders cause the partitioning of ethnic groups.[12] The conflict between Mali and Burkina Faso over the Agacher Strip, where the Bobo reside, illustrates the problems caused by poor demarcation. The escalation of minor conflicts that started after independence resulted in a fully blown war in 1985.[13] Imprecise colonial treaties seem to have contributed to conflict

[10] Wimmer et al. (2009) estimate that around 20% of all civil wars in Africa have a secessionist demand. Other examples of secessions that have resulted in de facto autonomous and independently governed areas include the Western Sahara and the Somaliland (former British Somaliland).

[11] Renner (1985) writes "*Senegal itself became truncated, and could only be linked by traversing Gambia or by using the much lengthier overland route, The partition was undertaken (between the French and the British) without any consideration for cultural ties, economic viability or regional coherence.*"

[12] For example Englebert et al. (2002) write "*of all the territorial disputes brought before the International Court of Justice since 1960,* 57% *were African, while only* 33% *(*104 *out of* 315*) of all bilateral boundaries worldwide are in Africa.*"

[13] Eventually this dispute was settled in the International Court of Justice in the end of 1986. The court split the $3,000$ km of disputed territory almost equally between the two countries.

in Somalia (Higham (1985)), while the ambiguity of the tripartite treaty between Britain, Italy and Ethiopia of 1902 has also played a role in the Eritrea-Ethiopia war.

Fifth, Africa is characterized by patronage politics where dominant ethnic groups discriminate against minority groups (see i Miguel (2007) for a theoretical exposition and Wimmer et al. (2009) for empirical evidence). In many cases the central government tries to suffocate partitioned ethnicities, for example by seizing property and imposing higher taxation in the activities of specific groups (Bates (1981)). As a result, the neighboring country intervenes either to support their peers or to prevent migration and refugee flows. The conflict in the Alur-land exemplifies the case. The Alur were split between the Belgian Congo and the British Protectorate of Uganda during the late phase of the scramble for Africa ($1910 - 1914$). When the regime of Mobutu Sese Seko initiated the subjugation of many minority groups in Congo, a large portion of the Alur in Congo moved to their historical homeland in Uganda. This in turn generated opposition from the dominant Buganda group leading to civil conflict.

Sixth, due to their ethnic contacts across the border, partitioned ethnicities may engage in smuggling and other criminal activities. For example, in his analysis of the Anglo-French partitioning of the Sultanate of the Mandara in the Nigeria-Cameroon boundary, Barkindo (1985) writes that "*the most serious problem was the increase in crime and disputes across the border. The fact that the border divided people of the same family and settlements made it difficult to check crime and control smuggling.*" Collins (1985) also describes how smuggling allowed the Hausa to arbitrage price caps and other distortionary policies in Niger and Nigeria.

Seventh, border artificiality (though not partitioning itself) spurred conflict, because heterogeneous ethnic groups were forced to be part of the same usually large country. Many African scholars emphasize that civil conflict is more pervasive in large countries where it is hard for the state to broadcast political power and prevent secessionist movements among diverse ethnicities (e.g. Herbst (2000)). Indeed most long-lasting civil wars have taken place in the largest African countries, namely the Democratic Republic of Congo, Chad, Niger and Angola with Sudan being the most illustrative example. The ethnically, religiously, and racially distinct tribes of the North (that are part of the Nilo-Saharan families) and the South (that belong to the Afro-Asiatic family) resulted in a three-decade long civil war and an ongoing referendum for the independence of Southern Sudan.

Eighth, partitioning may lead to armed warfare by interacting with natural resources. If the historical homeland of a partitioned ethnic group is rich in natural resources then the benefit of secession increases; moreover in this case the central government is more likely to be oppressive. For example, armed conflict in the Cabinda enclave that is separated from the rest of Angola by a narrow strip of territory belonging to the Democratic Republic of the Congo

is driven by the interaction between the artificial border design, the vast oil fields, and the partitioning of the Bakongo people.

Related Literature

Our paper contributes to two main strands of the literature. First, our work relates to studies that aim to uncover the deep roots of African -and more broadly global- development. This literature has mainly focused on the impact of colonization mainly via early institutions (see for a review). Glaeser et al. (2004) and Easterly and Levine (2009) also stress the key role of colonization for long-run development, but emphasize the human capital channel. In contrast to this body of work, Gennaioli and Rainer (2006, 2007), Michalopoulos and Papaioannou (2010) and Nunn (2008) focus on the pre-colonial period stressing the role of pre-colonial ethnic institutions and the slave trades between the 15th-19th century respectively (see also Nunn and Puga (2011) and Nunn and Wantchekon (2011) on the latter). Of relevance to our work are also studies linking ethnic fragmentation/polarization with development (e.g. Easterly and Levine (1997); Alesina et al. (2003); Montalvo and Reynal-Querol (2005a); Alesina and Zhuravskaya (2011); see Alesina and Ferrara (2005) for a review).

Our paper contributes to this body of research, by emphasizing a somewhat neglected aspect of colonization; the drawing of political boundaries in the end of the 19th century that in the eve of African independence partitioned numerous ethnicities across the newly created African states. Thus our work is mostly related to Alesina et al. (2011), who show that "artificial states" defined as those with straight borders and those where a significant part of the population resides in more than one country, perform economically worse compared to countries with more organic (squiggly) borders. We focus on Africa, as the random design of colonial borders that endured after African independence allows us to identify the causal effect of partitioning. Moreover, besides reporting reduced-form estimates linking partitioning to under-development, we uncover the detrimental role of the border design in fomenting civil conflict. Our regional focus allows us to control at a very fine level for geography, the disease environment, natural resources, and other factors that a vast literature emphasizes as key determinants of economic development. Furthermore, we estimate country-fixed effects specifications to control for factors affecting economic development and civil war likelihood at the country-level (such as institutional quality, foreign aid, national policies, etc.).[14] We also condition on ethnic-family fixed-effects and thus account for broad cross-ethnicity differences in pre-colonial institutions, cultural traits, and economic well-being.

[14] Our within-country analysis is thus similar in spirit to recent works that assess the effect of institutions and historical features employing regional variation (e.g. Banerjee and Iyer (2005); Iyer (2010); Dell (2010); Huillery (2009); Acemoglu et al. (2008); Naritomi et al. (2009); Berger (2009); Arbesu (2011); Michalopoulos and Papaioannou (2010)).

Second, our work contributes to the literature on the origins of civil conflict (see Collier and Hoeffler (2007), Kalyvas (2007) and Blattman and Miguel (2010) for reviews and Collier and Sambanis (2005) for case study evidence on Africa). The literature has examined the role of many country characteristics, such as income, natural resources, colonization and ethnic composition on several aspects of civil conflict.[15] While the correlation between ethnic fragmentation and civil war is weak, recent studies document interesting cross-country correlations linking various aspects of the societal structure with armed conflict. Montalvo and Reynal-Querol (2005b) and Esteban et al. (2010) show a strong negative correlation between ethnolinguistic polarization and conflict. Wimmer et al. (2009) find that the likelihood of ethnic conflict increases when a large share of the population is excluded from power and when the ruling coalition consists of many ethnic groups. Matuszeski and Schneider (2006) document that the likelihood, duration, and intensity of civil wars is much higher in countries where ethnicities are clustered in specific areas within a country. The most closely related study is that of Englebert et al. (2002), who show a positive cross-country correlation between measures of suffocation and dismemberment and political violence, secession attempts, border disputes, and civil warfare.

The correlations found in studies exploiting cross-country variation are informative; yet due to a host of concerns, they cannot be causally interpreted (see Blattman and Miguel (2010)). The main reason is that the process of border drawing is ultimately related to the process of state formation and is thus associated invariably with both voluntary and forced movements of people. Our study accounts for some of the shortcomings of cross-country works. First, it establishes that African borders are random by showing that there are no systematic differences in numerous geographic, economic, institutional, and cultural characteristics between partitioned and non-partitioned groups. Second, the use of information on the spatial distribution of ethnicities in the end of 19th century well before the current national boundaries came into effect alleviates concerns related to the movement of people. Third, our micro approach allows us to condition on country fixed-factors and ethnic-family factors.[16] Finally, we are also well positioned to consider local factors that may affect civil conflict.[17]

[15] See among others Collier and Hoeffler (1998), Collier et al. (2004), and Fearon and Laitin (2003) for cross-country evidence.

[16] As the sensitivity analysis of Herge and Sambanis (2006)) shows the estimates of the cross-country works are quite sensitive to the employed set of conditioning variables. Since there are at most 180 country observations and potentially numerous variables affecting civil conflict, it is eminently difficult to isolate the effect of a single variable utilizing solely the cross-country dimension. Moreover, because the key correlates of civil conflict are themselves highly correlated (e.g. poverty, corruption, and fragmentation go in tandem) and contain non-negligible measurement error the estimates are plagued with multi-colinearity (see Levine and Renelt (1992), Martin et al. (2004), and Ciccone and Jarocinski (2010) for a discussion of these issues in the similar context of cross-country growth regressions.). The regional approach gives us more degrees of freedom and multi-colinearity concerns are mitigated because the main explanatory variables at this level of aggregation are not collinear.

[17] Buhaug and Rod (2006) study the local determinants of civil war in Africa using as a unit of analysis 100

Structure

In the next section we present the regional data on civil conflict and satellite light density. We also discuss some issues on estimation and inference. In Section 3 we establish the arbitrariness of the African border design. Section 4 reports the results from our analysis on the effect of partitioning on various aspects of civil conflict (number of war incidents, casualties, and duration). Section 5 gives our results on the effect of partitioning on regional economic development. Section 6 concludes.

2 Data and Empirical Methodology

2.1 Civil War Data

The main data source for the occurrence and duration of civil wars comes from the Uppsala Conflict Data Program (UCDP)/International Peace Research Institute, Oslo (PRIO) Armed Conflict Dataset, Version 4-2006, initially assembled by Petter et al. (2002) and extended to cover the period $1946-2005$.[18] We limit our analysis to armed conflicts that started after 1970 when the majority of African states had gained independence.[19] Following the literature on civil war we focus on conflicts between the government of a state and one or more internal opposition group(s) without intervention from other states (intrastate) and internationalized intrastate conflicts between the government of a state and one or more internal opposition group(s) with intervention from other states (secondary parties). According to the UCDP/PRIO dataset since 1970 in Africa there have been in total 49 civil wars, 7 are classified as internationalized internal (e.g. the civil war of 1990 in Rwanda when Congo intervened and the civil conflict in the late 1990s in Guinea-Bissau where Guinea and Senegal intervened), with the overwhelming majority (42) classified as internal armed conflicts.

We use the dataset of Raleigh et al. (2006) to obtain information on the spatial extent of each civil war. This dataset assigns to each conflict a centroid with a corresponding radius in kilometers. The coordinates represent general estimates of where battles have occurred whereas the radius indicates the largest geographic extent of the conflict zone.[20] The location of several

kilometer by 100 kilometer grids. While they do not control for country characteristics or fixed-effects, they find that conflicts are more likely to occur far from the capital, near national borders, and close to regions with diamond mines and petroleum fields.

[18] Armed Conflict is defined as "*a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths.*" A minimum of 25 battle-related deaths per year and per dyad is required. Government is the party controlling the capital of a state.

[19] In particular, we are considering conflicts that are classified with a start date as early as 1970 where the date represents when the conflict for the first time reached 25-battle-related deaths in a calendar year. The results are similar if we include the 1960's or limit our attention in the period $1980-2005$.

[20] There are limitations with respect to the georeferencing conflict. For example, the authors note that at a given point in time, the actual conflict zone might be more constrained than the maximum that is recorded.

conflicts has changed over time. For example, the long-lasting Liberian civil war is coded as having affected 3 different conflict zones.[21] Overall, the 49 African civil wars between 1970 and 2005 played out in 77 conflict zones.

We also construct tribe-specific measure of combat deaths with data from Lacina and Gleditsch (2005). This dataset reports for each conflict-year a low, a high and a best estimate of the number of civilians and combatants killed in the course of combat.[22] Using the event identifier we merge the three datasets, so as to examine the effect of partitioning on various measures of civil war.

2.1.1 Civil Conflict Incidence

Based on Raleigh et al. (2006) we generate a map depicting each conflict zone associated with a civil war from 1970 to 2005. Then we project the constructed map on top of Murdock's ethnolinguistic map (Figure 1a). This allows us to identify in a systematic way which ethnicities have been affected by civil conflict during this period. We construct two alternative indexes of civil war incidence at the ethnicity level. The first measure captures the number of civil wars.[23] If the civil war changed location over time, we combine all the conflict zones. Hence, the index capturing the number of civil wars is a coarse measure of civil war intensity, because it does not take into account that the zone of a conflict may migrate over time. Thus, we also calculate for each tribal area the number of conflict zones. Figures 2a-2b portray the spatial distribution of civil war incidents at our ethnic level of analysis.

Furthermore, the authors define a circular zone of conflict whereas the actual shape is more likely to follow the contours of international boundaries, mountains, rivers, etc. To mitigate such issues in the empirical analysis several indexes of war intensity are constructed by taking into account the size of the tribal area affected by armed conflict.

[21] Namely in 1980 the conflict zone centroid had the following coordinates: $latitude = 6.32$ and $longitude = -10.8$ and a radius of 50 kilometers. From 1989 to 1995 the conflict zone's centroid moved to $latitude = 6$ and $longitude = -10$ with a radius of 300 kilometers. The final part of the conflict between 2000 and 2003 was centred on $latitude = 7.5$ and $longitude = -10.5$ with a radius of 150 km.

[22] We use the best estimate death measure as our benchmark number for battle fatalities; yet we also explore the sensitivity of our results to the low and high estimate. For two incidents the best estimate is unavailable. In this case we replace it with the average of the high and low estimates.

[23] Murdock's map includes 843 tribal areas (the mapped ethnicities correspond roughly to levels $7-8$ of the Ethnologue's language family tree); yet 8 areas are classified as uninhabited upon colonization and are therefore not considered in our analysis. Also, we eliminate the Guanche, a small group in the Madeira islands that is currently part of Portugal.

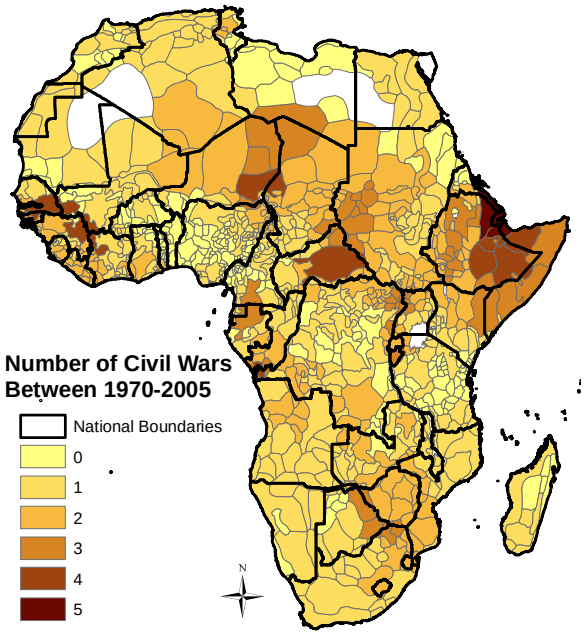


Figure 2*a*

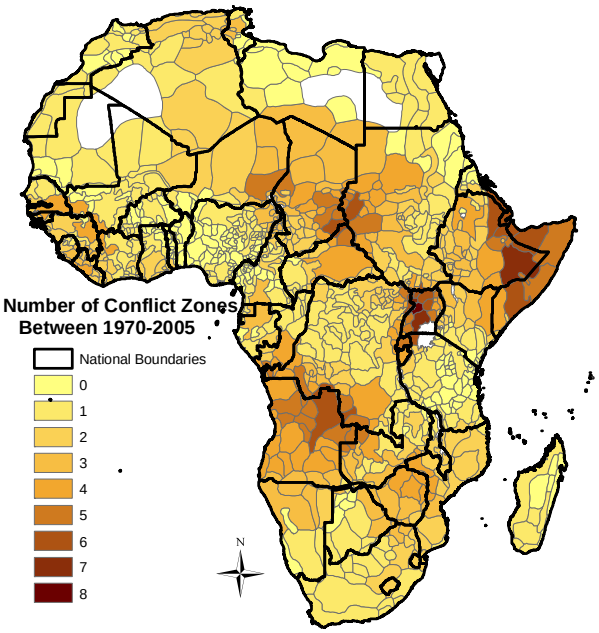


Figure 2*b*

2.1.2 Civil Conflict Casualties

We construct ethnic-specific estimates of civil war casualties with the following procedure. First, we calculate for each conflict zone the total number of casualties by summing up the battle related deaths across years. (For example in the case of the civil war in Sierra Leone (event ID 187) the total battle fatalities between 1991 and 2000 sum up to $12,997$). Second, we calculate the overall area that a conflict zone extends to. (The Sierra Leone civil war affected a total area of $54,287$ $km2$). Third, we estimate for each ethnic group how much of its homeland has been affected. (The Sierra Leone civil war involved $18,770$ $km2$ out of $22,946$ of the Mende's historical homeland. Thus, 35% of the civil war in Sierra Leone has taken place within the Mende's territory). Fourth, assuming that total casualties are distributed uniformly across the total area of the respective conflict zone, we derive ethnic-specific casualties by multiplying the total number of casualties with the fraction of each ethnicity's territory that falls within the conflict zone. (This implies that 35% of the $12,997$ battle deaths of the Sierra Leone civil war ($4,497$ casualties) have taken place in Mende's homeland). Fifth, we repeat this calculation for each conflict zone. For example, Mende's homeland takes up 29% of zone 1 of civil war in Liberia (event ID 146), that has caused a total of $4,058$ battle deaths; this translates into $1,175$ casualties to the Mende's historical homeland. Similarly, conflict zone 2 of the Liberian civil war (event ID 146) adds an additional 1011 casualties bringing the total number of civil war

fatalities of the Mende to 6,680 individuals.[24] Since larger tribes are more likely to be affected by conflict, we normalize battle deaths by the area of each group. The Mende, for example, have a casualty rate of 291 casualties per thousand of square kilometers. Figure 3*a* (below) plots civil conflict casualties for the 834 ethnic areas in our sample.

2.1.3 Conflict Duration

Analogously we obtain tribe-specific estimates on civil war duration. First, we calculate for each ethnic area the fraction of land that has been affected by a civil war. (In the case of the Mende, for example, 81% of its homeland has been affected by the Sierra Leone civil war (event ID 187) whereas 78% and 85% of its territory has been affected by zone 1 and zone 2 of the Liberian civil war (event ID 146) respectively). Second, we weigh the duration of each conflict with the fraction of the tribal area involved. (Since the civil war in Sierra Leone lasted for 10 years (1991-2000), the effective duration for Mende's homeland is $10 * 0.81 = 8.1$ years). Third, we sum the effective duration across all conflicts, so as to derive an estimate of civil war duration for each ethnic group. (Doing so shows that Mende were under civil war for 17.04 years). Figure 3*b* graphs the duration of civil conflict across ethnic groups.

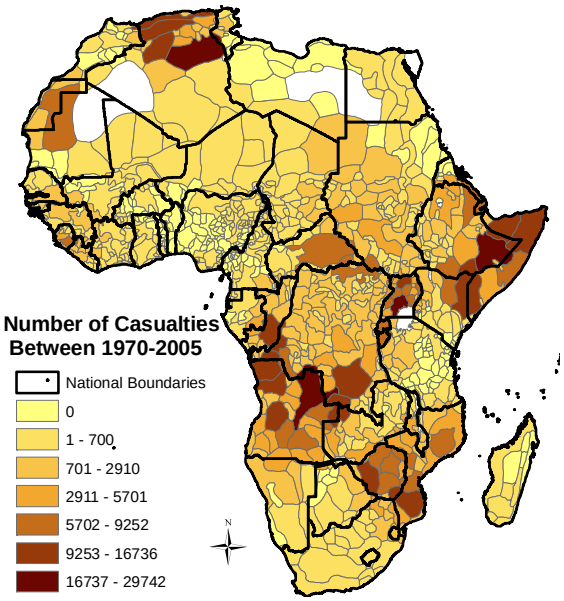


Figure 3*a*

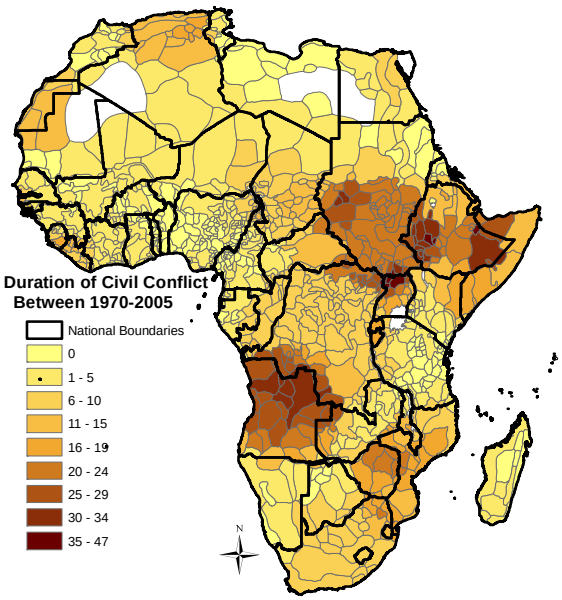


Figure 3*b*

2.2 Development Data

Given data unavailability on income or other development proxies for a large number of African countries at our ethnic level of analysis, we follow Henderson et al. (2009) and use satellite

[24] See Appendix Figures 1*a* and 1*b* portraying the conflict incidents on Mende's homeland.

data on light density at night to measure economic activity. Data come from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) that reports images of the earth at night captured from 20:00 to 21:30 local time. The measure is a six-bit $(0-63)$ digital number calculated for every 30-second area pixel (approximately 1 square kilometer), which is averaged across overlapping raw input pixels on all evenings in a year. To construct light density at the desired level of aggregation we average the digital number of luminosity across pixels that fall within the respective boundaries of an ethnic group (results are similar if we use the median value of light intensity). Using these data we construct average light density per square kilometer between 2007 and 2008.

Chen and Nordhaus (2010) argue that luminosity data can be quite useful for regional analyses in war prone countries and economies with data statistics of poor quality. Henderson et al. (2009) further show that satellite light density captures well abrupt changes in economic activity at a local scale in Africa (as for example during the Rwandan genocide or when large deposits of rubies and sapphires were accidently discovered in Madagascar). In the same vein, Michalopoulos and Papaioannou (2010) show a strong correlation between light density at night and GDP per capita across African countries as well as a significant within country negative correlation between light density and infant mortality. Figure 4 reports light density in $2007-2008$ for the 834 ethnic areas.

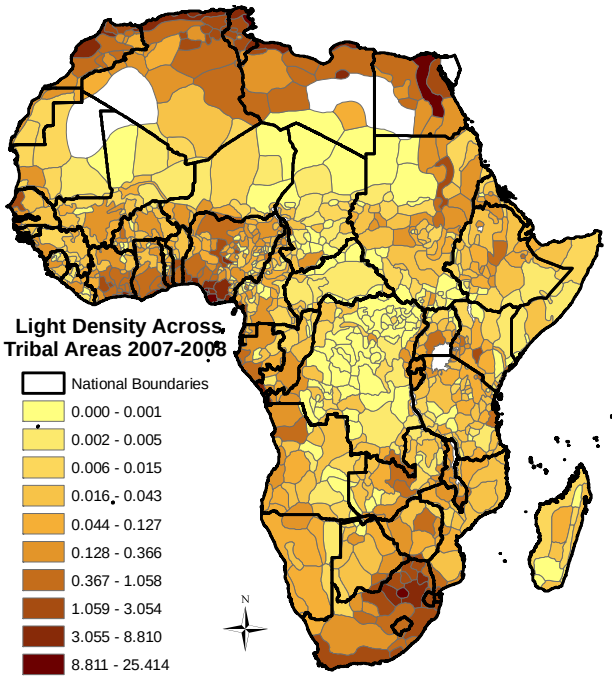


Figure 4

2.3 Preliminary Descriptive Evidence

Table 1*A* presents descriptive statistics of the civil war measures. There is large variation in the incidence of civil war. Out of the 834 tribes, 343 ethnicities suffered one civil war, but a non-negligible amount of 199 ethnicities experienced two distinct armed conflicts; 54 ethnicities experienced 3 civil wars, while 12 tribes were affected by four or even five conflicts. The groups with the highest incidence of civil war are the Afar and the Esa which during $1970-2005$ experienced 5 civil wars. Both groups have been greatly impacted by the border design with Afar being partitioned between Ethiopia, Eritrea and Djibouti and the Esa located in the border of Ethiopia and Somalia. The number of conflict zones also varies considerably (from 0 to 8). Not surprisingly, there is a strong correlation, 0.85, between civil war occurrences and conflict zones per ethnic area (Table 1*B*). The number of casualties per thousand of square kilometers is an indicator of how destructive civil conflict has been. The average casualty rate is 40 fatalities. The average rate however, masks a great deal of heterogeneity. In particular, 50% of tribal areas experienced less than 3 casualties per thousand kilometers. Ethnic groups like the Wanga and the Sabei partitioned between Kenya and Uganda, have been involved only in one civil war between 1970 and 2005 however they have lost thousands of people during this war. Conflict duration varies from 1 year (which is the case for 27 conflict zones) to 22 years (corresponding to the long-lasting civil war in Sudan (1983-2004)). There are tribal areas that have experienced incessant warfare with the case of the Alur group (partitioned between Zaire and Uganda) registering the highest duration of warfare across African tribes.[25] The modest correlation between the various measures of civil war (reported in Table 1*B*) suggests that each measure captures distinct aspects of civil war.

Before presenting our econometric framework we compare the means (results are the same with the medians) of the various measures of civil wars and development across different groups of ethnicities. In Table 1*C* we classify ethnicities in 3 groups, namely non-border tribes, border tribes with more than 90% of their homeland belonging to a single country, and partitioned ethnicities, defined as those with less than 90% of their territory found within a single country (using alternative thresholds for defining a partitioned group delivers similar results). Both border and partitioned tribes are systematically different compared to non-border groups. Border and partitioned groups have experienced more civil wars, have suffered higher casualty rates, and have seen their homelands under civil conflict for longer periods. Furthermore, economic performance is on average 60% lower.

[25] Note that because of how the duration variable is constructed ethnic groups whose homeland has been involved in overlapping conflict zones may have a duration of civil war larger than 35 years. This is the case for 7 ethnic groups.

2.4 Econometric Specification

We estimate the long-run effect of the scramble for Africa running variants of the following empirical specification

$$y_i^{f,c} = a_0 + \gamma SPLIT_i + \delta BORDER_i + X_i'\Phi + f(LOCUS_i) + a_c + a_f + \varepsilon_i^{f,c}. \quad (1)$$

The dependent variable, $y_{i,f,c}$, reflects civil conflict and economic development in the historical homeland of ethnic group i. Ethnicities are part of larger ethnic families (clusters) f and countries c.[26] $SPLIT$ is an indicator that takes the value 1 when a tribal area is partitioned by the political boundaries. Thus the coefficient γ captures the direct effect of partitioning. To identify partitioned ethnicities in our benchmark estimates we require that at least 10% of the historical homeland of each ethnic group falls into more than one countries.[27] $BORDER$ is an indicator that equals one when the historical homeland of an ethnic group is adjacent to the national border, but more than 90% of the land area of this group falls in the same country. We include the $BORDER$ variable for two reasons. First, to account for potential measurement error in identifying partitioned ethnic groups based solely on Murdock's map using the 10% threshold. Second, to capture potential spillover effects from areas where partitioned ethnicities reside ($SPLIT = 1$) to neighboring border regions ($BORDER = 1$). We return to the issue of spillovers below. Appendix Table 1 gives the list of all partitioned ethnic groups. Vector X_i' includes geographical controls, like land area and elevation; ecological features, such as a malaria stability index and land's suitability for agriculture; natural resources as proxied by the presence of diamond mines and petroleum fields; pre-colonial ethnic traits related to institutional structure and economic development. The Data Appendix gives detailed variable definitions and sources.

To capture unobservable characteristics and spatial effects in many specifications we control for smooth functions of geographic location by introducing a cubic polynomial ($f(LOCUS_{i,c})$) in the latitude and longitude of the centroid of each ethnic group.[28] In many specification we include country fixed-effects (a_c) and ethnic family fixed-effects (a_f). Country constants capture nation-wide factors related to post-independence policies, contemporary institutions, ethnolinguistic diversity and development that a vast literature on civil warfare identifies as significant

[26] Partitioned ethnicities are assigned to the country where the centroid of the historical homeland lies.

[27] We also experiment with other thresholds to identify partitioned ethnicities (5%, 20%, and 30%). The results are similar.

[28] Letting x denote latitude and y denote longitude the polynomial becomes $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. As Dell (2010) notices this parameterization is transparent and quite flexible. As there might be some concerns of overfitting, we report in all tables results without the polynomial. We also estimated specifications adding a cubic polynomial in distance to sea coast, finding similar results (not reported).

correlates of armed conflict. The ethnic family fixed-effects capture broad cultural, institutional, and other hard-to-observe and measure ethnic-specific factors (Murdock assigns the 834 groups into 96 ethnolinguistic clusters/families).

A notable fraction of the observations on civil warfare (number of civil wars, number of war zones, battle casualties and duration) and our proxy for regional development (light density) takes on the value of zero. Moreover the civil war measures and the satellite light density variable are highly skewed as we have many observations close to zero and a few extreme observations in the right tail of the distribution. (See Table $1A$). Hence, we estimate non-linear ML specifications with a Poisson and/or a Negative Binomial process (Wooldridge (2002)). The non-linear estimators are appealing, because they do not require log-linearizing the dependent variable and thus preserve the higher moments of the distribution, (see Silva and Tenreyro (2006) and Silva et al. (2010)). Yet we also report log-linear LS specifications taking the log of one plus the dependent variable.

In all specifications we account for spatially correlated residuals ($\varepsilon_i^{f,c}$) clustering standard errors at the country-level and at the ethnic-family level using the multi-way method of Miller et al. (2006). This correction also accounts for arbitrary residual correlation within each country and each ethnic family. We also estimated standard errors using Conley's GMM method to account for spatial dependence of an unknown form, finding similar results.

3 Are African Borders Artificial?

While the anecdotal and historical evidence point out that Europeans did not take into account local conditions and ethnic characteristics during the scramble for Africa, it is necessary to investigate whether there are systemic differences between partitioned ethnic groups, other border groups, and non-partitioned ethnicities.[29] We test whether there are geographic, economic, institutional, and other tribe-specific factors that predict whether an ethnic group is partitioned by the national borders, estimating probabilistic models of the following form:

$$SPLIT_i = a + X_i'\Psi + Z_i'\Theta + e_i \tag{2}$$

The dependent variable ($SPLIT_i$) equals one when at least 10% of the historical homeland of an ethnic group has been partitioned into more than one contemporary states. We explore the robustness of our results using a broader index that takes the value one when the

[29] Examining formally whether there are systematic differences on observable characteristics between partitioned and non-split ethnic groups is necessary, because in some cases Europeans did try taking into account local conditions (as for example when German West Africa was split into Urundi and Rwanda). Likewise in two cases (Cameroon-Nigeria; Ghana-Togo) there were referenda on the redrawing of borders at independence. Moreover we had the secession of Eritrea from Ethiopia (in 1993) and the unification of Tanganyika and Zanzibar (in 1964).

historical homeland of an ethnicity falls into more than one country irrespective of the size of the partitioning ($BROAD_i = SPLIT_i + BORDER_i$). X_i is a vector of geographic, ecological, and economic features at the ethnicity level; Z_i is a vector of ethnicity-specific pre-colonial institutional, cultural, and early development traits, extracted from Murdock (1967). Tables 2 and 3 present the results. The tables report probit marginal effects estimated at the mean of the other regressors and double-clustered at the country and the ethnic- family level standard errors. (The results are similar when we estimate linear probability or logit models).

In Table 2 we examine the role of geographical, ecological, and demographic features in the full sample (834 observations). In odd-numbered specifications the dependent variable is the $SPLIT_i$ index, while in even-numbered columns the dependent variable is the broad partitioning index, $BROAD_i$. The results in columns (1)-(2) show that ethnic groups spanning large territories in the pre-colonial period are more likely to be partitioned. This finding is in line with the historical evidence that colonizers drew arbitrary lines across imprecise maps. The estimates further show that ethnicities residing in areas with large water bodies (lakes and rivers) were more likely to find themselves split by national boundaries. This result is in line with the historical narrative that Europeans tried to use some natural barriers while delineating spheres of influence. Elevation, distance from the sea, land's agricultural quality, and malaria prevalence are not significant predictors of partitioning. In columns (3)-(4) we include region fixed effects to account for the somewhat different patterns of border design across the continent. The results remain unchanged.

In columns (5)-(6) we explore whether colonial powers took into account pre-existing development and the slave trades that preceded the scramble for Africa when designing the borders. Following Nunn and Watchenkon (2010) we proxy the pre-slave trade level of economic development using an indicator variable that equals one when a city with population exceeding $20,000$ people in 1400 AD was present in the historical homeland of an ethnicity. We also include a measure of ethnic-specific slave trades between the 16th-18th century using data from Nunn (2008). Neither the city in 1400 indicator nor the log number of slaves exported during the slave trades affect which ethnicities have been partitioned. The size of the historical homeland of each ethnic group and the amount of its water streams continue to be the only significant correlates of partitioning. The specifications in columns (5)-(6) also include indicators identifying ethnic areas with diamond mines and petroleum fields. While in the initial phase of colonization Europeans were mostly interested in agricultural goods and minerals, adding these two indicators allows us to investigate whether partitioned ethnicities differ from non-partitioned ethnic groups in terms of natural resources. This is important as one of the most robust correlates of civil warfare across countries is oil and other underlying

hydrocarbon deposits, as well as diamonds (e.g. Ross (2006); Herge and Sambanis (2006)). Again there is no systematic differences in natural resources among split and non-split ethnic groups. In columns (7)-(8) we add log population density around African independence (in 1960). In a Malthusian regime where richer areas are more densely populated, the insignificant estimate on log population density in 1960 implies that there were no systematic differences in economic performance and urbanization between partitioned and non-partitioned ethnicities. Hence, any contemporary divergence in economic performance between these two groups cannot be attributed to differences in economic conditions in the eve of African independence.

In Table 3 we examine whether ethnic-specific pre-colonial institutional, economic, and cultural traits affect which ethnic groups have been partitioned. The sample size drops because Murdock (1967) does not provide information for all the ethnicities in his Ethnographic Atlas (1959). In all specifications we include the region fixed-effects and the two geographical measures, size and land area under water, that were found to be significant predictors of partitioning. In columns (1)-(2) we investigate whether Europeans took into the account the degree of political centralization of the African ethnicities when designing the borders. Following Gennaioli and Rainer (2006, 2007), we proxy political centralization with an indicator variable that equals zero when Murdock assigns an ethnicity either as "*stateless*" or "*a petty chiefdom*" (e.g. Xam or the Ibo); and becomes 1 when the ethnicity is part of either a "*large paramount chiefdom*" or a "*large state*" (e.g. Thonga and Zulu). The political centralization enters with a negative coefficient, but the estimate is statistically indistinguishable from zero. In columns (3)-(6) we examine whether there were systematic differences in the degree of property rights protection between ethnicities that were partitioned and those that were not. In (3)-(4) we use Murdock's class stratification index and define a dummy variable that equals zero when the society is egalitarian "*without significant class distinctions*" (e.g. Fang in Nigeria or the Kikuyu in Kenya). The variable equals one when there are "*wealth or elite distinctions, a hereditary aristocracy is present, or there is complex class differentiation*" (e.g. the Yoruba in Nigeria or the Shilluk in Sudan).[30] In columns (5)-(6) we follow Fenske (2010) and proxy the presence of property rights institutions with a dummy variable that equals one when the society has some form of inheritance rules for property (e.g. the Ewe in Togo and Ghana or the Soga in Uganda) and zero otherwise (e.g. the Fang in Gabon or the Namshi in Cameroon). In all permutations neither proxy of pre-colonial property rights institutions enters with statistically significant coefficient. African scholars (e.g. Hopkins (1973); Austin (2008)) argue that economic and institutional development was higher in areas with intensive use of agriculture (Fenske (2010) provides empirical evidence supportive to this conjecture); thus in columns (7)-(8) we augment

[30] This transformation follows Gennaioli and Rainer (2006, 2007). The results are similar if in turn we use the original $0-4$ class stratification index.

the specification with a $0-10$ index measuring the importance of agricultural production at the ethnicity level, failing again to detect a significant effect. In columns (9)-(10) we include all the pre-colonial institutional, cultural, and economic characteristics in the specification. Again we are not able to find any significant variables affecting the likelihood of partitioning.

The results reported in tables 2 and 3 strongly support the notion of the arbitrary design of African borders. Overall Europeans did not take into consideration local geography, ethnic-specific institutional, cultural and economic features during the scramble for Africa. Out of dozens of potentially relevant variables, only land size and the presence of water streams correlate with partitioning. Moreover, the explanatory power of the models is poor. Mc Fadden's pseudo-R^2 (that compares the log likelihood value of the constant-only model with that of the full specification) is low across all permutations, at most 0.10. Likewise the R^2 of the linear probability models (not reported) is always below 0.12. The probit specifications perform quite poorly in predicting which ethnicities have been partitioned. For example, the full specifications in columns (7) and (8) of Table 2 predict correctly ($G(X_i'\Psi) > 0.5$) only 30 out of the 231 partitions with the benchmark index ($SPLIT$) and 164 out of the 358 partitions when we use the broad split index ($BROAD$).

4 Partitioning and Civil Conflict

4.1 Civil Conflict Incidence

In Table 4 we examine the effect of partitioning on the incidence on civil conflict. Both panels report Poisson ML estimates where the dependent variable equals the number of civil conflicts (Panel A) and the number of conflict zones (Panel B).

Specification (1) shows that, conditional on region fixed-effects, population density at independence, size and area under water (the only geographic variables that are correlated with partitioning) partitioned ethnicities are significantly more prone to civil conflict. Adding the cubic polynomial in latitude and longitude (column (2)) has little effect on the estimate. Column (3) includes a rich set of controls, reflecting local geography (land suitability for agriculture, elevation, distance to the sea coast), current and past urbanization (dummy variable that takes on the value one when a current capital city belongs to the historical homeland of an ethnic group and an analogous indicator that equals one if a major city was in the historical homeland in 1400 AD) and natural resources (indicators of a diamond mine or an oil deposit). Accounting for these factors seems a priori important, because the cross-country literature documents significant correlations between many of these variables and various aspects of civil warfare. For example Fearon and Laitin (2003) find that there is a higher likelihood of civil conflict in mountainous countries. Likewise both cross-country works (e.g. Ross (2006)) and

regional studies (e.g. Buhaug and Rød (2006); Bellows and Miguel (2009)) show that conflict is higher in areas with diamond mines and petroleum fields. Yet adding these controls has little impact on the coefficient on the partitioning index.[31] In column (4) we add ethnic family fixed-effects to account for hard-to-measure differences across African groups. The coefficient on the partitioning variable falls somewhat, though it retains significance at the 99% confidence level. In column (5) we add country fixed-effects.[32] The estimate on the partitioning dummy is significant at the 99%. The coefficient implies that partitioned ethnicities experience an increase of approximately 0.25 log points in the number of conflicts. This translates into a 28% higher increase in civil conflict activity ($\exp(0.25) - 1 = 0.28$) in areas where partitioned ethnicities reside.[33]

The estimates further show that there is a higher degree of civil war occurrence in border areas populated by ethnic groups that were affected by the border design, but to a smaller degree. Yet the economic effect implied by the coefficient of the $BORDER$ index is much smaller than the coefficient of $SPLIT$. Moreover, the coefficient on $BORDER$ becomes insignificant when we condition on the rich set of controls and/or fixed-effects.

4.2 War Casualties

In Table 5 we study the effect of the artificial border design on the number of casualties. Panel A reports negative binomial (NB) ML estimates, where the dependent variable is the number of casualties standardized by land area. Due to overdispersion in the dependent variable, specification tests suggest that the negative binomial model is preferable to the Poisson model. To illustrate the robustness of our results, Panel B reports analogous LS estimates, where the dependent variable is the log of one plus the number of casualties per thousand of square kilometers.

The results in column (1) show that casualties from armed conflict have been much higher in areas where partitioned ethnic groups reside; casualties are also higher in other border regions. The estimates of the negative binomial specification imply that casualties are by 83.5%

[31] In all specifications the capital city dummy variable enters with a positive and significant estimate indicating that there is a higher intensity of civil warfare in the capitals. The natural resource measures enter with positive estimates that in most (though not all) specifications are significant. There is also some evidence that civil conflict is higher in more mountainous regions.

[32] When we add country fixed-effects we lose variation from Rwanda, Swaziland, Burundi and the Comoros, since in these countries we have just one ethnic group.

[33] Ordered probit estimation yields similar results. The coefficient on the $SPLIT$ index is highly significant (t-stat higher than 4 in all permutations). In all ordered probit specifications the coefficient on $SPLIT$ is twice as large as the coefficient on the $BORDER$ index. As our Poisson results in columns (1)-(4) the estimate on $BORDER$ is statistically significant at standard confidence levels in the simple specifications, but drops considerably and becomes insignificant when we condition on the rich set of controls (as in column (4)). We prefer the Poisson specifications to avoid the incidental variables problem of the ordered probit specification when we introduce ethnic-family and country fixed-effects.

$(\exp(0.6072)-1=0.835)$ and by 74% $(\exp(0.555)-1=0.74)$ higher in the historical homeland of partitioned ethnicities and of the other border ethnic groups. The estimates in (1) do not take into account the potential spillovers in civil conflict from areas were partitioned ethnicities reside (the "treatment" group) to areas where non-border ethnicities reside ("control" group). Spillovers may emerge for numerous reasons. First, the battleground between a partitioned ethnic group and another tribe or the central government might take place outside the historical homeland of the split ethnicity. Second, in many cases the conflict in the area of the partitioned ethnicity leads to displacement and refugee flows to nearby areas, which in turn may spur conflict (e.g. Salehyan and Gleditsch (2006) and Blattman and Miguel (2010)). Third, tribes in the same ethnolinguistic family as the partitioned ethnic group might get involved in the conflict to support their peers. For example many Ugandan tribes assisted the ethnically similar Southern Sudanese troops in their long fight against the government troops of the ethnically and religiously distinct Northern Sudan.

Fourth, we have introduced some spatial correlation with the way the dependent variable is generated. If spillovers are present, the estimates in Table 5 column (1) are lower bounds because of non-accounted externalities to ethnic areas where non-partitioned ethnic groups reside (see Miguel and Kremer (2004)). In other words if there are war spillovers to areas where non-split ethnic group reside, then civil war activity in the control group of ethnicities will be higher as compared to the case where partitioning did not have any effect. In columns (2)-(3) we thus include as controls the total number of casualties per square kilometer in each country and each ethnic-family, netting out the number of casualties of each ethnic group's own homeland; this allows us to account for spillovers in civil conflict within countries and within ethnic families. The results show that there are sizable spillovers in both dimensions. For example the estimates in Panel B column (3) imply that a 10% increase in civil war casualties in the country where an ethnicity resides is associated with at 2.5% increase in casualties in the homeland of this ethnic group. Likewise a 10% increase in the number of casualties in nearby areas of the same ethnic cluster is associated with 1.8% increase in casualties for each ethnic group. In Appendix Table A we report estimates from a general spatial lag model. These specifications that model explicitly spatial correlation corroborate the finding that partitioned ethnic groups are more prone to conflict compared to the other border and non-border ethnicities.

In columns (4) and (5) we include ethnic-family and country fixed-effects respectively, while in column (6) we include both. The ML-NB coefficient on the partitioning indicator remains stable (0.61). When we control for ethnic-family factors, countrywide characteristics, and the rich set of controls, the estimate on the $BORDER$ index drops considerably and in the LS specification becomes indistinguishable from zero. In the NB-ML specification with country

fixed-effects and ethnic family fixed-effects, the estimate on $BORDER$ retains statistical significance, but the size of the coefficient is half of that for the partitioning index. (The difference between the two coefficients is statistically significant at the 90% level.) In columns (7) and (8) we report double fixed-effects estimates using the low estimate and the high estimate of conflict casualties with similar results. The coefficient on $SPLIT$ is positive and highly significant. Moreover, the estimate on $SPLIT$ is two times larger than the $BORDER$ coefficient.

4.3 Conflict Duration

In Table 6 we examine the effect of partitioning on the duration of civil conflict. Panel A gives negative binomial ML estimates, while Panel B reports LS estimates where the dependent variable is the log of one plus the war duration. The simple specifications in column (1) suggest that the duration of civil conflict is significantly higher in regions where partitioned ethnicities reside. The NB-ML estimate implies that civil conflicts last on average 38% longer in the historical homeland of partitioned ethnic groups, as compared to regions where non-partitioned groups reside ($\exp(0.32) - 1 = 0.377$). The coefficient on the $BORDER$ variable that captures the degree of civil war duration for ethnic groups that have been affected to a lesser extent by the artificial border design is positive, but smaller and (marginally) insignificant (p-value: 0.15). In columns (2)-(3) we control for spatial spillovers augmenting the specification with the average duration of civil conflict in all other groups within each country and all groups within each ethnic family (netting out the duration in each ethnicity's own location). The LS estimates suggest that a 10% increase in the duration of armed civil conflict in the same country and ethnic-family is associated with approximately 5% increase in local civil war duration.

In columns (4)-(5) we add a vector of ethnic-family and country fixed-effects respectively. The coefficient on the $SPLIT$ indicator retains its economic and statistical significance. Column (6) reports the most restrictive specification, where we control for geography, natural resources, ecology, ethnic-family unobservables and country fixed-factors. The results suggest that civil wars last 35% more in areas of partitioned ethnic groups, as compared to regions where non-partitioned, non-border groups reside ($\exp(0.30) - 1 = 0.35$). The NB-ML estimates further show that compared to non-partitioned non-border ethnic groups, civil conflict duration is higher by 18.5% in border areas where ethnicities modestly affected by the artificial border design reside ($\exp(0.173) - 1 = 0.188$). Our estimates are in line with the cross-country results of Fearon (2004) and Fearon and Laitin (2010), showing that wars involving land conflict between a peripheral ethnic minority and state-supported migrants of a dominant ethnic group are on average quite long-lived (see also Wimmer, *et al.* (2009)).

5 Partitioning and Development

The results in Tables $2-6$ reveal that civil war intensity is much higher in the historical homeland of ethnic groups that have been partitioned by national borders. The results also show that border ethnicities less impacted by the border drawing do suffer more from civil conflict compared to non–partitioned groups, though to a lesser extent. Does this imply that regional development is lower in areas where partitioned ethnic groups reside? This is the question we tackle in Table 7.

Table 7 reports NB-ML estimates (in Panel A) and log-linear LS estimates (in Panel B), where we use luminosity to proxy for regional development. The estimates in column (1) show that regional development is significantly lower in areas of ethnic groups that have been affected by the artificial border design. In column (2) we control for geographic location adding the RD-type cubic polynomial in latitude and longitude. The coefficients on both the $SPLIT$ and the $BORDER$ variables are negative and statistically significant. In column (3) we control for natural resources, geographical features, ecological characteristics, historical and current urbanization. While many of these variables enter with significant coefficients, the estimates on the partitioning indicator variables are not affected. This is in line with the evidence produced in Table 2 that there are no systematic differences with respect to geographic endowments between partitioned ethnicities and non-partitioned ones.

Cross-ethnicity differences in regional development (and electrification) are also driven by country-level factors, such as income, institutional quality, foreign aid, education, country size, etc. Moreover Michalopoulos and Papaioannou (2010) show that ethnic-specific institutional and social features correlate significantly with luminosity. Thus in columns (4) and (5) we augment the specification with ethnic-family fixed-effects and country fixed-effects. Examining the effect of partitioning within countries and within ethnic families, while conditioning on a rich set of local controls, allows us to isolate the direct effect of partitioning on development. The estimate on $SPLIT$ in the NB-ML specification drops in absolute value compared to the specification without the ethnic-family and the country fixed-effects (from -0.835 to -0.486); yet the coefficient retains its significance at the 99% confidence level. It implies that regional development is approximately 60% lower in areas where partitioned ethnic groups reside, compared to non-border ethnic groups ($\exp(0.485)-1=0.62$). The light density data contain some extreme observations. Thus in columns (6) and (7) we report specifications where we trim and winsorize the top 1% of the distribution, respectively.[34] The coefficient on the $SPLIT$ dummy in the NB-ML specifications retains significance at the 99% level. The $BORDER$ dummy also enters with a negative and significant estimate, suggesting that regional development is

[34] The results are similar if we trim or winsorize the data at the top 2% or top 5% of the lights distribution.

significantly lower in all ethnic areas were marginally affected by the artificial border design.

6 Further Robustness Checks

We have performed many sensitivity checks to investigate the robustness of our results reported in the supplementary Appendix Tables A, B, and C.

In Appendix Table A we show that our results associating civil conflict and regional development with partitioning are strong when we estimate spatial lag models that account for spatial interdependencies in the dependent variable and the error term.

In Panel A of Appendix Table B we report NB-ML estimates excluding ethnicities in Northern African countries (Algeria, Egypt, Morocco, Tunisia, Libya). This is a necessary robustness check because Europeans had long contacts with Northern Africa and were familiar with local ethnic and geographic conditions. The results are similar to the ones presented in Tables 4-7. In Panel B of Appendix Table B we exclude from the estimation countries in Southern Africa (Malawi, Mozambique, Namibia, South Africa, and Zimbabwe). This is important because the apartheid governments of South Africa intervened in most civil wars in the neighboring countries. Again all our results remain unchanged.

In Appendix Table C we report specifications associating civil war and development with a continuous measure of ethnic partitioning that captures the probability that a square kilometer of an ethnic homeland falls to different countries. According to this continuous index the most fractionalized ethnicity is the Malinke (0.77), a group partitioned between 6 different countries. In all permutations tribal land fractionalization enters with a significant estimate implying that ethnic partitioning leads to more devastating and prolonged civil conflict and to a lower level of development.[35]

7 Conclusion

We examine the economic consequences of a neglected aspect of colonization, the artificial drawing of African political boundaries among European powers in the end of the 19th century which led to the partitioning of several ethnicities across African states.

In the first part of our paper we formally establish the random nature of African political boundaries. Using regional data from Murdock (1959) on the spatial distribution of ethnicities

[35] We prefer the binary index of ethnic partitioning for two reasons. First, all studies in African historiography and political science we reviewed suggest that what matters for civil conflict and development is whether an ethnicity has been partitioned or not rather than the degree of the split. Second, from a theoretical point of view a priori there is no clear guidance as to whether propensity to conflict should monotonically increase with the degree of partitioning. Finally, to the extent that Murdock's tribal map certainly contains measurement error, small differences in this continuous measure are more likely to be affected by it.

at the time of colonization, we estimate probabilistic models associating ethnic partitioning with various geographical, ecological, and ethnic-specific precolonial characteristics. With the sole exceptions of the size of the historical homeland and the magnitude of water bodies, there are no other significant differences between partitioned and non-partitioned ethnicities. These results support African historiography on the accidental and artificial drawing of the colonial/national borders.

Second, we examine the effect of ethnic partitioning on various aspects of civil conflict (incidence, casualties, duration), as this has been hypothesized to be the major effect of the scramble for Africa. In contrast to most works on the long-run effects of colonization on civil conflict (and development), our analysis is based on detailed regional data spanning 834 ethnic areas across Africa. Our rich dataset allows us to control at a very fine level for natural resources, geography and early development. We also include country and ethnic-family fixed-effects to account for national factors and broad cultural characteristics respectively. We find that partitioned ethnicities compared to tribes that have not been directly affected by the improper border design, have experienced more civil war incidents that lasted longer and were more devastating in terms of casualties.

Third, using satellite light density data at night to proxy for development at the ethnicity level, we show that partitioning is negatively correlated with luminosity. While one could always argue that another ethnic-specific feature is driving the significant negative correlation between partitioning and development, the largely arbitrary design of contemporary African borders together with the rich set of conditioning factors at a fine regional level and the inclusion of country and ethnic-family fixed-effects suggests that these correlations imply a causal relationship. The uncovered contemporary divergence in comparative economic performance between partitioned and non-partitioned groups becomes more dramatic when viewed in light of the finding that partitioned and non-partitioned ethnic groups were economically similar both in the eve of colonization and at the time of African independence. Our work suggests that the scramble for Africa by partitioning ethnicities in different countries spurred civil conflict and unrest shaping their economic trajectory.

8 Data Appendix

Light Density at Night: Light Density is calculated at a ethnicity level averaging light density observations across pixels that fall within the historical homeland of each of the 834 ethnic groups in Murdock's Atlas. To smooth weather variation we use the average of the values in 2007 and 2008.
Source: Available at http://www.ngdc.noaa.gov/dmsp/global_composites_v2.html.

Population Density: Log of population density per sq. km. in 1960 plus one. *Source: UNESCO (1987). Available at: http://na.unep.net/datasets/datalist.php.*

Civil Conflict Casualties (Best.High/Low Estimate): See text

Number of Civil Conflicts: See text

Number of Civil War Conflict Zones: See text

Civil Conflict Duration: See text

Land Area: Log surface area of the historical homeland of each ethnic group in 1000*s* of sq. km. *Source: Global Mapping International, Colorado Springs, Colorado, USA.*

Water Area: Log of one plus the total area of the historical homeland of each ethnic group covered by rivers or lakes in sq. km. *Source: Constructed using the "Inland water area features" dataset from Global Mapping International, Colorado Springs, Colorado, USA.*

Elevation: Average elevation in km. *Source: National Oceanic and Atmospheric Administration (NOAA) and U.S. National Geophysical Data Center, TerrainBase, release 1.0 (CD-ROM), Boulder, Colorado. http://www.sage.wisc.edu/atlas/data.php?incdataset=Topography*

Land Suitability for Agriculture: Average land quality for cultivation within the area of each ethnic-country observation. The index is the product of two components reflecting the climatic and soil suitability for cultivation. *Source: Michalopoulos (2008); Original Source: Atlas of the Biosphere. Available at http://www.sage.wisc.edu/iamdata/grid_data_sel.php.*

Malaria Stability Index: The index takes into account the prevalence and type of mosquitoes indigenous to a region, their human biting rate, their daily survival rate, and their incubation period. The index has been constructed for 0.5 degree by 0.5 degree grid-cells globally. *Source: Kiszewski et al. (2004)*

Capital City: Indicator variable that equals one when the capital city falls into the historical homeland of each ethnic group area and zero otherwise.

Sea Distance: The geodesic distance of the centroid of the historical homeland of each ethnic group from the nearest coastline, measured in 1000*s* of km's. *Source: Global Mapping International, Colorado Springs, Colorado, USA. Series name: Global Ministry Mapping System. Series issue: Version 3.0*

Petroleum: Indicator variable that takes on the value of one if an oil field is found in

the historical homeland of an ethnic group and zero otherwise. *Source: The Petroleum Dataset v.1.1 contains information on all known on-shore oil and gas deposits throughout the world. http://www.prio.no/CSCW/Datasets/Geographical-and-Resource/Petroleum-Dataset/Petroleum-Dataset-v11/*

Diamond: Indicator variable that takes on the value of one if a diamond mine is found in the historical homeland of an ethnic group and zero otherwise. *Source: Map of Diamond Resources.* www.prio.no/CSCW/Datasets/Geographical-and-Resource/Diamond-Resources/

City in 1400: Indicator variable that takes on the value of one if a city with a population larger than $20,000$ in 1400 was in the historical homeland of an ethnic group and zero otherwise. *Source: Chandler (1987).*

Split: Indicator variable that equals 1 if at least 10% of the historical homeland of an ethnic group is partitioned into different countries. *Source: Calculated intersecting Murdock's (1959) ethnic map of Africa with the Digital Chart of the World (DCW) shapefile. The latter contains the polygons delineating the international boundaries in 1992. Appendix Table 1 reports partitioned ethnicities.*

Border: Indicator variable that equals one when a miniscule or small part ($> 0\%$ and $< 10\%$) of the historical homeland of an ethnic group falls into different countries. *Source: Calculated intersecting Murdock's (1959) ethnic map of Africa with the Digital Chart of the World (DCW) shapefile. The latter contains the polygons delineating the international boundaries in 1992. Appendix Table 2 reports partitioned ethnicities.*

Continuous Measure of Partitioning: If an ethnic homeland belongs to n countries with π_i denoting the fraction of the tribal homeland belonging to country i then tribal land fractionalization is calculated as: $1 - \sum_{i=1}^{n} \pi_i^2$

Latitude: *Source: Constructed using ArcGis Software.*

Longitude: *Source: Constructed using ArcGis Software.*

Regional Indicators: There are five regional indicator variables, North Africa, Western Africa, Central Africa, Eastern Africa, and Southern Africa. *Source: Nunn (2008).*

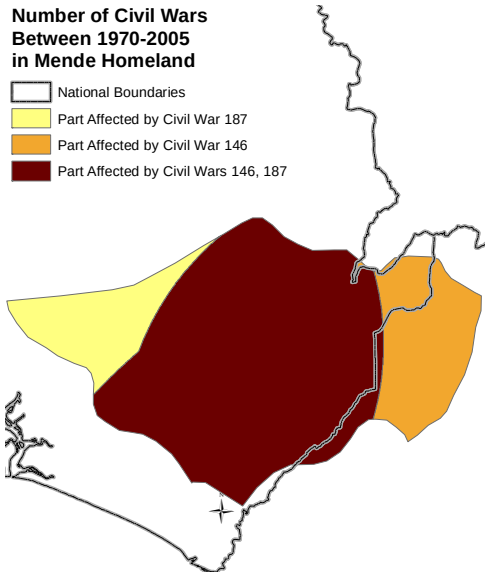
Slavery: Number of persons of each ethnic group that were shipped during the trans-Atlantic and Indian Ocean slave trades. Following Nunn (2008) in the regressions we use the log of one plus the number of slaves standardized by the land area of each ethnic group's historical homeland. *Source: Nunn (2008).*

Political Centralization: The binary index is constructed using Murdock's Jurisdictional Hierarchy beyond Local Community $0-4$ index that indicates the number of jurisdictional levels (political complexity) in each society above the local level**.** The political centralization index takes the value 0 if the Jurisdictional Hierarchy beyond Local Community variable equals

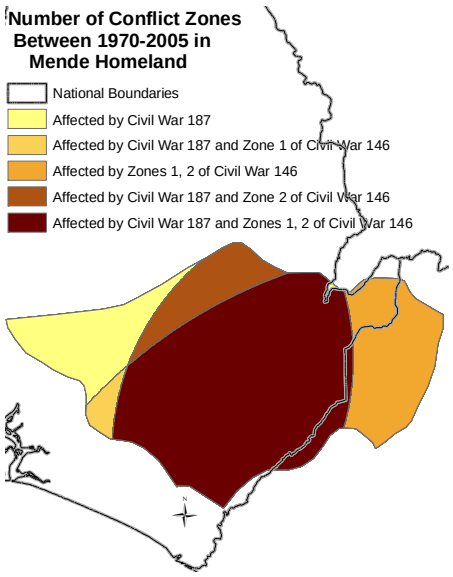
0 or 1 (when the society is classified as either stateless or forming a small chiefdom). The index takes on the value 1 if the Jurisdictional Hierarchy beyond Local Community variable equals 2, 3, and 4 (when the society is classified as being part of large paramount chiefdom or a large state). This aggregation follows Gennaioli and Rainer (2006, 2007). *Source: Murdock (1967). A revised version of Murdock's Atlas has been made available by J. Patrick Gray at: http://eclectic.ss.uci.edu/~drwhite/worldcul/EthnographicAtlasWCRevisedByWorldCultures.sav.*

Class Stratification: This binary index reflects "*the degree of class differentiation, excluding purely political and religious statuses*". A zero score indicates "*absence of significant class distinctions among freemen, ignoring variations in individual repute achieved through skill, valor, piety, or wisdom.*" A score of 1 indicates some degree of stratification based on wealth or elite distinctions, the presence of hereditary aristocracy, or complex class differentiation. The index is based on Murdock's $0-4$ class stratification index. This aggregation follows Gennaioli and Rainer (2006, 2007). *Source: Murdock (1967).*

Property Rights. A binary index that equals one when inheritance rights for land exist and zero otherwise. This aggregation follows Fenske (2010). *Source: Murdock (1967).*



Appendix Figure 1*a*



Appendix Figure 1*b*

References

Acemoglu, Daron, Maria Angelica Bautista, Pablo Querubin, and James A. Robinson, "Economic and Political Inequality in Development: The Case of Cundinamarca, Colombia," in Elhanan Helpman, ed., *Institutions and Economic Performance*, Harvard University Press, 2008.

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

___, ___, and ___, "Reversal Of Fortune: Geography And Institutions In The Making Of The Modern World Income Distribution," *Quarterly Journal of Economics*, 2002, *107* (4), 1231–1294.

___, ___, and ___, "Institutions as a Fundamental Cause of Long-Run Growth," *Handbook of Economic Growth*, 2005, *91*, 386–472.

Alesina, Alberto and Ekaterina Zhuravskaya, "Segregation and the Quality of Government in a Cross-Section of Countries," *American Economic Review, forthcoming*, 2011.

___ and Eliana La Ferrara, "Fractionalization," *Journal of Economic Literature*, 2005, *43*, 762–800.

___, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg, "Fractionalization," *Journal of Economic Growth*, 2003, *8*, 155–194.

___, William Easterly, and Janina Matuszeski, "Artificial States," 2011. Journal of the European Economic Association, forthcoming.

Arbesu, Horacio Alejandro Larreguy, "Colonialism and Economic Development: Evidence from a Natural Experiment in Colonial Nigeri," 2011. mimeo MIT.

Asiwaju, A., "The Conceptual Framework," in "Partitioned Africans," New York: St. Martin Press, 1985, pp. 1–18.

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

Barkindo, Bawuro M., "The Mandara Astride the Nigeria-Cameroon Boundary," in A.I. Asiwaju, ed., *Partitioned Africans*, New York, NM: St. Martin's Press,, 1985, pp. 155–194.

Bates, Robert H., in "States and Markets in Africa," Berkeley, CA: University of California Press, 1981.

Bellows, John and Edward Miguel, "War and Local Collective Action in Sierra Leone," *Journal of Public Economics*, 2009, *93*, 1144–1157.

Berger, Daneil, "Taxes, Institutions and Governance: Evidence from Colonial Nigeria," 2009. mimeo, New York University.

Blattman, Chris and Edward Miguel, "Civil War," *Journal of Economic Literature*, 2010, *48* (1), 3–57.

Buhaug, Halvard and Jan Ketil Rod, "Local Determinants of African Civil Wars," *Political Geography*, 2006, *25* (1), 315–335.

Chandler, Tertius, *Four Thousand Years of Urban Growth: An Historical Census*, Lewiston, NY: Edwin Mellon Press, 1987.

Chen, Xi and William D. Nordhaus, "The Value of Luminosity Data as a Proxy for Economic Statistics," 2010. Working Paper, NBER, No. 16317.

Ciccone, Antonio and Marek Jarocinski, "Determinants of Economic Growth: Will Data Tell?," *American Economic Journal Macroeconomics*, 2010, *2*, 222–246.

Collier, Paul and Anke Hoeffler, "On Economic Causes of Civil War," *Oxford Economic Papers*, 1998, *50* (4), 563–573.

___ and ___ , "Civil War," in Todd Sandler and Keith Hartley, eds., *Handbook of Defense Economics, Volume 2, Defense in a Globalized World,*, Amsterdam and Oxford: Elsevier, North Holland, 2007, pp. 711–740.

___ and Nicholas Sambanis, *Understanding Civil War. Volume 1: Africa*, Washington, DC: World Bank, 2005.

___ , Anke Hoeffler, and Mans Soderbom, "On the Duration of Civil War," *Journal of Peace Research*, 2004, *41* (3), 253–273.

Collins, David, "Partitioned Culture Areas and Smuggling: The Hausa and Groundnut Trade across the Nigeria-Niger Boundary up to the 1970s," in A.I. Asiwaju, ed., *Partitioned Africans*, New York, NM: St. Martin's Press„ 1985, pp. 155–194.

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

Dell, Melissa, "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 2010, *78*.

Dowden, Richard, *Africa: Altered States, Ordinary Miracles*, London, UK: Portobello Books Ltd, 2008.

Easterly, William and Ross Levine, "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, 1997, *112* (4), 1203–1250.

____ and ____ , "The European Origins of Economic Development," 2009. Working Paper, mimeo Brown University.

Englebert, Pierre, Stacy Tarango, and Matthew Carter, "Dismemberment and Suffocation: A Contribution to the Debate on African Boundaries," *Comparative Political Studies*, 2002, *35* (10), 1093–1118.

Esteban, Joan, Laura Mayoral, and Debraj Ray, "Ethnicity and Conflict: An Empirical Investigation," 2010. mimeo NYU Department of Economics.

Ethnologue, Languages of the World, SIL International; Fifteenth edition 2005.

Fearon, James, "Why Do Some Civil Wars Last So Much Longer Than Others?," *Journal of Peace Research*, 2004, *41*, 275–302.

____ and David Laitin, "Ethnicity, Insurgency and Civil War," *American Political Science Review*, 2003, *97*, 75–90.

____ and ____ , "Sons of the Soil, Migrants, and Civil War," 2010. World Development, forthcoming.

Fenske, James, "Does Land Abundance Explain African Institutions," 2010. mimeo Yale University.

Gennaioli, Nicola and Ilia Rainer, "Precolonial Centralization and Institutional Quality in Africa," in Mark Gradstein and Kai Konrad, eds., *Institutions and Norms in Economic Development*, MIT Press, 2006.

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

Glaeser, Edward L., Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "Do Institutions Cause Growth?," *Journal of Economic Growth*, 2004, *9* (3), 271–303.

Hargreaves, J. D., *West Africa Partitioned, Vol. II. The Elephants adn the Grass*, London: Mc Millan, 1985.

Henderson, Vernon J., Adam Storeygard, and David N. Weil, "Measuring Economic Growth from Outer Space," 2009. NBER Working Paper No. 15199.

Herbst, Jeffrey, *State and Power in Africa*, Princeton, NJ: Princeton University Press, 2000.

Herge, Havard and Nicholas Sambanis, "Sensitivity Analysis of Empirical Results on Civil War Onset," *Journal of Conflict Resolution*, 2006, *50*, 508–535.

Higham, Said S., "The Somali Dilema: Nation in Search for State," in A.I. Asiwaju, ed., *Partitioned Africans*, New York, NM: St. Martin's Press,, 1985, pp. 155–194.

Hopkins, A. G., *An Economic History of West Africa*, New York, NY: Columbia University Press, 1973.

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

i Miguel, Gerard Padro, "The Control of Politicians in Divided Societies: The Politics of Fear," *Review of Economic Studies*, 2007, *74* (4), 1259–1274.

Iyer, Lakshmi, "Direct versus Indirect Colonial Rule in India: Long-term Consequences," *Review of Economics and Statistics*, 2010, *forthcoming.*

Kalyvas, Stathis, "Civil Wars," in Carles Boix and Susan Stokes, eds., *Handbook of Political Science*, New York: Oxford University Press, 2007, pp. 416–434.

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

Lacina, Bethany and Nils Petter Gleditsch, "Monitoring Trends in Global Combat: A New Dataset of Battle Deaths," *European Journal of Population*, 2005, *21* (2-3), 145–116.

Levine, Ross E. and David Renelt, "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review*, 1992, *84* (2), 942–963.

Martin, Xavier Sala I., Gernot Doppelhofer, and Ronald I. Miller, "Determinants of Long-Run Growth: a Bayesian Averaging of Classical Estimates (BACE) Approach," *American Economic Review*, 2004, *94* (4), 813–835.

Matuszeski, Janina and Frank Schneider, "Patterns of Ethnic Group Segregation and Civil Conflict," 2006. Working Paper, mimeo Harvard University.

Michalopoulos, Stelios, "The Origins of Ethnic Diversity," 2008. Working Paper, mimeo Tufts University.

___ and Elias Papaioannou, "Divide and Rule or the Rule of the Divided? Evidence from Africa," 2010. CEPR Discussion Paper, CEPR DP No. 6414.

Miguel, Edward and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 2004, *72*, 159–217.

Miller, Douglas, Colin A.Cameron, and Jonah Gelbach, "Robust Inference with Multi-Way Clustering," 2006. NBER Technical Working Paper No. T0327.

Montalvo, José G. and Marta Reynal-Querol, "Ethnic Diversity and Economic Development," *Journal of Development Economics*, 2005, *76*, 293–323.

___ and ___ , "Ethnic Polarization, Potential Conflict and Civil War," *American Economic Review*, 2005, *95*, 796–816.

Murdock, George P., *Africa: Its Peoples and Their Culture History*, New York: McGraw-Hill Book Company, 1959.

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

Naritomi, Joana, Rodrigo R. Soares, and Juliano J. Assunção, "Institutional Development and Colonial Heritage within Brazil," 2009. Working Paper, IZA DP No. 4276.

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

___ **and Diego Puga**, "Ruggedness: The Blessing of Bad Geography in Africa," *Review of Economics and Statistics, forthcoming*, 2011.

___ **and Leonard Wantchekon**, "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review, forthcoming*, 2011.

Petter, Gleditsch Nils, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Havard Strand, "Armed Conflict 1946-2001: A New Datase," *Journal of Peace Research*, 2002, *39* (5), 615–637.

Raleigh, Clionadh, David Cunningham, Lars Wilhelmsen, and Nils Petter Gleditsch, "Conflict Sites: 1946-2005. v2.0," 2006. Centre for the Study of Civil War, PRIO.

Renner, F.A., "Ethnic Affinity: Partition and Political Integration in Senegambia," in A.I. Asiwaju, ed., *Partitioned Africans*, New York, NM: St. Martin's Press,, 1985, pp. 71–86.

Ross, Michael L., "A Closer Look on Oil, Diamonds, and Civil War," *Anuual Review of Political Science*, 2006.

Salehyan, Idean and Kristian Skrede Gleditsch, "Refugees and the Spread of Civil War," *International Organization*, 2006, *60* (02), 335–366.

Silva, J.M.C Santos and Silvana Tenreyro, "The Log of Gravity," *Review of Economics and Statistics*, 2006, *88* (4), 641–658.

___ , ___ **, and Frank Windmeijer**, "Is it Different for Zeros? Discriminating Between Models for Non-Negative Data with Many Zeros," 2010. CeMMAP working papers CWP20/10.

Wesseling, H. L., *Divide and Rule: The Partition of Africa, 1880-1914*, Westport, Conn: Praeger Publishers, 1996.

Wimmer, Andreas, Lars-Erik Cederman, and Brian Min., "Ethnic politics and Armed Conflict. A Configurational Analysis of a New Global Dataset," *American Sociological Review*, 2009, *74*, 316–337.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, MIT Press, 2002.

Table 1A : Summary Statistics

	Obs.	mean	st. dev.	min	p25	median	p75	max
Panel A: Outcome Measures								
Number of Civil Conflicts	834	1.14	0.95	0.00	0.00	1.00	2.00	5.00
Number of Civil War Conflict Zones	834	1.58	1.52	0.00	0.00	1.00	2.00	8.00
Civil Conflict Casualties (Best Estimate)	834	39.62	82.25	0.00	0.00	3.21	40.12	589.53
Civil Conflict Casualties (Low Estimate)	834	33.57	77.59	0.00	0.00	1.90	40.12	561.34
Civil Conflict Casualties (High Estimate)	834	45.23	91.47	0.00	0.00	4.32	40.68	639.37
Civil Conflict Duration	834	7.75	9.39	0.00	0.00	3.86	11.63	46.79
Light Density (Development)	834	0.46	1.83	0.00	0.00	0.03	0.17	25.41
Panel B: Control Variables								
Land Area	834	34.06	58.97	0.24	6.13	14.44	35.94	604.90
Land Area under Water	834	0.86	2.25	0.00	0.01	0.17	0.66	27.66
Latitude	834	4.46	13.11	-33.09	-2.92	6.59	11.08	36.58
Longitude	834	17.85	15.47	-16.41	7.32	18.40	31.22	49.25
Population Density at Independence	834	17.47	25.71	0.00	3.17	9.26	20.47	321.53
Mean Elevation	834	0.62	0.43	0.00	0.30	0.49	0.93	2.17
Land Suitability for Agriculture	834	0.41	0.24	0.00	0.25	0.42	0.57	0.98
Malaria Stability Index	834	0.75	0.36	0.00	0.57	0.98	1.00	1.00
Average Distance to the Seacoast	834	600.76	435.52	1.49	209.33	552.27	924.67	1697.01
Diamond Mine Indicator	834	0.12	0.33	0.00	0.00	0.00	0.00	1.00
Oil Deposit Indicator	834	0.12	0.40	0.00	0.00	0.00	0.00	4.00
Capital City Indicator	834	0.06	0.23	0.00	0.00	0.00	0.00	1.00
Major City in 1400 AD Inidcator	834	0.04	0.20	0.00	0.00	0.00	0.00	2.00
Panel C: Pre-colonial Ethnic-Specific Variables								
Slave Exports	834	551.70	3289.21	0.00	0.00	0.00	17.67	41045.08
Pre-colonial Political Centralization	440	0.34	0.48	0.00	0.00	0.00	1.00	1.00
Pre-colonial Class Stratification	396	0.54	0.50	0.00	0.00	1.00	1.00	1.00
Pre-colonial Property Rights Indicator	375	0.93	0.25	0.00	1.00	1.00	1.00	1.00
Pre-colonial Share of Agriculture	490	5.63	1.79	0.00	5.00	6.00	7.00	9.00

The table reports descriptive statistics for all variables employed in the empirical analysis. The Data Appendix gives detailed variable definitions and data sources.

Table 1B : Correlation Structure Dependent Variables

Civil Conflict Casualties (Best Estimate)	1						
Civil Conflict Casualties (Low Estimate)	0.9780*	1					
Civil Conflict Casualties (High Estimate)	0.9787*	0.9344*	1				
Number of Civil Conflicts	0.3669*	0.3383*	0.3926*	1			
Number of Civil War Conflict Zones	0.5851*	0.5560*	0.5900*	0.8484*	1		
Civil Conflict Duration	0.5445*	0.5081*	0.5392*	0.6411*	0.7101*	1	
Light Density (Development)	-0.0494	-0.0667	-0.0475	-0.0631	-0.0743*	-0.0819*	1

The table reports the correlation structure among the outcome measures (civil conflict and economic development). * indicates statistical significance at the 95% confidence level. The Data Appendix gives detailed variable definitions and data sources.

Table 1C: Differences in Outcomes between Partitioned and Non-Split Ethnic Groups

	Non-Border Groups	Border Groups		Partitioned Groups	
	(1)	(2)		(3)	
	mean	mean	difference	mean	difference
Number of Civil Conflicts	0.968	1.315	0.346	1.407	0.438
	(0.038)	(0.093)	(0.036)	(0.067)	(0.072)
			0.00		*0.00*
Number of Civil War Conflict Zones	1.2710	1.9921	0.7211	1.9827	0.7117
	(0.0575)	(0.1697)	(0.1792)	(0.1064)	(0.1210)
			0.00		*0.00*
Civil Conflict Casualties	29.702	57.383	27.682	50.276	20.574
(Best Estimate)	(62.115)	(10.175)	(10.566)	(94.004)	(6.809)
			0.01		*0.00*
Civil Conflict Casualties	26.006	49.374	23.368	40.458	14.452
(Low Estimate)	(58.237)	(9.805)	(10.162)	(88.344)	(6.396)
			0.03		*0.02*
Civil Conflict Casualties	32.532	64.107	31.575	61.005	28.474
(High Estimate)	(67.244)	(10.814)	(11.244)	(109.771)	(7.853)
			0.01		*0.00*
Civil Conflict Duration	7.229	8.705	1.476	8.290	1.061
	9.346	(0.864)	(0.964)	(9.245)	(0.744)
			0.13		*0.15*
Light Density (Development)	0.611	0.228	-0.383	0.273	-0.337
	(2.319)	(0.057)	(0.121)	(0.800)	(0.119)
			0.00		*0.00*
Observations	476	127		231	

The table reports summary statistics and test of means for the outcome measures (civil conflict and economic development). Column (1) gives summary statistics for non-border ethnic groups. Column (2) gives summary statistics for ethnic groups whose historical homeland is adjacent to the national border, but is not significantly partitioned. Column (3) gives summary statistics for partitioned ethnic groups, where at least 10% of the historical homeland falls into more than one contemporary country. For each category of ethnicities we report the mean and the standard deviation in parentheses. The number of observations is reported in italics. The table also reports the difference, the standard error of the difference, and the corresponding p-value of a test of mean equality (assuming unequal variances) between the non-border groups and the other two categories. The Data Appendix gives detailed variable definitions and data sources.

Table 2 - Are African Borders Artificial?

	SPLIT	BROAD	SPLIT	BROAD	SPLIT	BROAD	SPLIT	BROAD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Land Area	0.1256**	0.2243***	0.1149**	0.2462***	0.0998*	0.2204***	0.0931*	0.2194***
	(0.0600)	(0.0654)	(0.0526)	(0.0584)	(0.0502)	(0.0597)	(0.0560)	(0.0639)
Land Area under Water	0.3910***	0.4878***	0.3711***	0.4778***	0.3908***	0.4888***	0.3925***	0.4891***
	(0.0803)	(0.1340)	(0.0808)	(0.1274)	(0.0845)	(0.1295)	(0.0861)	(0.1311)
Elevation	-0.0610	-0.0044	-0.0608	0.0676	-0.0933	0.0398	-0.0814	0.0418
	(0.2443)	(0.2462)	(0.2557)	(0.2736)	(0.2582)	(0.2812)	(0.2522)	(0.2727)
Suitability for Agriculture	0.3956	0.3250	0.5867	0.4169	0.6138	0.4271	0.6399	0.4313
	(0.3802)	(0.3288)	(0.3954)	(0.3816)	(0.3860)	(0.3657)	(0.3868)	(0.3573)
Malaria Stability Index	0.2504	0.2705	-0.1065	0.2044	-0.1225	0.1827	-0.1274	0.1817
	(0.2514)	(0.2593)	(0.3583)	(0.3714)	(0.3591)	(0.3751)	(0.3549)	(0.3718)
Distance to the Seacoast	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Diamond					0.1509	0.2523	0.1527	0.2526
					(0.1740)	(0.1681)	(0.1758)	(0.1699)
Oil					-0.0689	0.0192	-0.0639	0.0202
					(0.1679)	(0.1476)	(0.1685)	(0.1491)
Major City in 1400AD					-0.0089	-0.3583	-0.0065	-0.3569
					(0.2165)	(0.2439)	(0.2166)	(0.2447)
Slave Exports					0.0122	-0.007	0.013	-0.007
					(0.0310)	(0.0336)	(0.0306)	(0.0333)
Population Density 1960							-0.0228	-0.0038
							(0.0713)	(0.0717)
Region Fixed-Effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Log Likelihood	-472.817	-527.644	-463.695	-518.69	-463.057	-517.348	-462.961	-517.345
pseudo R-squared	0.039	0.074	0.058	0.090	0.059	0.092	0.059	0.092
Observations	834	834	834	834	834	834	834	834

The table reports probit marginal effects associating whether a group is partitioned with ethnic-specific features. The dependent variable is a dummy variable that equals one when an ethnicity is partitioned by the national border. In odd numbered columns partitioned ethnicities are identified as those with at least 10% of the historical homeland belonging to more than one contemporary countries (SPLIT=1). In even numbered columns we broaden the definition and identify partitioned ethnicities with those impacted by the national borders irrespective of the degree of partitioning (BROAD=1). The specifications in columns (3)-(8) include a set of region fixed-effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 3 - Do Ethnic Characteristics Predict Partitioning?

	SPLIT	BROAD	SPLIT	BROAD	SPLIT	BROAD	SPLIT	BROAD	SPLIT	BROAD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Land Area	0.1672**	0.2997***	0.1788**	0.2465***	0.1300*	0.2552***	0.1721**	0.2834***	0.1230	0.2583**
	(0.0609)	(0.0805)	(0.0691)	(0.0900)	(0.0714)	(0.0889)	(0.0598)	(0.0722)	(0.0890)	(0.1062)
Land Area under Water	0.2452**	0.3634***	0.3274***	0.4526***	0.3047***	0.4934***	0.2803**	0.4140***	0.4732***	0.6136***
	(0.0976)	(0.1243)	(0.0944)	(0.1426)	(0.1087)	(0.1525)	(0.0952)	(0.1290)	(0.1514)	(0.1956)
Political Centralization	-0.1865	-0.2367							0.1023	-0.1377
	(0.1668)	(0.1679)							(0.2246)	(0.2167)
Class Stratification			-0.2433	-0.0866					-0.2814	-0.0565
			(0.2433)	(0.0866)					(0.2165)	(0.1845)
Property Rights					0.0671	0.2392			-0.0245	0.1224
					(0.2675)	(0.2746)			(0.3685)	(0.3065)
Dependence on Agriculture							0.0280	0.0195	0.0354	0.0247
							(0.0328)	(0.0340)	(0.0474)	(0.0421)
Region Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log Likelihood	-251.81	-267.26	-216.82	-238.48	-207.39	-225.26	-274.00	-296.36	-165.34	-175.90
pseudo R-squared	0.07	0.12	0.08	0.12	0.07	0.13	0.07	0.12	0.09	0.14
Observations	440	440	396	396	375	375	490	490	295	295

The table reports probit marginal effects associating whether a group is partitioned with pre-colonial ethnic-specific features, reflecting political centralization (in (1), (2), (9), and (10)), class stratification (in (3), (4), (9), and (10)), property rights protection (in (5), (6), (9), and (10)), and historical economic development reflected in the dependence on agriculture (in (7)-(10)). The dependent variable is a dummy variable that equals one when an ethnicity is partitioned by the national border. In odd numbered columns partitioned ethnicities are identified as those with at least 10% of the historical homeland belonging to more than one contemporary countries (SPLIT=1). In even numbered columns we broaden the definition and identify partitioned ethnicities with those impacted by the national borders irrespective of the degree of partitioning (BROAD=1). All specifications include a set of region fixed-effects (constants not reported), log land area, and log land area under water (lakes, rivers, and other streams).

The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 4: Partitioning and the Incidence of Civil Conflict (Poisson ML Estimates)

	(1)	(2)	(3)	(4)	(5)
	Panel A - Dependent Variable: Number of Civil Conflicts				
SPLIT - Partitioning (>10%)	0.3316***	0.3213***	0.3139***	0.2750***	0.2336***
	(0.0752)	(0.0649)	(0.0621)	(0.0537)	(0.0526)
BORDER - Partitioning (0% - 10%)	0.2255**	0.1826**	0.1368*	0.0914	0.0980
	(0.0919)	(0.0786)	(0.0758)	(0.0627)	(0.0597)
Log Likelihood	-1034.05	-995.85	-965.83	-855.41	-825.33
R-squared	0.222	0.316	0.406	0.670	0.738
	Panel B - Dependent Variable: Number of Conflict Zones				
SPLIT - Partitioning (>10%)	0.4194***	0.4212***	0.4128***	0.2872***	0.2573***
	(0.0981)	(0.0815)	(0.0799)	(0.0545)	(0.0538)
BORDER - Partitioning (0% - 10%)	0.3864***	0.3720***	0.3175***	0.1038*	0.1449***
	(0.1317)	(0.1070)	(0.1023)	(0.0616)	(0.0538)
Log Likelihood	-1269.93	-1218.97	-1164.74	-969.33	-925.11
R-squared	0.269	0.320	0.435	0.752	0.808
Region Fixed-Effects	Yes	Yes	Yes	Yes	No
Geography	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	Yes	Yes	Yes	Yes
Additional Controls	No	No	Yes	Yes	Yes
Ethnic Family Fixed-Effects	No	No	No	Yes	Yes
Country Fixed-Effects	No	No	No	No	Yes
Observations	834	834	834	834	834

The table reports Poisson ML estimates associating civil war with partitioning and other ethnicity-specific measures. In Panel A the dependent variable is the number of civil wars that have taken place in the historical homeland of an ethnic group between 1970 and 2005. In Panel B the dependent variable is the number of conflict zones associated with the civil wars that have affected the historical homeland of an ethnic group during the period 1970-2005. SPLIT is an indicator variable that identifies partitioning ethnicities as those with at least 10% of the historical homeland belonging to more than one contemporary country. BORDER is an indicator that identifies ethnic groups residing by the border. These groups also fall into more than one country but more than 90% of the historical homeland lies in one country. All specifications include a set of region fixed-effects (constants not reported), log land area, log land area under water (lakes, rivers, and other streams), and population density around independence (in 1960). Columns (2)-(6) include a regression discontinuity (RD) cubic polynomial in latitude and longitude of the centroid of each ethnic group. Columns (3)-(6) include a rich set of control variables (land suitability for agriculture, elevation, a malaria stability index, an indicator of early development that equals one when a major city was in the ethnicity's historical homeland in 1400, an oil indicator and a diamond mine indicator). Columns (4)-(5) include a set of ethnic-family fixed-effects (constants not reported). Column (5) includes a set of country fixed-effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 5 - Partitioning and Civil War Casualties

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Negative Binomial ML				
SPLIT	0.6072***	0.9134***	0.6834***	0.6160***	0.5858***	0.6102***	0.5791***	0.6228**
Partitioning (>10%)	(0.2006)	(0.2373)	(0.1621)	(0.1896)	(0.1914)	(0.1775)	(0.1861)	(0.1786)
BORDER	0.5550***	0.6173***	0.3893**	0.2032	0.3689***	0.3134***	0.2888***	0.3228**
Partitioning (0% - 10%)	(0.1403)	(0.1170)	(0.1539)	(0.1460)	(0.1316)	(0.1389)	(0.1358)	(0.1410)
Country Casualties		0.0001***	0.0001***	0.0001***				
		(0.0000)	(0.0000)	(0.0000)				
Ethnic Family Casualties		0.0001***	0.0001***		0.0001***			
		(0.0000)	(0.0000)		(0.0000)			
Log Likelihood	-3192.98	-3108.18	-2916.83	-2677.24	-2745.51	-2556.05	-2343.46	-2667.39
				Panel B: Log Linear OLS				
SPLIT	0.6053**	0.7638***	0.6779***	0.4768***	0.3458**	0.3558**	0.3209**	0.3905**
Partitioning (>10%)	(0.2488)	(0.1674)	(0.1537)	(0.1618)	(0.1512)	(0.1648)	(0.1525)	(0.1684)
BORDER	0.4836**	0.4191***	0.3418***	0.1133	0.1935*	0.1527	0.1352	0.1695
Partitioning (0% - 10%)	(0.2430)	(0.1246)	(0.1072)	(0.1678)	(0.1024)	(0.1375)	(0.1248)	(0.1461)
Country Casualties		0.3283***	0.2460***	0.1686***				
		(0.0682)	(0.0617)	(0.0570)				
Ethnic Family Casualties		0.2362***	0.1768***		0.1045***			
		(0.0588)	(0.0538)		(0.0366)			
adjusted R-squared	0.221	0.529	0.618	0.769	0.782	0.840	0.852	0.833
Region Fixed-Effects	Yes	Yes	Yes	Yes	Yes	No	No	No
Geography	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Family Fixed-Effe	No	No	No	Yes	No	Yes	Yes	Yes
Country Fixed-Effects	No	No	No	No	Yes	Yes	Yes	Yes
Observations	834	834	834	834	834	834	834	834

Table 5 - Notes

Panels A and B report Negative Binomial ML and OLS estimates respectively, associating civil conflict casualties with ethnic partitioning. The dependent variable is the number of casualties from civil conflict between 1970 and 2005 that have occurred in the historical homeland of an ethnic group divided by the respective tribal area. In the OLS specifications the dependent variable is the log of one plus the number of casualties per thousand of square kilometers, while in the Negative Binomial ML specifications the casualty rate is expressed in levels. In columns (1)-(6) we use the best estimate for casualties from the PRIO database, while in column (7) and (8) we use the high and the low war casualties estimates, respectively. SPLIT is an indicator variable that identifies partitioning ethnicities as those with at least 10% of the historical homeland belonging to more than one contemporary country. BORDER is an indicator that identifies ethnic groups residing by the border. These groups also fall into more than one country but more than 90% of the historical homeland lies in one country.

In columns (2)-(5) we control for spatial spillovers including the number of civil war casualties per thousand kilometers at the country-level and the ethnic-family level (excluding ethnicity i), respectively. All specifications include a set of region fixed-effects, log land area, log land area under water (lakes, rivers, and other streams), and population density around independence (in 1960). Columns (3)-(8) include a regression discontinuity (RD) cubic polynomial in latitude and longitude of the centroid of each ethnic group. Columns (4)-(8) include a rich set of control variables (land suitability for agriculture, elevation, malaria stability index, an early development indicator whether a major city was in the ethnicity's historical homeland in 1400, an oil indicator and a diamond mine indicator). Columns (4), (6), (7), and (8) include a set of ethnic-family fixed-effects (constants not reported). Columns (5)-(8) include a set of country fixed-effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension.

***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 6 - Partitioning and Civil War Duration

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Negative Binomial ML			
SPLIT	0.3206*	0.3807**	0.3837**	0.3129**	0.4225**	0.3055**
Partitioning (>10%)	(0.1787)	(0.1702)	(0.1527)	(0.1265)	(0.1814)	(0.1453)
BORDER	0.1958	0.1894*	0.1601**	0.1737**	0.1600**	0.1729*
Partitioning (0% - 10%)	(0.1478)	(0.1064)	(0.0758)	(0.0862)	(0.0797)	(0.0935)
Total Duration at Country		0.0912***	0.0763***	0.0571***		
		(0.0197)	(0.0099)	(0.0159)		
Total Duration at Ethnic Family		0.0834***	0.0696***		0.0558***	
		(0.0155)	(0.0117)		(0.0121)	
Log Likelihood	-2377.90	-2195.90	-2088.96	-1884.05	-1972.15	-1758.84
			Panel B: Log Linear OLS			
SPLIT	0.3039**	0.2724***	0.2927***	0.2763***	0.2673***	0.2585**
Partitioning (>10%)	(0.1258)	(0.0819)	(0.0756)	(0.0913)	(0.1001)	(0.1048)
BORDER	0.2240	0.0912*	0.0834	0.0721	0.0523	0.0909
Partitioning (0% - 10%)	(0.1377)	(0.0485)	(0.0520)	(0.0904)	(0.0615)	(0.0898)
Total Duration at Country		0.5355***	0.5383***	0.4484***		
		(0.0770)	(0.0722)	(0.0997)		
Total Duration at Ethnic Family		0.6210***	0.5240***		0.4251***	
		(0.0575)	(0.0517)		(0.0736)	
adjusted R-squared	0.235	0.678	0.717	0.794	0.791	0.847
Region Fixed-Effects	Yes	Yes	Yes	Yes	No	No
Geography	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	No	Yes	Yes	Yes	Yes
Additional Controls	No	No	Yes	Yes	Yes	Yes
Ethnic Family Fixed-Effects	No	No	No	Yes	No	Yes
Country Fixed-Effects	No	No	No	No	Yes	Yes
Observations	834	834	834	834	834	834

Table 6 - Notes

Panels A and B report Negative Binomial ML and OLS estimates respectively, associating civil war duration with ethnic partitioning. The dependent variable is the total number of years that an area of an ethnic group has been under civil conflict over the period 1970-2005. In the OLS specifications the dependent variable is the log of one plus the number of number of years under conflict, while in the Negative Binomial ML civil war duration is expressed in levels. In columns (2)-(5) we control for spatial spillovers including the civil war duration at the country-level and the ethnic-family level (excluding ethnicity i), respectively. SPLIT is an indicator variable that identifies partitioning ethnicities as those with at least 10% of the historical homeland belonging to more than one contemporary country. BORDER is an indicator that identifies ethnic groups residing by the border. These groups also fall into more than one country but more than 90% of the historical homeland lies in one country.
All specifications include a set of region fixed-effects, log land area, log land area under water (lakes, rivers, and other streams), and population density around independence (in 1960). Columns (3)-(6) include a regression discontinuity (RD) cubic polynomial in latitude and longitude of the centroid of each ethnic group. Columns (4)-(6) include a rich set of control variables (land suitability for agriculture, elevation, malaria stability index, an early development indicator whether a major city was in the ethnicity's historical homeland in 1400, an oil indicator and a diamond mine indicator). Columns (4) and (6) include a set of ethnic-family fixed-effects (constants not reported). Columns (5) and (6) include a set of country fixed-effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 7 - Partitioning and Development (Satellite Light Density)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Panel A: Negative Binomial ML			
SPLIT	-0.7827***	-0.7018***	-0.8353***	-0.6913***	-0.4856***	-0.4000***	-0.3854***
Partitioning (>10%)	(0.6591)	(0.1737)	(0.1807)	(0.1275)	(0.1590)	(0.1060)	(0.1198)
BORDER	-0.7854***	-0.4533**	-0.3907**	-0.1334	-0.0959	-0.3124***	-0.2033*
Partitioning (0% - 10%)	(0.2242)	(0.1821)	(0.2050)	(0.1365)	(0.1345)	(0.0524)	(0.0849)
Log Likelihood	-478.964	-438.123	-393.55	-339.236	-321.435	-284.317	-309.682
				Panel B: Log Linear OLS			
SPLIT	-0.1048***	-0.0677**	-0.0756**	-0.0604**	-0.0536*	-0.0306	-0.0422
Partitioning (>10%)	(0.0375)	(0.0337)	(0.0309)	(0.0283)	(0.0299)	(0.0241)	(0.0271)
BORDER	-0.1168***	-0.0825**	-0.0709**	-0.0674**	-0.0665**	-0.0585**	-0.0616***
Partitioning (0% - 10%)	(0.0305)	(0.0330)	(0.0276)	(0.0283)	(0.0261)	(0.0230)	(0.0239)
adjusted R-squared	0.399	0.478	0.563	0.730	0.764	0.751	0.785
Region Fixed-Effects	Yes	Yes	Yes	Yes	No	No	No
Geography	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	No	No	Yes	Yes	Yes	Yes	Yes
Ethnic Family Fixed-Effects	No	No	No	Yes	Yes	Yes	Yes
Country Fixed-Effects	No	No	No	No	Yes	Yes	Yes
Observations	834	834	834	834	834	825	834

Panels A and B report Negative Binomial ML and OLS estimates respectively associating regional development, as reflected in satellite light density at night, with ethnic partitioning. In the OLS specifications the dependent variable is the log of one plus light density, while in the Negative Binomial ML satellite light density is expressed in levels. SPLIT is an indicator variable that identifies partitioning ethnicities as those with at least 10% of the historical homeland belonging to more than one contemporary country. BORDER is an indicator that identifies ethnic groups residing by the border. These groups also fall into more than one country but more than 90% of the historical homeland lies in one country.

All specifications include a set of region fixed-effects, log land area, log land area under water (lakes, rivers, and other streams), and population density around independence (in 1960). Columns (2)-(7) include a regression discontinuity (RD) cubic polynomial in latitude and longitude of the centroid of each ethnic group. Columns (3)-(7) include a rich set of control variables (land suitability for agriculture, elevation, malaria stability index, an early development indicator whether a major city was in the ethnicity's historical homeland in 1400, an oil indicator and a diamond mine indicator). Columns (4)-(7) include a set of ethnic-family fixed-effects (constants not reported). Columns (5)-(7) include a set of country fixed-effects (constants not reported). In column (6) we drop observations at the top 1% of the distribution on light density, while in column (7) we winsorise the dependent variable at the 1%. The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Appendix Table A - Sensitivity Analysis
Spatial Lag Specifications

	Casualties		Duration		Development	
	(1)	(2)	(3)	(4)	(5)	(6)
SPLIT	0.5243***	0.5029***	0.2677***	0.2813***	-0.0942***	-0.0647***
Partitioning (>10%)	(0.1114)	(0.1025)	(0.0754)	(0.0658)	(0.0270)	(0.0240)
BORDER	*0.3702***	*0.2731***	*0.1564**	*0.1188*	*-0.0951****	*-0.0599***
Partitioning (0% - 10%)	(0.1351)	(0.1233)	(0.0919)	(0.0792)	(0.0329)	(0.0289)
rho (spatial lag)	1.658	1.642	1.659	1.645	1.603	1.575
	(0.015)	(0.031)	(0.015)	(0.028)	(0.070)	(0.097)
chi2	447.17	226.47	343.96	237.90	106.54	82.52
Log Likelihood	-1406.76	-1318.38	-1085.07	-949.07	-227.30	-105.94
Region Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
Geography	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	Yes	No	Yes	No	Yes
Additional Controls	No	Yes	No	Yes	No	Yes
Observations	834	834	834	834	834	834

The table reports general spatial lag specifications, associating partitioning with civil war intensity (in (1)-(2)), civil war duration (in (3)-(4)), and regional development as proxied by satellite light density (in (5)-(6)). The spatial lag is decaying linearly in the distance from the centroid of each ethnic group. Estimation is with maximum likelihood. The table also reports the coefficient and the corresponding standard error of the spatial lag (rho). SPLIT is an indicator variable that identifies partitioning ethnicities as those with at least 10% of the historical homeland belonging to more than one contemporary country. BORDER is an indicator that identifies ethnic groups residing by the border. These groups also fall into more than one country but more than 90% of the historical homeland lies in one country.

All specifications include a set of region fixed-effects, log land area, log land area under water (lakes, rivers, and other streams), and population density around independence (in 1960). Columns (2), (4), and (6) include a regression discontinuity (RD) cubic polynomial in latitude and longitude of the centroid of each ethnic group and a rich set of control variables (land suitability for agriculture, elevation, a malaria stability index, an early development indicator whether a major city was in the ethnicity's historical homeland in 1400, an oil indicator and a diamond mine indicator). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for spatial correlation. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Appendix Table B - Sensitivity Analysis
Excluding Northern African and Southern African Countries

	Casualties		Duration		Development	
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Excluding Northern Africa Countries					
SPLIT	0.6483***	0.4832***	0.4338***	0.2319**	-0.9342***	-0.4144***
Partitioning (>10%)	(0.1762)	(0.1300)	(0.1495)	(0.1019)	(0.2080)	(0.1359)
BORDER	0.4699***	0.3270**	0.2374	0.1786**	-0.5529***	-0.2811**
Partitioning (0% - 10%)	(0.1642)	(0.1546)	(0.1527)	(0.0815)	(0.1706)	(0.0734)
Log Likelihood	-2700.54	-2331.48	-2051.15	-1582.62	-294.75	-234.46
Observations	770	770	770	770	770	770
	Panel B: Excluding Southern Africa Countries					
SPLIT	0.6496***	0.6957***	0.3904***	0.3530**	-0.9189***	-0.4100**
Partitioning (>10%)	(0.1391)	(0.1869)	(0.1462)	(0.1545)	(0.2185)	(0.1689)
BORDER	0.3539**	0.2939***	0.1369	0.1856***	-0.3840**	-0.0684
Partitioning (0% - 10%)	(0.1720)	(0.1313)	(0.1165)	(0.0794)	(0.2314)	(0.1481)
Log Likelihood	-2516.23	-2170.84	-1839.07	-1479.81	-329.33	-270.56
Observations	747	747	747	747	747	747
Region Fixed-Effects	Yes	No	Yes	No	Yes	No
Geography	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed-Effects	No	Yes	No	Yes	No	Yes
Ethnic Family Fixed-Effects	No	Yes	No	Yes	No	Yes

The table reports negative binomial ML specifications, associating partitioning with civil war intensity (in (1)-(2)), civil war duration (in (3)-(4)), and regional development as proxied by satellite light density (in (5)-(6)). In Panel A we exclude ethnic areas in Northern Africa, while in Panel B we exclude ethnic areas in Southern Africa. SPLIT is an indicator variable that identifies partitioning ethnicities as those with at least 10% of the historical homeland belonging to more than one contemporary country. BORDER is an indicator that identifies ethnic groups residing by the border. These groups also fall into more than one country but more than 90% of the historical homeland lies in one country.

All specifications include log land area, log land area under water (lakes, rivers, and other streams), population density around independence (in 1960), a regression discontinuity (RD) cubic polynomial in latitude and longitude of the centroid of each ethnic group, land suitability for agriculture, elevation, malaria stability index, an early development indicator whether a major city was in the ethnicity's historical homeland in 1400, an oil indicator and a diamond mine indicator. Columns (1), (3), and (5) include a set of region fixed-effects (constants not reported). Columns (2), (4), and (6) include a set of ethnic-family fixed-effects and a set of country fixed-effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Appendix Table C - Sensitivity Analysis
Employing a Continuous Measure of Ethnic Partitioning

	Casualties		Duration		Development	
	(1)	(2)	(3)	(4)	(5)	(6)
FRACT	1.1693**	1.0899**	0.7731**	0.6302**	-1.5200***	-1.2020***
Continuous Partitioning	(0.4901)	(0.4176)	(0.3290)	(0.3008)	(0.5043)	(0.3930)
Log Likelihood	-2963.87	-2562.32	-2240.25	-1761.59	-406.13	-321.32
Observations	834	834	834	834	834	834
Region Fixed-Effects	Yes	No	Yes	No	Yes	No
Geography	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed-Effects	No	Yes	No	Yes	No	Yes
Ethnic Family Fixed-Effects	No	Yes	No	Yes	No	Yes

The table reports negative binomial ML specifications, associating a contninuous measure of ethnic partitioning with civil war intensity (in (1)-(2)), civil war duration (in (3)-(4)), and regional development as proxied by satellite light density (in (5)-(6)).

All specifications include log land area, log land area under water (lakes, rivers, and other streams), population density around independence (in 1960), a regression discontinuity (RD) cubic polynomial in latitude and longitude of the centroid of each ethnic group, land suitability for agriculture, elevation, malaria stability index, an early development indicator whether a major city was in the ethnicity's historical homeland in 1400, an oil indicator and a diamond mine indicator. Columns (1), (3), and (5) include a set of region fixed-effects (constants not reported). Columns (2), (4), and (6) include a set of ethnic-family fixed-effects and a set of country fixed-effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Appendix Table 1 - Partitioned Ethnicities and Countries they Belong to

Ethnicity Name	% of Initial Homeland	Country	# of Partitions	Ethnicity Name	% of Initial Homeland	Country	# of Partitions
ABABDA	0.72	EGY	2	LAKA (ADAMAWA)	0.69	TCD	3
ABABDA	0.28	SDN	2	LAKA (ADAMAWA)	0.20	CMR	3
ADELE	0.48	GHA	2	LAKA (ADAMAWA)	0.11	CAF	3
ADELE	0.52	TGO	2	LAMBA	0.39	ZAR	2
AFAR	0.17	DJI	3	LAMBA	0.61	ZMB	2
AFAR	0.22	ERI	3	LAMBYA	0.17	MWI	3
AFAR	0.61	ETH	3	LAMBYA	0.33	TZA	3
ALUR	0.16	ZAR	2	LAMBYA	0.50	ZMB	3
ALUR	0.84	UGA	2	LIGBI, DEGHA (SE)	0.72	GHA	2
AMBA	0.87	ZAR	2	LIGBI, DEGHA (SE)	0.28	CIV	2
AMBA	0.13	UGA	2	LOBI	0.42	CIV	2
AMBO	0.41	AGO	2	LOBI	0.58	BFA	2
AMBO	0.59	NAM	2	LUGBARA	0.45	ZAR	3
AMER	0.56	ERI	2	LUGBARA	0.04	SDN	3
AMER	0.44	SDN	2	LUGBARA	0.51	UGA	3
ANA	0.33	BEN	2	LUNGU	0.31	TZA	2
ANA	0.67	TGO	2	LUNGU	0.69	ZMB	2
ANUAK	0.75	ETH	2	LUVALE	0.81	AGO	3
ANUAK	0.25	SDN	2	LUVALE	0.01	ZAR	3
ANYI	0.42	GHA	2	LUVALE	0.17	ZMB	3
ANYI	0.58	CIV	2	MADI	0.42	SDN	2
ASBEN	0.89	NER	2	MADI	0.58	UGA	2
ASBEN	0.11	DZA	2	MAKONDE	0.56	MOZ	2
ASSINI	0.51	GHA	2	MAKONDE	0.44	TZA	2
ASSINI	0.49	CIV	2	MALINKE	0.03	GMB	6
ATTA	0.51	MAR	2	MALINKE	0.13	CIV	6
ATTA	0.49	DZA	2	MALINKE	0.27	MLI	6
ATYUTI	0.13	GHA	2	MALINKE	0.04	GNB	6
ATYUTI	0.87	TGO	2	MALINKE	0.25	GIN	6
AULLIMINDEN	0.55	MLI	3	MALINKE	0.29	SEN	6
AULLIMINDEN	0.40	NER	3	MAMBILA	0.57	CMR	2
AULLIMINDEN	0.05	DZA	3	MAMBILA	0.43	NGA	2
AUSHI	0.27	ZAR	2	MANDARA	0.35	CMR	2
AUSHI	0.73	ZMB	2	MANDARA	0.65	NGA	2
AVATIME	0.51	GHA	2	MANGA	0.60	NER	2
AVATIME	0.49	TGO	2	MANGA	0.40	NGA	2
AZANDE	0.62	ZAR	3	MANYIKA	0.39	MOZ	2
AZANDE	0.15	CAF	3	MANYIKA	0.61	ZWE	2
AZANDE	0.23	SDN	3	MASAI	0.38	KEN	2
AZJER	0.24	LBY	3	MASAI	0.62	TZA	2
AZJER	0.00	NER	3	MASALIT	0.13	TCD	2
AZJER	0.75	DZA	3	MASALIT	0.87	SDN	2

BABUKUR	0.82	ZAR	2	MASHI	0.12	AGO	2
BABUKUR	0.18	SDN	2	MASHI	0.88	ZMB	2
BAJUN	0.37	KEN	2	MASINA	0.82	MLI	3
BAJUN	0.63	SOM	2	MASINA	0.09	BFA	3
BALANTE	0.73	GNB	2	MASINA	0.09	MRT	3
BALANTE	0.27	SEN	2	MATAKAM	0.70	CMR	2
BANYUN	0.48	GNB	2	MATAKAM	0.30	NGA	2
BANYUN	0.52	SEN	2	MBERE	0.02	TCD	3
BANZIRI	0.14	ZAR	2	MBERE	0.24	CMR	3
BANZIRI	0.86	CAF	2	MBERE	0.74	CAF	3
BARABRA	0.31	EGY	2	MBUKUSHU	0.74	AGO	3
BARABRA	0.69	SDN	2	MBUKUSHU	0.15	BWA	3
BARARETTA	0.18	ETH	3	MBUKUSHU	0.12	NAM	3
BARARETTA	0.44	KEN	3	MBUNDA	0.89	AGO	2
BARARETTA	0.38	SOM	3	MBUNDA	0.11	ZMB	2
BARGU	0.77	BEN	4	MENDE	0.18	LBR	3
BARGU	0.03	NER	4	MENDE	0.82	SLE	3
BARGU	0.19	NGA	4	MINIANKA	0.01	CIV	3
BARGU	0.02	BFA	4	MINIANKA	0.72	MLI	3
BASHI	0.09	BDI	3	MINIANKA	0.27	BFA	3
BASHI	0.83	ZAR	3	MOMBERA	0.72	MWI	2
BASHI	0.08	RWA	3	MOMBERA	0.28	ZMB	2
BATA	0.29	CMR	2	MPEZENI	0.11	MWI	2
BATA	0.71	NGA	2	MPEZENI	0.89	ZMB	2
BAYA	0.20	CMR	2	MUNDANG	0.80	TCD	2
BAYA	0.80	CAF	2	MUNDANG	0.20	CMR	2
BERABISH	0.80	MLI	2	MUNDU	0.30	ZAR	2
BERABISH	0.20	MRT	2	MUNDU	0.70	SDN	2
BERTA	0.75	ETH	2	MUSGU	0.76	TCD	2
BERTA	0.25	SDN	2	MUSGU	0.24	CMR	2
BIDEYAT	0.21	LBY	4	NAFANA	0.74	GHA	2
BIDEYAT	0.40	TCD	4	NAFANA	0.26	CIV	2
BIDEYAT	0.03	EGY	4	NALU	0.41	GNB	2
BIDEYAT	0.36	SDN	4	NALU	0.59	GIN	2
BIRIFON	0.52	GHA	3	NAMA	0.18	ZAF	2
BIRIFON	0.47	BFA	3	NAMA	0.82	NAM	2
BOBO	0.20	MLI	2	NAUDEBA	0.87	BEN	2
BOBO	0.80	BFA	2	NAUDEBA	0.13	TGO	2
BOKI	0.22	CMR	2	NDAU	0.86	MOZ	2
BOKI	0.78	NGA	2	NDAU	0.14	ZWE	2
BONDJO	0.14	ZAR	2	NDEMBU	0.26	AGO	3
BONDJO	0.86	COG	2	NDEMBU	0.39	ZAR	3
BONI	0.67	KEN	2	NDEMBU	0.35	ZMB	3
BONI	0.33	SOM	2	NDOGO	0.01	ZAR	3
BORAN	0.46	ETH	2	NDOGO	0.18	CAF	3
BORAN	0.54	KEN	2	NDOGO	0.81	SDN	3
BRONG	0.84	GHA	2	NDUKA	0.23	TCD	2
BRONG	0.16	CIV	2	NDUKA	0.77	CAF	2
BUEM	0.40	GHA	2	NGAMA	0.30	TCD	2

BUEM	0.60	TGO	2	NGAMA	0.70	CAF	2
BULOM	0.85	SLE	2	NGERE	0.65	CIV	3
BULOM	0.15	GIN	2	NGERE	0.29	LBR	3
BUSA	0.14	BEN	2	NGERE	0.06	GIN	3
BUSA	0.86	NGA	2	NGUMBA	0.65	CMR	2
BWAKA	0.81	ZAR	3	NGUMBA	0.35	GNQ	2
BWAKA	0.15	CAF	3	NGWAKETSE	0.86	BWA	2
BWAKA	0.04	COG	3	NGWAKETSE	0.14	ZAF	2
CHAGA	0.24	KEN	2	NSENGA	0.15	MOZ	3
CHAGA	0.76	TZA	2	NSENGA	0.78	ZMB	3
CHAKOSSI	0.27	GHA	2	NSENGA	0.06	ZWE	3
CHAKOSSI	0.73	TGO	2	NSUNGLI	0.78	CMR	2
CHEWA	0.34	MWI	3	NSUNGLI	0.22	NGA	2
CHEWA	0.50	MOZ	3	NUKWE	0.44	AGO	4
CHEWA	0.16	ZMB	3	NUKWE	0.24	BWA	4
CHIGA	0.12	RWA	3	NUKWE	0.05	ZMB	4
CHIGA	0.87	UGA	3	NUKWE	0.26	NAM	4
CHOKWE	0.81	AGO	2	NUSAN	0.30	BWA	3
CHOKWE	0.19	ZAR	2	NUSAN	0.37	ZAF	3
COMORIANS	0.82	COM	2	NUSAN	0.33	NAM	3
COMORIANS	0.18	MYT	2	NYAKYUSA	0.12	MWI	2
DAGARI	0.67	GHA	2	NYAKYUSA	0.88	TZA	2
DAGARI	0.33	BFA	2	NYANGIYA	0.17	SDN	2
DARI	0.78	TCD	2	NYANGIYA	0.83	UGA	2
DARI	0.22	CMR	2	NYANJA	0.64	MWI	2
DAZA	0.27	TCD	2	NYANJA	0.36	MOZ	2
DAZA	0.73	NER	2	NYASA	0.05	MWI	3
DELIM	0.55	ESH	2	NYASA	0.68	MOZ	3
DELIM	0.45	MRT	2	NYASA	0.27	TZA	3
DENDI	0.60	BEN	3	NZANKARA	0.14	ZAR	2
DENDI	0.39	NER	3	NZANKARA	0.86	CAF	2
DIALONKE	0.36	MLI	3	PANDE	0.38	CAF	2
DIALONKE	0.58	GIN	3	PANDE	0.62	COG	2
DIALONKE	0.06	SEN	3	POPO	0.72	BEN	2
DIDINGA	0.04	KEN	3	POPO	0.28	TGO	2
DIDINGA	0.89	SDN	3	PUKU	0.31	CMR	3
DIDINGA	0.07	UGA	3	PUKU	0.49	GNQ	3
DIGO	0.62	KEN	2	PUKU	0.19	GAB	3
DIGO	0.38	TZA	2	REGEIBAT	0.34	ESH	2
DIOLA	0.14	GMB	3	REGEIBAT	0.66	MRT	2
DIOLA	0.07	GNB	3	RESHIAT	0.83	ETH	3
DIOLA	0.78	SEN	3	RESHIAT	0.06	KEN	3
DUMA	0.63	GAB	2	RESHIAT	0.11	SDN	3
DUMA	0.37	COG	2	RONGA	0.60	MOZ	3
DZEM	0.74	CMR	3	RONGA	0.35	ZAF	3
DZEM	0.03	GAB	3	RONGA	0.05	SWZ	3
DZEM	0.24	COG	3	RUANDA	0.02	BDI	5
EGBA	0.41	BEN	3	RUANDA	0.06	ZAR	5
EGBA	0.52	NGA	3	RUANDA	0.89	RWA	5

EGBA	0.07	TGO	3	RUANDA	0.02	TZA	5
EKOI	0.38	CMR	2	RUANDA	0.02	UGA	5
EKOI	0.62	NGA	2	RUNDI	0.76	BDI	4
ESA	0.03	DJI	3	RUNDI	0.04	RWA	4
ESA	0.52	ETH	3	RUNDI	0.20	TZA	4
ESA	0.44	SOM	3	RUNGA	0.74	TCD	3
EWE	0.44	GHA	2	RUNGA	0.26	CAF	3
EWE	0.56	TGO	2	SABEI	0.56	KEN	2
FANG	0.37	CMR	4	SABEI	0.44	UGA	2
FANG	0.07	GNQ	4	SAHO	0.43	ERI	2
FANG	0.54	GAB	4	SAHO	0.57	ETH	2
FANG	0.02	COG	4	SAMO	0.12	MLI	2
FON	0.86	BEN	3	SAMO	0.88	BFA	2
FON	0.14	TGO	3	SANGA	0.26	CMR	3
FOUTADJALON	0.01	MLI	4	SANGA	0.19	CAF	3
FOUTADJALON	0.11	GNB	4	SANGA	0.55	COG	3
FOUTADJALON	0.88	GIN	4	SEKE	0.34	GNQ	2
FOUTADJALON	0.01	SEN	4	SEKE	0.66	GAB	2
FUNGON	0.81	CMR	2	SHAMBALA	0.10	KEN	2
FUNGON	0.19	NGA	2	SHAMBALA	0.90	TZA	2
GADAMES	0.25	LBY	3	SHEBELLE	0.58	ETH	2
GADAMES	0.27	TUN	3	SHEBELLE	0.42	SOM	2
GADAMES	0.48	DZA	3	SHUWA	0.62	TCD	3
GIL	0.80	MAR	2	SHUWA	0.17	CMR	3
GIL	0.20	DZA	2	SHUWA	0.21	NGA	3
GOMANI	0.86	MWI	2	SONGHAI	0.57	MLI	3
GOMANI	0.14	MOZ	2	SONGHAI	0.36	NER	3
GREBO	0.33	CIV	2	SONGHAI	0.07	BFA	3
GREBO	0.67	LBR	2	SONINKE	0.68	MLI	3
GRUNSHI	0.68	GHA	2	SONINKE	0.03	SEN	3
GRUNSHI	0.32	BFA	2	SONINKE	0.29	MRT	3
GUDE	0.83	CMR	2	SOTHO	0.24	LSO	2
GUDE	0.17	NGA	2	SOTHO	0.76	ZAF	2
GULA	0.61	TCD	2	SUBIA	0.11	BWA	4
GULA	0.39	CAF	2	SUBIA	0.53	ZMB	4
GUN	0.48	BEN	2	SUBIA	0.06	ZWE	4
GUN	0.52	NGA	2	SUBIA	0.30	NAM	4
GURENSI	0.74	GHA	3	SUNDI	0.37	ZAR	2
GURENSI	0.13	TGO	3	SUNDI	0.63	COG	2
GURENSI	0.13	BFA	3	SURI	0.71	ETH	2
GURMA	0.15	BEN	4	SURI	0.29	SDN	2
GURMA	0.12	NER	4	SWAZI	0.45	ZAF	2
GURMA	0.01	TGO	4	SWAZI	0.55	SWZ	2
GURMA	0.72	BFA	4	TABWA	0.57	ZAR	2
GUSII	0.53	KEN	2	TABWA	0.43	ZMB	2
GUSII	0.47	TZA	2	TAJAKANT	0.15	MAR	4
HAMAMA	0.80	TUN	2	TAJAKANT	0.14	ESH	4
HAMAMA	0.20	DZA	2	TAJAKANT	0.66	DZA	4
HAUSA	0.14	NER	2	TAJAKANT	0.05	MRT	4

HAUSA	0.86	NGA	2	TAMA	0.30	TCD	2
HIECHWARE	0.81	BWA	2	TAMA	0.70	SDN	2
HIECHWARE	0.19	ZWE	2	TAWARA	0.57	MOZ	2
HLENGWE	0.82	MOZ	3	TAWARA	0.43	ZWE	2
HLENGWE	0.00	ZAF	3	TEDA	0.34	LBY	3
HLENGWE	0.18	ZWE	3	TEDA	0.35	TCD	3
HOLO	0.84	AGO	2	TEDA	0.31	NER	3
HOLO	0.16	ZAR	2	TEKE	0.31	ZAR	3
IBIBIO	0.11	CMR	2	TEKE	0.03	GAB	3
IBIBIO	0.89	NGA	2	TEKE	0.66	COG	3
IFORA	0.30	MLI	2	TEKNA	0.53	MAR	2
IFORA	0.70	DZA	2	TEKNA	0.47	ESH	2
IMRAGEN	0.10	MAR	3	TEM	0.17	BEN	2
IMRAGEN	0.74	ESH	3	TEM	0.83	TGO	2
IMRAGEN	0.16	MRT	3	TENDA	0.57	GIN	2
ISHAAK	0.20	ETH	2	TENDA	0.43	SEN	2
ISHAAK	0.80	SOM	2	THONGA	0.58	MOZ	3
IWA	0.33	TZA	2	THONGA	0.42	ZAF	3
IWA	0.67	ZMB	2	TIENGA	0.22	NER	3
JERID	0.90	TUN	2	TIENGA	0.78	NGA	3
JERID	0.10	DZA	2	TIGON	0.32	CMR	2
JIE	0.24	KEN	2	TIGON	0.68	NGA	2
JIE	0.76	UGA	2	TIGRINYA	0.51	ERI	3
KABRE	0.39	BEN	2	TIGRINYA	0.44	ETH	3
KABRE	0.61	TGO	2	TIGRINYA	0.05	SDN	3
KANEMBU	0.73	TCD	3	TLOKWA	0.14	BWA	3
KANEMBU	0.25	NER	3	TLOKWA	0.77	ZAF	3
KANEMBU	0.02	NGA	3	TLOKWA	0.09	ZWE	3
KAONDE	0.21	ZAR	2	TOMA	0.29	LBR	2
KAONDE	0.79	ZMB	2	TOMA	0.71	GIN	2
KAPSIKI	0.65	CMR	2	TONGA	0.84	ZMB	2
KAPSIKI	0.35	NGA	2	TONGA	0.16	ZWE	2
KARA	0.85	CAF	2	TRIBU	0.25	GHA	2
KARA	0.15	SDN	2	TRIBU	0.75	TGO	2
KARAMOJONG	0.27	KEN	2	TRIPOLITANIANS	0.74	LBY	2
KARAMOJONG	0.73	UGA	2	TRIPOLITANIANS	0.26	TUN	2
KARE	0.75	ZAR	2	TUBURI	0.25	TCD	2
KARE	0.25	CAF	2	TUBURI	0.75	CMR	2
KGATLA	0.13	BWA	2	TUKULOR	0.39	SEN	2
KGATLA	0.87	ZAF	2	TUKULOR	0.61	MRT	2
KISSI	0.12	LBR	3	TUMBUKA	0.74	MWI	2
KISSI	0.02	SLE	3	TUMBUKA	0.26	ZMB	2
KISSI	0.86	GIN	3	TUNISIANS	0.87	TUN	2
KOBA	0.89	BWA	2	TUNISIANS	0.13	DZA	2
KOBA	0.11	NAM	2	UDALAN	0.82	MLI	3
KOMA	0.57	ETH	2	UDALAN	0.05	NER	3
KOMA	0.43	SDN	2	UDALAN	0.13	BFA	3
KOMONO	0.49	CIV	2	VAI	0.76	LBR	2
KOMONO	0.51	BFA	2	VAI	0.24	SLE	2

KONGO	0.77	AGO	3	VENDA	0.70	ZAF	2
KONGO	0.23	ZAR	3	VENDA	0.30	ZWE	2
KONJO	0.81	ZAR	2	VILI	0.20	AGO	4
KONJO	0.19	UGA	2	VILI	0.22	ZAR	4
KONKOMBA	0.24	GHA	2	VILI	0.11	GAB	4
KONKOMBA	0.76	TGO	2	VILI	0.47	COG	4
KONO	0.74	SLE	2	WAKURA	0.28	CMR	2
KONO	0.26	GIN	2	WAKURA	0.72	NGA	2
KONYANKE	0.30	CIV	2	WANGA	0.79	KEN	2
KONYANKE	0.70	GIN	2	WANGA	0.21	UGA	2
KORANKO	0.39	SLE	2	WUM	0.88	CMR	2
KORANKO	0.61	GIN	2	WUM	0.12	NGA	2
KOTA	0.41	GAB	2	YAKA	0.16	AGO	2
KOTA	0.59	COG	2	YAKA	0.84	ZAR	2
KOTOKO	0.67	TCD	2	YAKOMA	0.40	ZAR	2
KOTOKO	0.33	CMR	2	YAKOMA	0.60	CAF	2
KPELLE	0.48	LBR	3	YALUNKA	0.25	SLE	2
KPELLE	0.52	GIN	3	YALUNKA	0.75	GIN	2
KRAN	0.16	CIV	2	YAO	0.13	MWI	3
KRAN	0.84	LBR	2	YAO	0.65	MOZ	3
KREISH	0.10	CAF	2	YAO	0.22	TZA	3
KREISH	0.90	SDN	2	YOMBE	0.13	AGO	3
KUNDA	0.84	MOZ	3	YOMBE	0.48	ZAR	3
KUNDA	0.15	ZMB	3	YOMBE	0.39	COG	3
KUNG	0.10	BWA	2	ZAGHAWA	0.14	TCD	2
KUNG	0.90	NAM	2	ZAGHAWA	0.86	SDN	2
KUNTA	0.85	MLI	2	ZEKARA	0.83	MAR	2
KUNTA	0.15	DZA	2	ZEKARA	0.17	DZA	2
KWANGARE	0.84	AGO	2	ZIMBA	0.16	MWI	2
KWANGARE	0.16	NAM	2	ZIMBA	0.84	MOZ	2

Appendix Table 1 reports the name of partitioned ethnic groups (as coded by Murdock (1959)) and the percentage of the historical homeland of the split ethnic groups that fall into more than one country. Section 2 in the paper gives details on our approach in identifying partitioned ethnicities.