THE LONG-RUN EFFECTS OF THE SCRAMBLE FOR AFRICA

Stelios Michalopoulos
Elias Papaioannou

The Long-Run Effects of the Scramble for Africa
Stelios Michalopoulos and Elias Papaioannou
NBER Working Paper No. 17620
November 2011, Revised June 2012
JEL No. N17,N47,O10,Z10

<u>ABSTRACT</u>

We examine the long-run consequences of the scramble for Africa among European powers in the late 19th century and uncover the following empirical regularities. First, utilizing information on the spatial distribution of African ethnicities before colonization, we show that borders were artificially drawn. Apart from the land mass and water area of an ethnicity's historical homeland, no other geographic, economic, and historical trait predicts partitioning by the national borders. Second, we exploit a detailed geo-referenced database on various types of conflict across African regions and show that civil conflict is concentrated in the historical homeland of partitioned ethnicities. We further document that violence against civilians and territorial changes between rebel groups, militias, and government forces are systematically higher in the homelands of split groups. These results are robust to a rich set of local controls, the inclusion of country fixed effects, and alternative data sources. The uncovered evidence thus identifies a sizable causal impact of the scramble for Africa on contemporary political violence and conflict.

Stelios Michalopoulos
Brown University
Department of Economics
64 Waterman Street
Providence, RI 02912
and NBER
smichalo@brown.edu

Elias Papaioannou
Department of Economics
Dartmouth College
6106 Rockefeller Hall
Hanover, NH 03755
and NBER
papaioannou.elias@gmail.com

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), Nunn and Wantchekon (2011)). Yet in the period between the ending of the slave trades and the colonial rule, another major event took place in 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 were designed in European capitals at a time when Europeans had barely settled in Africa and had little knowledge of local conditions. Despite their arbitrariness these boundaries endured after African independence in the 1960s. As a result in many African countries a significant fraction of the population belongs to ethnic groups that have been partitioned by the national border.[1] A considerable body of work in African historiography (e.g. Asiwaju (1985); Dowden (2008); Wesseling (1996); Herbst (2000)) argues that the main channel of Europeans' influence on African development was not colonization per se, but the improper border design. Partitioning, the argument goes, has led to ethnic struggles, patronage politics, and conflict. Yet there is little work that formally examines the impact of ethnic partitioning.[2]

This study is a first step to empirically assess the long-run effects of the scramble for Africa. While there is little disagreement among historians that colonial borders were arbitrarily drawn, we start our analysis establishing formally their artificiality. With the sole exceptions of the size of the historical homeland and area under water, we are unable to detect any other significant differences between partitioned and non-partitioned ethnicities with respect to geography (elevation, distance to the coast, soil quality), the disease environment (malaria), natural resources (diamond mines and oil fields), and measures of early contact with colonizers. We further show that there are no systematic differences between split and non-split groups, across several pre-colonial ethnic-specific institutional, cultural, and economic features, such as the size of settlements, the type of the subsistence economy, etc. (Murdock (1967)).

We then employ the scramble for Africa as a quasi-natural experiment and assess the

---

[1] Asiwaju (1985) identifies **177** partitioned ethnic groups that span all African borders. Englebert, Tarango, and Carter (2002) estimates that partitioned ethnic groups constitute on average 40% of the total population. Alesina, Easterly, and Matuszeski (2011) estimate that in several African countries the percentage 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] The cross-country studies of Alesina, Easterly, and Matuszeski (2011) and Englebert, Tarango, and Carter (2002) do touch upon this issue. We discuss the relationship of our work with these studies below.

impact of ethnic partitioning on regional civil conflict. Using a new rich dataset that reports detailed geo-referenced information for $1997-2010$ on the exact location of more than $43,000$ incidents of political violence including battles between government forces, rebel groups and militias, changes of territorial control, as well as violence against civilians (the latter includes murders, abductions, child soldiering raids, rapes, mutilations), we show that civil conflict is concentrated in the historical homeland of partitioned ethnicities. Our regional focus allows us to explore within-country variation, accounting for the numerous country-wide factors that interact with civil conflict. Moreover, we obtain similar results when we restrict estimation to ethnic areas close to national borders. Our most conservative estimates suggest that civil conflict intensity is approximately 50% higher in areas where partitioned ethnicities reside as compared to the homelands of ethnic groups that have not been separated by the national borders. We further find that homelands of partitioned groups experience a 5% to 10% higher likelihood of a territorial control change between the government and rebel groups. It is not only army fighting that is concentrated in the homelands of partitioned groups. Violence against civilians is also roughly 40% higher where split groups reside. The evidence thus uncovers the ongoing violent repercussions of the colonial border design.

Historical Background

The "Scramble for Africa" starts in 1860s - 1870s when the French and the British begin the systematic exploration of Western Africa and sign bilateral agreements assigning to each other spheres of influence. In the next 30 years, European powers signed hundreds of treaties that partitioned the largely unexplored continent into protectorates, free-trade areas, and colonies. 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 discussed only 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.[3] The key consideration was to preserve the "status quo" preventing conflict among Europeans for Africa (as the memories of the European wars of the 18th-19th century were still alive). As a result, European powers drew borders without taking into account local conditions and the ethnic composition. African leaders were not invited and had

---

[3]Three major principles emerged from the Berlin Conference. 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 had to base their claim on treaties with local tribal 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 to the coastline.

no say.[4] European leaders were in such a rush that they didn't wait for the new information arriving from explorers, geographers, and missionaries.

There is wide agreement among African historians that border design was to a great extent arbitrary (see Asiwaju (1985) and Englebert (2009) for references). As the British prime minister at the time Lord Salisbury put it, "*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.*" Asiwaju (1985) summarizes that "*the study of European archives supports the accidental rather than a conspiratorial theory of the marking of African boundaries.*"[5] In line with the historical evidence, Alesina, Easterly, and Matuszeski (2011) document that eighty percent of African borders follow latitudinal and longitudinal lines, more than in any other part of the world.

Several factors have been proposed to rationalize the arbitrary border design. First, at the time Europeans had limited knowledge of local geographic conditions, as with the exception of some coastal areas, the continent was largely unexplored. Second, Europeans were not drawing borders of prospective states or -in many cases- even colonies. Third, there was a constant imperialist back and forth with European powers swapping pieces of land with limited (at best) idea of what they were worth of.[6] Fourth, while in most cases the treaties indicated that the exact boundaries would be set by special commissions, demarcation was poor. Fifth, Europeans were not willing to sacrifice their commitment not to go to war for any part of Africa.[7] 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, as there was an implicit agreement between Europeans that ethnicities could freely move across colonial borders, African leaders did not oppose the colonial design,

---

[4] 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*".

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

[6] An illustrative example is the annexation of Katanga in Congo Free State that turned out to be the richest province. King Leopold demanded and eventually got Katanga in exchange for the Niari-Kwilu area that the French insisted of getting themselves. Wesseling (1996) writes "*what impelled him [Leopold] was a general imperialist surge, the desire for compensation for the Niari-Kwilu, and the objective of making the new state as large as possible and filling as much of the Congo basin as possible.*"

[7] 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.*"

as little changed on the ground.[8]

The other major event in recent African history, the wave of independence, was also rapid. The independence of Northern African countries in the 1950s was soon followed by Ghana's and Guinea's 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 did not touch the issue. The leaders of African independence believed that nation building and industrialization would sideline ethnic divisions. Europeans' main objective was to maintain their special rights and corporate deals with former colonies, and as such, they were reluctant to open the border issue.[9]

Case Studies - Channels

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

First, in several 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. The five-pointed star in the flag of Somalia symbolizes the five regions inhabited by Somali tribes (Somalia, North Kenya, Southern Ethiopia, Djibouti, and Eritrea); at least three long-lasting post-independence wars have been (partly at least) driven by the desire of Somalis in Ethiopia, Djibouti, and Kenya to become part of Somalia (e.g. Meredith (2005)). In line with this, in our sample that covers the period $1997 - 2010$, the bulk of battles and violent events against civilians have taken place in the Ogaden region in Southern Ethiopia where Somali tribes reside. Specifically, in Ethiopia (that in Murdock's map comprises 48 ethnic homelands), 33% of a total of 961 battles between government forces, rebel groups and militias as well as 19% of 295 violent events against civilians occurred in the Ogaden region where the partitioned Afar and the other Somali tribes are located.

Second, partitioned ethnicities may fight to gain independence or obtain autonomy. Wimmer, Cederman, and Min (2009) estimate that around 20% of all civil wars in Africa have a secessionist demand. Compared to non-split groups, partitioned ethnicities can get assistance from their peers on the other side of the border. An illustrative example is the recurring conflict in the Casamance region in Southern Senegal, where the partitioned Diola (Jola) reside. As Gambia effectively splits Senegal into a Northern and a Southern part, the Casamance province

---

[8]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.*"

[9]Almost all African countries accepted the colonial borders when signing the Charter of the Organization of African Union in 1964. Only Somalia and Morocco did not accept the colonial borders. Ghana and Togo raised also objections on their boundary that splits the Ewe.

is disconnected from the central government in Dakar. The independence "Movement of the Democratic Forces of Casamance" was supported by the neighboring Guinea-Bissau (and to a lesser extent by Gambia), where the Diola exert a significant influence.[10] Our results are in line with these arguments. In Senegal Murdock (1959) maps 12 ethnic homelands. In our sample 40% of a total of 198 battles and 40% of 140 violent events against civilians have taken place in the homeland of the partitioned Diola.

Similarly, compared to non-split ethnicities, ethnic militias and rebel groups that support partitioned ethnicities shelter and regroup on the other side of the border. The notorious Lord's Resistance Army (LRA) militia, which was founded to protect the marginalized by ethnic partitioning Acholi group offers an illustration. Whenever the Uganda government forces were defeating the LRA, the rebel group and its leader Joseph Kony were sheltering, regrouping, and rearming in the Acholi homeland in Southern Sudan and the Central African Republic. Nowadays, and in spite of the improving political and economic situation in Uganda, the Acholi-land is the most conflict-prone region in Africa recording more than 1,500 incidents of conflict.

Third, African borders are poorly delineated due to the imprecise colonial treaties. This has resulted in many border disputes, especially when poorly demarcated borders cause the partitioning of ethnic groups.[11] This imprecision seems to have contributed to conflict in Somalia (S.Samatar (1985)), whereas the ambiguity of the tripartite treaty of 1902 between Britain, Italy and Ethiopia has also played a role in the Eritrea-Ethiopia war.

Fourth, Africa is characterized by patronage politics where dominant ethnic groups discriminate against minority groups. In case of partitioned groups the neighboring country intervenes either to support its peers or to prevent migration and refugee flows. 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 conflict in the subsequent years. Our data are in line with this argument. While the civil war is long over, we still observe violence against civilians and (relatively minor) conflict in the homeland of the Ewe both in Ghana and in Togo, although overall conflict in both countries has been minimal in the past decade.[12]

---

[10]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.*"

[11]Englebert, Tarango, and Carter (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.*"

[12]The conflict in the Alur-land offers another illustration of this type of violence. The Alur had been split between the Belgian Congo and the British Protectorate of Uganda during the late phase of the scramble for Africa (1910 − 1914). After independence when the regime of Mobutu Sese Seko initiated the subjugation of many minority groups in Congo, a large portion of the Alur moved to their homeland in Uganda. This in turn

Fifth, due to the poor institutional infrastructure and 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 provides an illustration of how smuggling allowed the Hausa to arbitrage price caps and other distortionary policies in Niger and Nigeria. In line with these arguments over our sample period the Hausa-land in Nigeria has experienced 111 conflict incidents, while the average (median) number of conflict across the 112 ethnic regions is 20 (3).

Sixth, partitioning and border artificiality may lead to armed warfare by interacting with natural resources. 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 (see Caselli, Morelli, and Rohner (2012) for a theoretical exposition and cross-country-pair evidence).

Finally, the artificial border design may have contributed to underdevelopment and civil conflict via channels beyond ethnic partitioning. In particular, the colonial border drawing shaped a host of country-specific geographical and cultural characteristics including a country's ethnic diversity, size, access to the sea, etc. For example, Herbst (2000) argues that civil conflict is more pervasive in large African countries due to geographic inequalities that make it harder for the state to broadcast political power and prevent secessionist movements (see also Michalopoulos and Papaioannou (2012a)). Collier and Venables (2008) observe that the border design resulted in Africa having the largest proportion of landlocked countries limiting their growth potential. While our analysis focuses on one aspect of the scramble for Africa, namely the effect of ethnic partitioning on civil conflict, we are able to account for these other aspects of European's influence with the inclusion of country fixed effects that account for all time-invariant, country-specific characteristics.

Related Literature

Our paper contributes to two main strands of literature. First, it 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 the formation of early institutions (e.g. Acemoglu, Johnson, and Robinson (2005)) or via human capital (e.g. Glaeser, LaPorta,

---

generated opposition from the Buganda (the main group in Uganda) leading to civil conflict.

6

de Silanes, and Shleifer (2004) and Easterly and Levine (2009)). In contrast to this body of work, Gennaioli and Rainer (2006, 2007) and Michalopoulos and Papaioannou (2012b) focus on the pre-colonial period and show that deeply-rooted ethnic institutions correlate significantly with contemporary economic development. In the same vein Besley and Reynal-Querol (2012) show that contemporary conflict correlates with pre-colonial conflict.[13]

Our study contributes to this body of research, by emphasizing a neglected aspect of colonization; the drawing of political boundaries in the end of the 20th century that resulted in a large number of partitioned ethnicities. As such our work is mostly related to Alesina, Easterly, and Matuszeski (2011) who show that "artificial states" with straight borders and 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 the independence allows us to identify the causal effect of partitioning.

Second, our work contributes to the literature on the origins of civil conflict that mainly examines the role of country-level characteristics, such as income and natural resources (see Collier and Hoeffler (2007) and Blattman and Miguel (2010) for reviews and Collier and Sambanis (2005) for case studies in Africa). Of most relevance to the present study are works that link a country's ethnic composition to civil war, Caselli and Coleman (2012). While the correlation between ethnic fragmentation and civil war is weak (Fearon and Laitin (2003)), recent studies document interesting cross-country correlations associating various aspects of the societal structure with armed conflict. Montalvo and Reynal-Querol (2005) and Esteban, Mayoral, and Ray (2012) show a strong negative correlation between ethnic polarization and conflict. Wimmer, Cederman, and Min (2009) find that the likelihood of ethnic conflict increases when a large share of the population is excluded from power. Matuszeski and Schneider (2006) document that civil warfare is much higher in countries where ethnicities are clustered in specific areas. Englebert, Tarango, and Carter (2002) show a positive cross-country correlation between proxy measures of suffocation and dismemberment and political violence, secession attempts, border disputes, and civil warfare.

The correlations found in studies exploiting cross-country variation in border design and the distribution of ethnicities are informative; yet they cannot be causally interpreted (see Blattman and Miguel (2010) for a discussion). The main reason is that the process of border

---

[13] Of some relevance to our work are studies showing a significant negative association between ethnic fragmentation/polarization and development (see Alesina and Ferrara (2005) for a review). Ethnic fragmentation tends to lower public goods provision (Alesina, Baqir, and Easterly (1999), LaPorta, de Silanes, Shleifer, and Vishny (1999)), fuel authoritarianism (Aghion, Alesina, and Trebbi (2004)), and increase the likelihood of secession (e.g. Alesina and Spolaore (2003)), especially when ethnicities are segregated (Alesina and Zhuravskaya (2011)) or are economically unequal (Alesina, Michalopoulos, and Papaioannou (2012)).

drawing is historically related to the process of state formation and is thus associated with both voluntary and forced peoples' movements. Our study accounts for some of the shortcomings of cross-country studies. First, it establishes that African borders are to a great degree artificial by showing that there are no systematic differences in geographic, economic, institutional, and cultural characteristics between partitioned and non-split ethnicities. 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 migratory flows ignited by the border design. Third, we can control for country fixed effects as well as local geography, the disease environment, natural resources, and other factors that a vast literature has emphasized as key determinants of civil conflict and under-development.

Structure

In the next section we discuss how we identify partitioned ethnic groups and present the geo-referenced civil conflict data. We give the descriptive statistics illustrating the cross-ethnicity and within-country variation in conflict intensity. We also report test of means/medians that illustrate the significant differences in the likelihood and intensity of armed conflict between partitioned and non-split groups. In Section 3 we examine whether there are systematic differences between partitioned and non-partitioned ethnicities with respect to an array of geographic and historical features. Section 4 reports our baseline estimates on the effect of ethnic partitioning on various aspects of civil conflict (number of conflict incidents, violence against the civilian population, total battles as well as battles resulting in territorial changes between the government and rebel groups). In Section 5 we report the results of our sensitivity analysis. In Section 6 we summarize discussing possible avenues for future research.

2   Data

2.1   Identifying Partitioned Ethnic Groups

We identify partitioned groups projecting contemporary national borders, as portrayed in the 2000 Digital Chart of the World on George Peter Murdock's Ethnolinguistic Map (1959) that depicts the spatial distribution of African ethnicities at the time of European colonization in the mid/late 19th century (Figure 1a).[14] Murdock's map divides Africa into 843 ethnic regions. The mapped ethnicities correspond roughly to levels $7-8$ of the Ethnologue's language family tree. 8 areas are "uninhabited upon colonization" and are therefore not considered in our analysis. We also drop the Guanche, a small group in the Madeira islands that is currently

---

[14] When we intersect Murdock's ethnolinguistic map with the 2000 Digital Chart of the World we drop resulting partitions of less than 100 square kilometers, as such tiny partitions are most likely due to the lack of precision in the underlying mapping of ethnicities.

part of Portugal. Out of a total of 834 ethnicities in Murdock's Map, the homeland of 358 groups falls into more than one contemporary country. Yet for several of these groups the overwhelming majority of their homeland's area (usually more than 99%) falls into a single country. For example, 99.5% of the total area of the Ahaggaren falls into Niger and only 0.5% falls into Algeria. Since Murdock's map is bound to be drawn with some error, we identify as partitioned groups those ethnicities with at least 10% of their total surface area belonging to more than one countries ($SPLIT$). As such the Ahaggaren is classified as a non-split group. There are 231 ethnic groups with at least 10% of their historical homeland falling into more than one contemporary states (Figure 1b). Appendix Table A lists all partitioned ethnic groups. When we use a broader threshold of 5% we identify 267 partitioned ethnicities. In our empirical analysis we also exclude 8 regions where population according to the earliest post-independence census is zero. Thus, in our baseline sample we have a total of 826 populated ethnic areas of which 230 are partitioned.[15]

Our procedure identifies most major ethnic groups that have been split by African borders. For example, the Maasai are partitioned between Kenya and Tanzania (shares 62% and 38%, respectively), the Anyi between Ghana and the Ivory Coast (shares 58% and 42%, respectively), and the Chewa between Mozambique (50%), Malawi (34%), and Zimbabwe (16%). Other examples include the Hausa (split between Nigeria and Niger), the Ababda (split between Egypt and Sudan), and the Bararetta Somali clans (split between Kenya and Somalia). We also checked whether our codification of partitioned ethnicities is in line with Asiwaju (1985), who provides the only (to our knowledge) codification of partitioned ethnicities in Africa. Our

---

[15] When we primarily explore within-country variation, we also lose observations in countries with either one ethnic group or without variability in partitioning, namely in Burundi, Djibouti, Swaziland, Comoros, Madagascar, and Western Sahara.

strategy identifies almost all ethnic groups that Asiwaju (1985) lists as partitioned.[16]
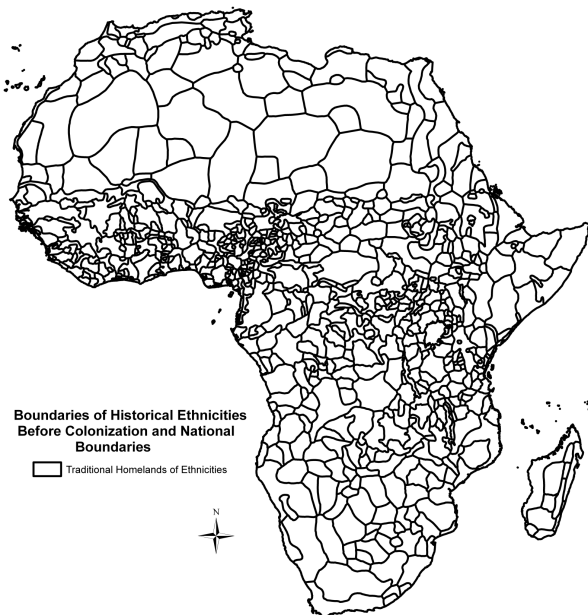
Figure 1a
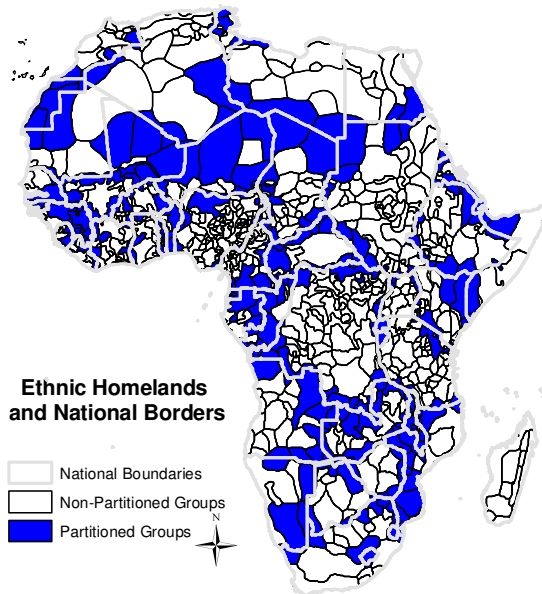
Figure 1b

We also construct a continuous index of partitioning in the spirit of the ethnic/linguistic fragmentation indicators (e.g. Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003)). The continuous index of partitioning reflects the probability that a randomly chosen pixel (area) of the historical homeland of an ethnic group falls into a different country. The ethnic groups with the highest score in this index are the Malinke, which are split into six different countries; the Ndembu, which are split between Angola, Zaire, and Zambia; and the Nukwe, which are split between Angola, Namibia, Zambia, and Botswana.[17]

2.2 Civil Conflict Data

The main data source for the location of armed conflict across African regions is the Armed Conflict Location and Event Dataset Raleigh, Linke, Hegre, and Karlsen (2010). The ACLED project provides geo-referenced data on the exact location (and some key characteristics) of politically violent events across all African countries since 1997. In contrast to other geo-referenced civil war datasets that only report the centroid and an approximate radius of the major war

---

[16]In Section 5 we report results using the Ethnologue's mapping of linguistic groups to identify partitioned ethnicities. The results are similar.

[17]We prefer the binary index of partitioning for several reasons. First, all studies in African historiography suggest that what matters for civil conflict is whether an ethnicity has been partitioned or not rather than the degree of the split. Second, there is no clear reason on why the propensity to conflict should monotonically increase with the degree of partitioning. Third, as Murdock's map certainly contains noise, this will be reflected more clearly in the continuous measure (as compared to the binary index). Nevertheless, to show that our results are not sensitive to the index of partitioning in Table 8 we report specifications with the continuous index.

incidents, ACLED provides precise locational data of armed violence. The ACLED database is quite rich as there are $43,271$ incidents of political violence over the period $1997 - 2010$ across all African countries. Political violence is understood as the use of force by a group with a political purpose or motivation. Conflict groups (actors) include rebels, militias, governments, and organized political or ethnic groups that interact violently over issues of political authority, such as territorial control, government control, access to resources, etc. ACLED categorizes armed conflict into 8 types. (1) Battles without change of control; (2) Battles where rebel groups gain control of the location; (3) Battles where the government regains control of a location; (4) Headquarter of base establishments, where rebel groups establish (via violent or non-violent means) their base; (5) Non-violent conflict events where rebel groups, militias or government forces proceed in non-violent actions (without active fighting) that are, however, within the context of an ongoing civil conflict and dispute (e.g. recruitment drives, incursions or rallies); (6) Riots and protests; (7) Violence again civilians, where armed groups (rebels, militias or government forces) attack unarmed civilians; (8) Non-Violent transfer of control.

Our benchmark index of civil conflict is the count of all types of armed conflict. We also examine the effect of ethnic partitioning on (i) the total number of battles, (ii) battles that resulted in a change in territorial control, and (iii) violent events against civilians. Battles among armed forces account for 43.2% of all incidents in the ACLED database; out of a total of $18,705$ battles, $2,324$ resulted in a change of territorial control. The dataset also records information on $15,844$ cases of violence against civilians (36.6% of all incidents of conflict); as such we can examine the effect of partitioning not only on fighting among organized armed actors, but also on violence against civilians. The latter includes incidents of rape, terrorism, mass killings of civilian populations, kidnaps, and fires.

Examples of battles between armed actors include the numerous fights between the Acholi-based Lord's Resistance Army (LRA), the Sudanese People's Liberation Army, and the Uganda's People Defence Force (UPDF); the constant fighting between the Rwandan military forces against the FDLR (Forces démocratiques de libération du Rwanda) Hutu rebels both in Rwanda and in Eastern Congo; the battles between the republican forces of Ivory Coast and various militias either supporting Alassane Quattera or Laurent Gambo; and the (relatively small scale) battles between Kikuyu rebel groups against Maasai militias. Battles result usually in many casualties; for example in a single event in September 1999 the Ugandan army killed 42 Pian warriors, coming from the Karamojong ethnic group that is split between Uganda, Sudan, and Kenya. Battles resulting in territorial change of control are usually even more devastating involving both a higher number of casualties and ambushes against the civilian population. For example, ACLED reports that in August 1997 when government forces of

11

the Democratic Republic of Congo retook control of the town of Watsa, close to the border with Uganda, where the partitioned Alur reside, this resulted in 800 casualties.[18]

Violent events against civilians include the raids of the Janjaweed militias against civilian population in the Darfur region in Eastern Sudan; the assaults and tortures against the civilian population by President's Mugabe's Central Intelligence Organization in Zimbabwe; the killings of civilians in Northern and Western Rwanda by the Interahamwe Hutu ethnic militias (that are raiding from their bases in Eastern Congo and Uganda); and the killings, abductions, rapes, and terrorist activities of numerous militia groups in the Democratic Republic of Congo. Violent events include also the burning of churches (by militias in Eastern Congo, Sudan, and Nigeria), tortures (that are widespread in the border of Zaire and Uganda), hostage-taking and child-soldiering raids. Going over the narratives of each event reveals that they may be also quite devastating. For example, in a single event in Eastern Congo in May 1997 "ADLF rebels moved in and took control of Mbandaka slaughtering 200 Rwandan Hutu refugees".

2.2.1 Data Patterns

Figure $2a$ maps the spatial distribution of all conflict events over the period $1997 - 2010$. There is significant heterogeneity in the incidence of political violence across Africa. There are numerous conflict events in Central Africa, mostly in Eastern Congo, Rwanda, Burundi, Sudan, and Uganda. In Western Africa, conflict and political violence are mostly present in Nigeria, especially in areas close to the Niger delta and in Sierra Leone. Moreover, political violence is pervasive in Somalia and Ethiopia, as well as in Zimbabwe and the Northern regions of South Africa. In contrast there are few events in Botswana, Zambia, Tanzania, Namibia, and Gabon. There is also considerable variation within countries. For example, while conflict incidence in Tanzania is low, there are quite a few violent incidents along the border with Kenya and Rwanda, though not as much as at the border with the Democratic Republic of Congo. Likewise, most of the conflict in Senegal is concentrated on the Southern region of Casamance, while most conflict in Angola is close to the border with Congo and in the Cabinda enclave.

To construct the conflict intensity per ethnic homeland we project ACLED's mapping of conflict events (Figure $2a$) on Murdock's ethnolinguistic map (Figure $1a$). Figure $2b$ portrays

---

[18]A proper analysis of the number of casualties is hindered by the fact that the ACLED dataset does not systematically report the number of casualties per incident.

the spatial distribution of all civil conflict incidents at the ethnic-homeland level.
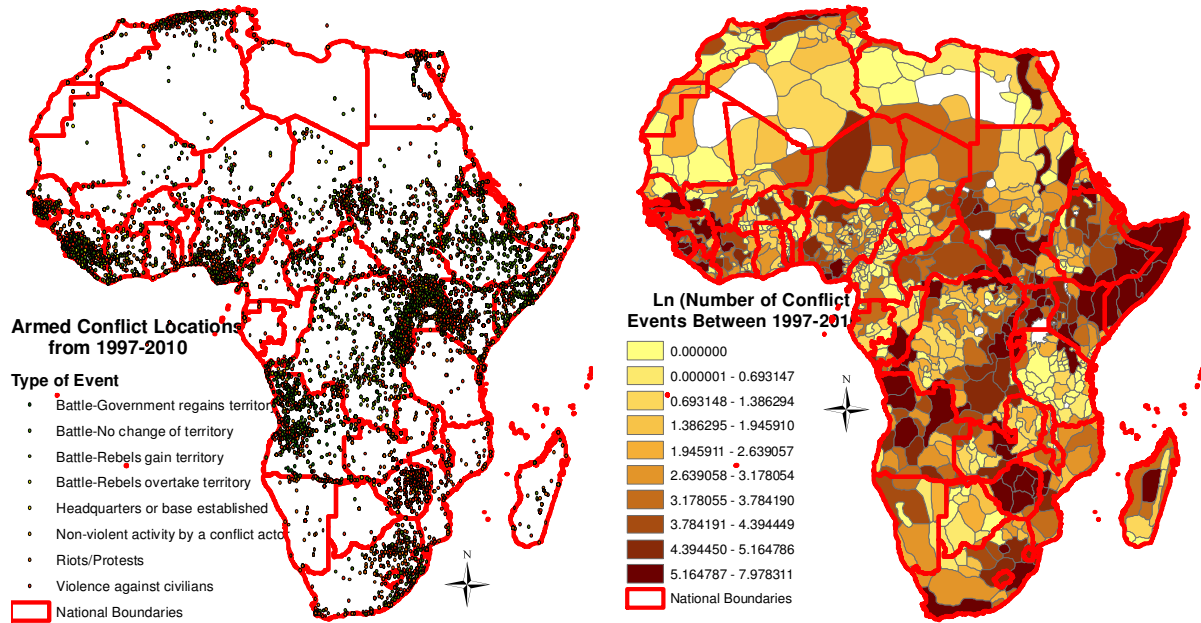
Figure 2a

Figure 2b

We also examine the effect of partitioning on the number of battles and violence against civilians, as these are the dominant and deadly types of conflict. Figures 3a and 3b plot the number of battles and an indicator that identifies ethnic homelands where a battle resulted in a territorial change, respectively. Figure 3c portrays the number of events involving violence against civilians. The correlation between battles' intensity and violence against civilians in high, but far from perfect (0.60). For example, in most areas in Zimbabwe we observe a large number of violent events against civilians with very few conflict incidents between military actors. Conversely in Ethiopia and Sudan we predominantly observe armed conflict between the government and rebel groups rather than unilateral violent events against civilians. In the Democratic Republic of Congo and in Uganda the number of battles go hand in hand with violence against civilians. The correlation between battles resulting in a territorial change and the total number of battles is 0.59 and with violent events against civilians is 0.22 (Appendix Table 1). This suggests that in our empirical analysis we can examine the effect of ethnic partitioning not only on overall political violence but also on its individual components.
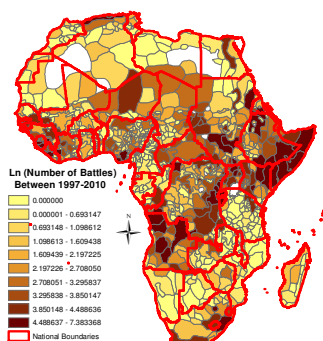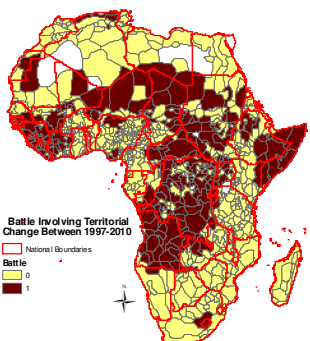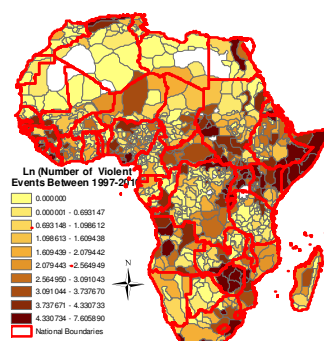
Figure 3a   Figure 3b   Figure 3c

2.2.2   Summary Statistics

Table 1 reports descriptive/summary statistics for the main outcome variables across the 826 ethnic homelands.[19] In Panel A we report summary statistics (mean and medians) across all ethnic homelands, while in Panel B we report statistics for homelands close to the national border (using the median value of distance to the border; 102 kilometers). This helps us isolate the role of ethnic partitioning from an overall border effect (which, nonetheless, may still be driven by partitioning).

     Civil Conflict Incidents: All Types   Three-fourths of all ethnic areas have experienced at least one conflict event over the period $1997-2010$ (column(1)). 84% of all partitioned ethnicities (193 out of 230) experienced some conflict, while the likelihood of a civil conflict incidence for non-partitioned ethnicities is 11 percentage points lower (73%). When we focus on groups close to the border (Table 1- Panel B), we observe similar -and if anything larger-differences; on average 65% of non-partitioned ethnic homelands experienced a conflict, while 83% of partitioned ethnicities suffered some type of conflict.

     Partitioned groups have also experienced more violent events than non-split ones. On average partitioned ethnicities experienced 64.75 incidents, while for the rest 47.6 incidents were recorded. This difference is not statistically significant because there are some extreme cases (outliers) both across partitioned and non-split ethnic groups (see Appendix Table 2). To account for outliers we exclude ethnic homelands where capitals fall (in columns (4) and (5)) and homelands where the number of armed conflicts exceeds the $99th$ percentile (in (6)-(7)). The differences in the mean number of conflict incidents between partitioned and non-split groups continue to be large (17.4 and 21.9) and are now statistically significant at the 99% level.

---

[19]In Supplementary Appendix Table 2 we report summary statistics for all outcome and control variables, both at the ethnic homeland level (Panel A) and at the country-ethnic-homeland level (Panel B).

The median of all conflict incidents across all ethnicities is 4. Again there are large differences between partitioned (median=13) and non-split ethnicities (median=3). The differences in conflict intensity between partitioned and non-split ethnic groups are also sizable when we focus on areas close to the border (Panel *B*). While the average (median) number of all civil conflict incidents for partitioned ethnicities is 66 (12), for non split ones the average (median) is 32.3 (1).[20]

Battles　　On average 59% of all ethnic homelands have experienced at least one battle between government forces, rebel groups, or militias. The corresponding likelihood for partitioned and non-split groups in the full sample is 63.5% and 57.7%, respectively. When we restrict estimation in areas close to the national border, the difference between partitioned and non-split groups in the probability of experiencing at least one battle is somewhat larger (15%). On average partitioned ethnic homelands have experienced ten more battles as compared to non-split groups (29.9 versus 19.8); and while due to outliers this difference is not statistically significant, once we exclude the top 1% of the distribution or ethnic regions where capitals fall, the difference is always significant at standard confidence levels. Likewise, the median value of battles for non-split ethnic groups close to the national border is zero, while the corresponding median value for partitioned ethnic homelands is 3.

Territorial Control Change　　The ACLED database also reports battles that resulted in territorial change of control. Focusing on such battles is interesting as in these cases the local population is likely to be more dramatically affected. 27% of all ethnic homelands experienced a territorial change of control; yet partitioned ethnic homelands were affected much more. The likelihood that a battle resulting in a change of territorial control for partitioned ethnic homelands is 36%, while the corresponding likelihood for non-split groups is 22.7%. This pattern suggests that partitioned ethnic groups are more likely to be traumatized as control oscillates between the government and rebel forces.[21]

Violence against Civilians　　The summary statistics of violence against civilians also reveal large and significant differences between partitioned and non-split groups. The likelihood that a partitioned ethnicity has experienced at least one violent event against the civilian

---

[20] The results are similar if we use a narrower threshold of distance to the national border to identify ethnic homelands close to the national border. For example when we use the 25% percentile of distance to the border (45 km), the average (median) number of civil conflict incidents for partitioned ethnicities is 47 (8.5) while for non-split ethnicities 4.6 (1).

[21] Similarly the likelihood that headquarters of rebel groups or militias are established in the homeland of partitioned ethnicities is 25.2% while the corresponding likelihood for non-split ethnic groups is 9.4%; when we focus in areas close to the national border the corresponding probabilities are 25.4% and 13.4%, respectively.

population is 0.70, while the corresponding likelihood for not split ethnicities is 0.57. The difference is even larger when we focus on ethnic homelands close to the national border (0.24). On average partitioned ethnic homelands experience 24 violent incidents against the civilian population, while the average for non-split ethnicities is 17.6. The median value of violence against civilians across partitioned ethnic homelands is three times the median value across non-split ethnic groups (3 versus 1) independently on whether we examine all ethnic homelands or we limit our attention to regions close to the national border.

3 Borders Artificiality

3.1 Empirical Specification

In this section we explore potential correlates of ethnic partitioning estimating models of the following form:[22]

$$SPLIT_i \ (FRAC_i) = a_r + X_i'\Psi + Z_i'\Theta + e_i. \tag{1}$$

The dependent variable, $SPLIT_i$, equals one when at least 10% of the historical homeland of an ethnic group i has been partitioned into more than one contemporary states. We also show results using the continuous measure of partitioning, $FRAC_i$. X_i is a vector of geographic, ecological, natural resource at the ethnicity level; Z_i is a vector of ethnic-specific pre-colonial institutional, cultural, and economic traits, extracted from Murdock's (1967) Ethnographic Atlas available for a subset of ethnicities. Appendix Table 2 gives summary statistics for all variables. In all specifications we include region-specific constants (a_r) to account for the somewhat different timing and patterns of colonization.

3.2 Results

Table 2 reports the results. Odd-numbered specifications report probit (maximum-likelihood) marginal effects, where the dependent variable is the benchmark binary partitioning index ($SPLIT$). Even-numbered columns report LS estimates, where the dependent variable is the continuous measure of partitioning ($FRAC_i$).[23]

Geographical, Ecological, and Natural Resource Measures In Panel A we explore the role of geographic, ecological, and natural resources. Specifications (1)-(2) show that

---

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

[23]The results are similar if we estimate Tobit models that account for truncation (at zero) of the continuous partitioning index.

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 borders in an arbitrary manner. The estimates further show that ethnicities residing in areas with larger water bodies (lakes and rivers) were more likely to find themselves split by the national boundaries. This result is in accord with the historical narrative that Europeans in some instances attempted to use natural barriers while delineating the spheres of their influence.

In columns (3)-(4) we augment the specification with an index reflecting the average land quality for agriculture and average elevation. We also add the respective standard deviations measures to proxy for ruggedness and the variance of land quality. All four geographic features enter with insignificant estimates.[24]

In columns (5)-(6) we examine whether partitioned and non-partitioned ethnic homelands differ with regards to ecological conditions, augmenting the empirical model with a malaria stability index (taken from Kiszewski *et al.* (2004)) and distance to the coast. Since Europeans settled mainly in areas by the coast and regions where malaria was less pervasive, these models also shed light on whether contact with colonizers affected partitioning. Both indicators enter with small and statistically indistinguishable from zero coefficients.

In columns (7)-(8) we 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 groups differ from non-partitioned ones in terms of natural resources. There are no systematic differences in this dimension.

Measures of European Contact and Early Development   In Panel *B* columns (1)-(4) we examine whether early contact with Europeans either during the slave trades or during the initial phase of colonization correlates with ethnic partitioning. In columns (1)-(2) we regress the partitioning indicators on the log number of slaves exported during the slave trades. In columns (3)-(4) we regress the partitioning measures on the average distance of each ethnic group to the main European exploration routes (using data from Nunn (2009)). Both variables enter with an insignificant estimate showing that early contact with colonizers is not related to ethnic partitioning.

---

[24]In some specifications mean land suitability for agriculture enters with a (weakly) significant estimate. We further explored the role of land quality and dependence on agriculture using alternative measures and numerous model permutations. Overall the correlation between ethnic partitioning and land quality is weak and in most specifications statistically indistinguishable from zero (see for example the results in Supplementary Appendix 3). Even in the models where some index of land's suitability for agriculture enters with a significant estimate, the economic magnitude is small. Moreover, in our analysis of the effect of ethnic partitioning on civil conflict we are reporting specifications accounting for land quality as well as numerous other local geographic controls, showing that this has no impact on the coefficient of the partitioning index.

In columns (5)-(6) we explore whether pre-colonial ethnic economic development was taken into account when the colonial borders were being designed. 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 and zero otherwise (using data from Chandler (1987)). There is no evidence that ethnicities with historical urban centers were differentially treated. Similarly, in columns (7)-(8), using Murdock's (1967) data, we show that there are no differences in the type of pre-colonial settlements. The settlement pattern variable ranges from 0 to 7 with higher numbers indicating more complex and thus more densely populated local communities.[25]

Ethnic-Specific Pre-colonial Traits In Panel C we examine whether other ethnic-specific pre-colonial institutional, cultural, and economic traits correlate with partitioning, using the rich information provided in Murdock's (1967) Ethnographic Atlas.

In columns (1)-(2) we investigate whether Europeans took into 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). In columns (3)-(4) we examine whether the societal structure correlates with partitioning using Murdock's class stratification index. The index ranges from zero, indicating societies without any class distinctions, to four for ethnicities with significant class and wealth distinctions. Class stratification may also proxy for institutional and economic development, since pre-colonially stratified societies were usually more developed.

African scholars argue that pre-colonial economic and institutional development was higher in areas with intensive use of agriculture (Fenske (2009) provides empirical evidence supportive of this conjecture); thus in columns (5)-(6) we augment the specification with a $0-10$ index measuring the importance of agriculture for subsistence at the ethnicity level. In columns (7)-(8) we examine whether ethnic partitioning is systematically related to pastoralism using a $0-10$ range index of ethnicity's dependence on animal husbandry. Finally, in Appendix Table 3 we further explore the association between partitioning and numerous other ethnic-specific variables from Murdock (1967) measuring the dependence of the economy on agriculture, fishing, hunting, the type of family organization, the presence of rules for

---

[25]We also regressed the ethnic partitioning measures on a dummy variable that identifies societies living in compact and/or relatively permanent or complex settlements, failing again to detect a significant correlation (Supplementary Appendix 3).

inheritance, the role of clans.

There is no evidence that partitioned ethnic groups differ along this host of pre-colonial traits.

Country-Fixed-Effects Estimates Since in most specifications below we associate civil conflict with ethnic partitioning exploring within-country variation, one would like to know whether post independence there are systematic differences between partitioned ethnicities and non-split groups within countries. Table 3 reports country-fixed-effects specifications associating geographical, ecological, and natural resource features with ethnic partitioning. In this case the unit of analysis becomes an ethnicity-country observation assigning each partition of a group to the corresponding country. The evidence in Table 3 suggests that there is no systematic association between these traits and ethnic partitioning. Interestingly, the positive correlation between surface and water area and partitioning found in Table 2 turns now insignificant. This is because after partitioning both the overall surface area and the area under water of split groups within a country are comparable to those of non-partitioned ethnicities.

Summary The results reported in Tables $2 - 3$ and Appendix Table 3 are broadly consistent with the historical narrative on the arbitrary design of African borders. Out of dozens of potentially relevant variables, only surface area and the presence of water streams correlate robustly with partitioning. Perhaps more importantly, the overall 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 0.07. Likewise, the R^2 of the OLS models is below 0.13. The probit specifications perform quite poorly in predicting which ethnicities have been partitioned. For example, the specification with all the geographical, ecological, and natural resource measures in Table 2-Panel A (not reported) predicts correctly ($G(X_i'\Psi + Z_i'\Theta + a_j) > 0.5$) only 29 out of the 231 partitions with the benchmark index ($SPLIT$). So, although we cannot rule out the possibility that some unobservable characteristic may correlate with partitioning, the evidence suggests that ethnic partitioning is not correlated with observable factors that may independently affect civil conflict.

4 Partitioning and Civil Conflict

4.1 Econometric Specification

We estimate the long-run effect of the scramble for Africa on contemporary civil conflict running variants of the following empirical specification:

$$y_{i,c} = a_c + \gamma SPLIT_i + X'_{i,c}\Phi + \varepsilon_{i,c}. \tag{2}$$

The dependent variable, $y_{i,c}$, reflects civil conflict in the historical homeland of ethnic group i in country c. In the country-fixed-effects specifications (with a_c), each partition of a group is assigned to the corresponding country c. For example, conflict in the part of the Lobi in Ivory Coast is assigned to Ivory Coast, while conflict in Lobbi's homeland in Burkina Faso is assigned to Burkina Faso.[26] The coefficient γ on $SPLIT$ captures the direct (local) effect of ethnic partitioning on civil conflict.

Vector $X'_{i,c}$ includes geographical controls, like surface and water area; ecological features, such as a malaria stability index and land's suitability for agriculture; natural resources reflecting the presence of diamond mines and petroleum fields; early development proxies such as having a major city in 1400. To further account for location characteristics, in almost all specifications we control for the distance to the coast, the distance to the national border, the distance to the capital city, and an indicator for regions where capital cities are located. To minimize concerns that the coefficient on the partitioning index captures an overall border effect (which however may itself be driven by partitioning), we also report specifications restricting estimation to ethnic homelands close to the national border.

Estimation and Inference Since the dependent variable (all conflict incidents, number of battles, number of violent events against civilians) is a count variable, we estimate negative binomial models with maximum likelihood (Wooldridge (2002)).[27] This negative binomial model accounts for the many zeros, as well as for the fact that there are a few extreme observations in the right tail of the distribution of the dependent variable. Moreover, the nonlinear estimator is appealing because it does not require log-linearizing the dependent variable and thus preserves the higher moments of the distribution (see Silva and Tenreyro (2006) and Silva, Tenreyro, and Windmeijer (2010)). To illustrate the robustness of our estimates, we also report log-linear LS specifications taking the log of one plus the respective civil conflict measure as the dependent variable.[28] To further account for outliers, we report specifications excluding homelands hosting capital cities or groups where the dependent variable exceeds the top 1%.

In all specifications we account for spatially correlated residuals clustering standard errors at the country level and at the ethnic-family level using the multi-way method of Cameron,

---

[26] In the previous draft, where the unit of analysis was the entire ethnic homeland, we assigned partitioned ethnicities to the country where the centroid of the historical homeland falls, finding similar results.

[27] Due to overdispersion in the number of battles and the number of violent events against civilians, specification tests reject the Poisson model, favoring the negative binomial model.

[28] Standardizing the dependent variable with land area or population yields similar results.

Gelbach, and Miller (2011). This correction also accounts for arbitrary residual correlation within each country and within each ethnic family. Moreover, double clustering accounts for spatial correlation.[29] We also estimated standard errors using Conley's method to account for spatial dependence of an unknown form, finding similar (and if anything less conservative) standard errors.

4.2 Cross-Sectional Estimates

We start our analysis estimating the relationship between partitioning and civil conflict across the 826 ethnic homelands (without country fixed effects). Table 4 reports the results. In column (1) we simply control for log population using the first post-independence census estimate (for most countries in the 1960s or 1970s), the log of surface area, and the log of area under water, the only variables found to correlate with partitioning in Table 2. In line with the descriptive analysis, the coefficient on the partitioning index is positive (0.76) and highly significant. Adding region constants (in (2)) has little effect on the estimate. In column (3) we control for location augmenting the specification with the distance from the centroid of each ethnic homeland to (i) the national border, (ii) the sea coast, and (iii) the capital city. We also include an indicator for ethnic homelands where capitals fall. Overall, distance to the sea enters with a positive and significant estimate suggesting that there is less conflict in areas close to the sea. Distance to the capital enters with a positive estimate suggesting that there is more conflict in regions further from the capitals, though the coefficient is not always significant. Distance to the border enters with a negative sign; yet the coefficient is not statistically significant. The capital city indicator enters with a positive and highly significant coefficient. This is not surprising as violent events against civilians, riots, and protests often take place in the capitals. In spite of the inclusion of these significant covariates, the partitioning indicator drops only slightly (0.674) and retains significance at the 99% level.

Column (4) includes a rich set of controls, reflecting geography (land suitability for agriculture, elevation, malaria) and natural resources (indicators for diamond mines or oil deposits). Accounting for these factors seems a priori important, because the cross-country literature documents significant correlations between 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 Rod (2006); Bellows and Miguel (2009)) show that conflict is higher in areas

---

[29]Cameron, Gelbach, and Miller (2011) explicitly cite spatial correlation as an application of the multi-way clustering method. See Spolaore and Wacziarg (2009) and Nunn and Wantchekon (2011) for analogous applications of the multi-way clustering method in accounting for spatial correlation. Murdock (1959) assigns the 834 ethnic groups into 96 ethnolinguistic clusters/families.

with diamond mines and petroleum fields. Moreover, Cervellati, Sunde, and Valmori (2011) document a strong positive correlation between the disease environment and civil conflict. The magnitude on the partitioning index remains unaffected. This is consistent with the findings in Table 2 showing that partitioning is uncorrelated with these characteristics.[30] In column (5) we drop outliers (top 1% of the dependent variable), while in column (6) we exclude ethnic regions where capitals fall. This has little effect on the ethnic partitioning index. The most conservative estimate implies that partitioned ethnicities experience an increase of approximately 162 log points in the number of civil conflict incidents. This translates into an 85% increase in civil conflict activity ($\exp(0.62) - 1 = 0.85$) in areas where partitioned ethnicities reside (as compared to the homelands of non-split ethnicities). The effect of ethnic partitioning on civil conflict is quantitatively as strong as the effect of the petroleum indicator, that enters in almost all specifications with a positive and significant coefficient.

In (7)-(12) we restrict estimation to ethnic areas close to the national border. This allows us to compare civil conflict intensity between partitioned ethnicities and other at-the-border ethnic groups that were not directly affected by the artificial border design. We now have a more balanced sample with 213 partitioned ethnicities and 200 non-split ethnic groups. Across all permutations the coefficient on the partitioning index is positive and highly significant, reassuring that our estimates in the full sample are not capturing an overall border effect.[31]

4.3 Within-Country Analysis

Baseline Country Fixed Effects Estimates The positive association between ethnic partitioning and civil conflict shown in Table 4 (and the descriptive analysis in Table 1) may be driven by hard-to-account-for country-wide factors. In Table 5 we thus estimate country-fixed-effects specifications associating civil conflict across ethnicity-country homelands with partitioning. Columns (1)-(6) report estimates in the full sample, while columns (7)-(12) present results across homelands that are close to the national border (using the median value of distance to the national border, which at the ethnicity-country homeland level is $61km$).

The coefficient on the ethnic partitioning index in (1) and (2) is positive and more than two standard errors larger than zero. The estimate in column (2) implies that on average civil conflict intensity is higher in homelands of partitioned groups by approximately 60%

---

[30]In all specifications the natural resource measures enter with a positive statistically significant estimate. There is also some weak evidence that civil conflict is higher in mountainous regions.

[31]We also estimated specifications trying to account for externalities using a similar to Miguel and Kremer (2004) specification; there is some weak, though insignificant, evidence of spatial spillovers from the historical homeland of partitioned ethnicities to adjacent ethnic regions. It should be noted that if there are spillovers (externalities) from the historical homelands of partitioned ethnic groups (the "treatment" group) to neighboring regions where non-split ethnicities reside (the "control" group), then the estimate on the partitioning index will be a lower bound of the true effect.

($exp(0.47) - 1 = 0.60$). In column (3) we control for distance to the national border, distance to the sea coast, distance to the capital, and the capital city dummy. The coefficient, if anything, increases in absolute value, and becomes more precisely estimated. Conditioning on the rich set of controls and accounting for outliers either by excluding observations where capitals fall or by dropping areas where the dependent variable exceeds the top 1% has no effect on the estimated magnitude. In columns (7)-(12) we restrict estimation across ethnic areas that are close to the national border. Across all specifications the coefficient on ethnic partitioning is positive and highly significant.[32]

Ethnic Partitioning and Type of Civil Conflict In Table 6 we examine the effect of ethnic partitioning on the different types of conflict.[33] The coefficients on ethnic partitioning in columns (1)-(2) imply that fighting between government forces, militias, and rebel groups is more pervasive in the historical homelands of partitioned groups; the estimated magnitudes suggest that on average partitioned groups experience approximately 80% ($exp(0.60) - 1 = 0.82$) more battles as compared to non-split ethnic groups. Limiting our focus to ethnic areas close to the national border has little effect on the estimate.

Ethnic partitioning is also systematically linked with violence against civilians by the government forces or rebel groups. This shows that partitioning has not only resulted into more warfare between armed forces, but has been particularly devastating for the civilian populations. The estimate in column (3) implies that there are 65% ($exp(0.50) - 1 = 0.65$) more violent incidents against civilians in the homelands of partitioned ethnicities.

In columns (5), (6), (11), and (12) we report linear probability models associating partitioning with the likelihood that a change in territorial control occurs in an ethnic homeland. The estimates show that partitioned ethnic homelands are more likely to swing between different control groups. The coefficient in (12), where we restrict estimation to ethnic areas close to the national border, implies that there is a 5% higher likelihood that a battle resulting in a change of territorial control occurs in the homeland of a partitioned ethnicity. This effect is not small, as in the country-ethnic homeland sample, the overall likelihood that a territorial change takes place is around 20%.

Linear Specifications In Table 7, columns (1)-(3) and (6)-(8) we report linear specifications using the natural log of one plus the total number of civil conflict incidents as the

---

[32] We also identified at-the-national-border ethnic homelands as those where the distance of the closest edge of each ethnic polygon from the national border is zero. Thus in these models we compare partitioned ethnic homelands with those that are by the border, but not partitioned. The coefficient on the ethnic partitioning index enters with a positive and significant estimate.

[33] Appendix Table 4 reports analogous cross-sectional specifications at the ethnic homeland level.

dependent variable. The estimate on the partitioning index is positive and highly significant both in the full sample and when we restrict estimation in areas close to the border. Columns (4) and (9) report linear probability models where the dependent variable is a dummy identifying areas that have experienced some civil conflict. While by solely looking at the "extensive" margin of civil conflict, we do not exploit the richness of the data, we further account for the nonlinear nature of the dependent variable. Moreover, these specifications shed light on whether the effect of ethnic partitioning on civil conflict (shown in Tables $4 - 6$) operates at the intensive or extensive margin of conflict. The estimate on ethnic partitioning implies that there is an 8% higher likelihood that a partitioned group will suffer at least one civil conflict. We also estimated linear probability models using as the cutoff threshold the median number of civil conflicts (median equals 2). Thus in columns (5) and (10) the dependent variable takes on the value of one for ethnic homelands that experienced more than two civil conflict incidents during the period $1997 - 2010$ and zero otherwise. The coefficient on the ethnic partitioning index retains its economic and statistical significance.

4.4 Example: Conflict in East-Central Africa

East-Central Africa, one of the most conflict-prone regions in the world, offers an illustration of our empirical results.

Let us start from Tanzania, a country with little overall conflict; in the 69 ethnic regions of Tanzania there have been 175 conflict incidents over the period $1997 - 2010$. The mean (median) conflict per ethnic homeland is 2.5 (0). Most conflict (19 incidents) occurs at the border with Rwanda where the partitioned Rundi tribes reside. Conflict also spreads to the nearby homeland of the (ethnically similar) Ha, where both militias based in nearby Rwanda and Burundi raid against the civilian population looting, raping, and terrorizing the local population. In contrast, at the border with Uganda where the non-split Haya group resides there are only 4 conflict incidents. Interestingly, there is zero conflict in the non-split homelands of the Bende and the Pipa, although both groups reside at the border with the Democratic Republic of Congo, the country with the highest conflict intensity in Africa. This is because lake Tanganyika is the natural border between Tanzania and Zaire. Focusing now on the northern border of Tanzania with Kenya, there is recurring conflict in the homeland of the partitioned Maasai (in total 10 incidents) and (to a lesser extent) in the partitioned Digo along the Indian Ocean. For example, ACLED documents a deadly fight resulting in the death of (at least) 30 farmers by Maasai militia on December 8th 2000. In contrast, there is zero conflict in the Eastern part of the Tanzania-Kenya border where the non-split Pare group resides. Moreover, and in line with our results, there is no conflict at the Kenyan side of the border populated by

the non-split Teita group.

Focusing now on the Democratic Republic of Congo, there are $4,333$ conflict events across the 102 ethnic regions (mean=26; median=4). In the homelands of the three partitioned Ruanda ethnicities (of the Interlacustrine Bantu - Rwanda family) we have 946 incidents (i.e. more than 20% of all conflict in Congo); and in the two adjacent (non-split), but ethnically similar Rwanda groups of the Hunde and Toro we have 409 and 27 events, respectively. Going over the event narratives reveals that conflict in Eastern Congo is (partly at least) driven by partitioning, as it involves ethnic militias (such as the FDLR) that constantly move across the border between Rwanda, Burundi, and Uganda. The FDLR and other Hutu-based militias, which fled Rwanda after the 1994 genocide and sought shelter in their homeland in Eastern Congo, had played a major role in the recent civil conflict in Congo. Not only there are hundreds of conflict incidents in the homeland of the partitioned Rwandan tribes, but going over the event narratives reveals that these events had been very devastating.[34]

In contrast there is no conflict at the (non-split) Holoholo ethnic homeland on the Western (Zairian) bank of the Tanganyika lake; yet the adjacent to the Holoholo ethnic group in the South, the Tabwa (a Bantu tribe) has experienced 91 conflict events, 10 of those being major battles resulting in a change of territorial control. This is not surprising since, although the Eastern border of Congo with Tanzania is organic (Tanganyika lake), the Southern one with Zambia follows a latitudinal line that splits the homeland of the Tabwa almost equally between Congo and Zambia. Moreover, there is zero conflict in the nearby (further to the south) border group of the Shila in Congo that has not been affected by the artificial line splitting Zambia and Congo. In line also with our results there is no conflict on the other side of the border opposite to the Shila in Zambia where the non-split Luapula group resides.

---

[34]For example, in early 2007 in just one event, FDLR groups raided two villages killing 17 civilians and wounding 19. In the following day FDLR militias looted 18 houses in a nearby town. Also in an single event in Fendula in Eastern Congo, (at least) 30 civilians were burned alive and 50 wounded by Rwanda militias. Prunier (2009) provides a detailed narrative of how the partitioning of the Rwandan tribes and the genocide in Rwanda spread to Congo leading to the recent devastating civil war.
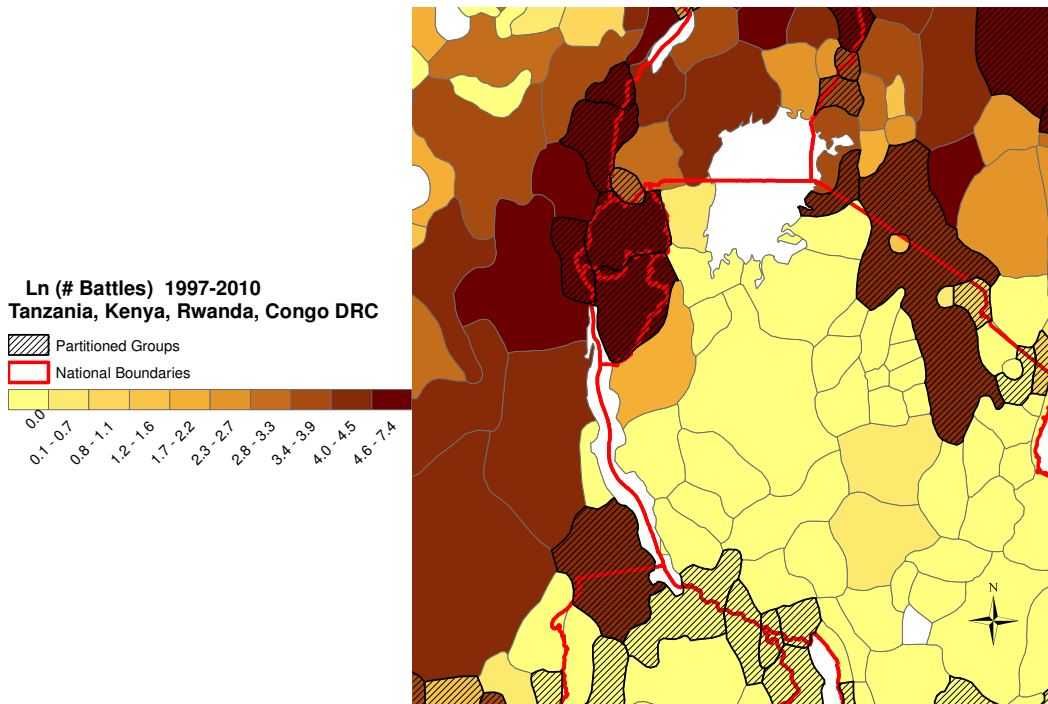
Figure 4

5 Sensitivity Analysis

We perturbed the empirical model in various ways to explore the robustness of our results. In this section we report the main sensitivity checks.

Alternative Measures of Partitioning   First, we repeated estimation using alternative measures of ethnic partitioning. Table 8 report the results. Columns (1)-(3) and (7)-(9) present results associating civil conflict with the continuous index of partitioning ($FRAC$) that reflects the likelihood that a historical ethnic homeland falls into more than one contemporary state. The ethnic partitioning measure enters with a positive and highly significant coefficient, implying that a higher degree of ethnic partitioning is associated with a higher likelihood of civil conflict. In columns (4)-(6) and (10)-(12) we regress civil conflict on a binary index of ethnic partitioning using a 5% threshold to identify split groups. This has little effect on our baseline results. The estimate implies that there is a 65% higher incidence of civil conflict events in the homelands of partitioned groups.

Unobservables   Second, to further account for unobservable characteristics that vary smoothly in space (see Dell (2010)), in Table 9A, columns (1)-(4) we augment the specification

26

with a cubic polynomial in latitude and longitude of the centroid of an ethnic group in each country.[35] The coefficient on the partitioning index remains virtually unaffected. Moreover, the estimate retains significance at the 99% confidence level. In columns (5)-(8) we include ethnic-family fixed effects to account both for local conditions and broad cultural, institutional, and other hard-to-observe ethnic-specific factors (see Nunn (2012)). Examples of ethnic families include the Bedouin Arabs, the Tuareg, and the Southwestern Bantu. The estimates suggest that, even when we solely examine within-country, within-ethnic-family variation, civil conflict is significantly more pervasive in border areas belonging to partitioned ethnicities. In columns (9)-(12) we report specifications with both ethnic-family fixed effects and the third-order polynomial in latitude and longitude; while we may be over-fitting, the coefficient on the ethnic partitioning index retains its economic and statistical significance.

Location   Third, we estimated models dropping iteratively ethnic homelands from each of the 5 African regions to investigate whether the results are driven by a particular part of the continent. Table $9B$ reports the results. Odd-numbered columns show results in the full sample and even-numbered columns report results for ethnic areas close to the national border. In (1)-(2) we exclude North Africa to account for the fact that Europeans had contacts with the northern part of the continent since the ancient times. In (3)-(4) we drop Southern African countries. In columns (5)-(6) we drop Western African countries because some of the contemporary African borders in this region correspond to internal administrative borders of the Federation of the French West Africa. In (7)-(8) and (9)-(10) we exclude ethnic areas in East Africa and Central Africa, respectively. This allows us to examine the robustness of our results to influential observations, as the most deadly and prolonged conflicts have taken place in Ethiopia, Somalia, Sudan, Rwanda, and the Democratic Republic of Congo. The results show that the strong positive effect of ethnic partitioning on civil conflict is not driven by a particular region. In all specifications the partitioning index enters with a highly significant estimate, similar in magnitude to the (more efficient) estimates in Table 5.

Partitioning and Civil Wars   In the previous draft of the paper we used geo-referenced data from the Peace Research Institute of Oslo (PRIO; Raleigh, Cunningham, Wilhemsen, and Gleditsch (2006)) on African civil wars over the $1970 - 2005$ finding (as in the current draft) that ethnic partitioning is associated with a higher number and more prolonged civil conflicts. However, a limitation of the PRIO mapping of conflicts is that it reports rough approximations of the location of civil wars using a centroid for each war and a coarse radius. However as the data code-book states each conflict, instead of having a circular zone (as the mapping sug-

---

[35]Letting x denote latitude and y denote longitude the polynomial reads: $x+y+x^2+y^2+xy+x^3+y^3+x^2y+xy^2$.

gests), the actual shape is more likely to follow the contours of mountains and rivers. Moreover, the PRIO identifies (major and minor) civil wars using information only from battle-related casualties.[36]

Despite these obvious shortcomings in Table 10 we reproduce the basic findings. Overall there are 49 civil and internationalized internal wars over the period $1970 - 2005$; these wars played out in 77 conflict zones. Out of the 826 ethnicities 343 experienced a civil war in their historical homeland; 199 ethnicities experienced two distinct incidents of armed conflict; 54 ethnicities experienced 3 civil wars while 12 regions were affected by four or even five conflicts.[37] Table 10 reports Poisson maximum likelihood estimates using the number of civil wars (in (1)-(2)) and the number of war zones (in (3)-(4)) at the ethnic homeland level as the dependent variable. To illustrate the robustness of our results we report specifications controlling for location, geographical features, as well as unobserved ethnic-family and spatial characteristics. Moreover, in the country fixed effects specifications we assign an ethnic homeland to a single country using the centroid of each homeland. The estimates show that partitioned ethnicities are significantly more likely to experience major civil wars.

Alternative Mapping of Ethnic Homelands and Ethnic Partitioning Fifth, we repeated the analysis using Ethnologue's database, that reports the spatial distribution of linguistic groups in the early/mid 1990s. One advantage is that the Ethnologue explicitly maps linguistic homelands within each country making the identification of partitioned linguistic groups straightforward. Moreover, using a contemporary dataset is useful because it contains less error than Murdock's pre-colonial map. The disadvantage is that the current location of ethnic groups is likely to have been affected by the border drawing, civil conflict, as well as numerous country-level characteristics that affect conflict. Ethnologue maps 2405 linguistic groups in Africa out of which 821 are explicitly mapped in more than one countries. Combining the Ethnologue with the ACLED mapping of conflicts we obtain the conflict statistics for each linguistic homeland. Partitioned groups are 48% likely to experience at least one conflict compared to 33% for non-partitioned groups. Moreover, conditional on having at least one conflict split groups experience an average of 61 conflict incidents whereas non-split ones register 34 conflict events. Overall, partitioned groups have suffered an average of 30 conflict incidents in-between $1997 - 2010$ whereas those language groups located in a single country

---

[36]PRIO identifies wars 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.*"

[37]The groups with the highest incidence of civil war are the Afar and the Esa, which during the period $1970 - 2005$ have experienced 5 civil wars. Both groups have been greatly impacted by the artificial border design with the Afar being partitioned between Ethiopia, Eritrea and Djibouti, and the Esa being split between Ethiopia and Somalia.

have experienced on average 10 conflict events. All differences are statistically significant at conventional levels. An illustrative example is that of the Acholi group partitioned between Sudan and Uganda. Both Acholi partitions have suffered significantly from conflict. The Sudanese partition has experienced 14 conflict incidents whereas the part of the Acholi in Uganda has experiences a staggering 1583 events of political violence.

Table 11 reports cross-sectional and within-country specifications using the Ethnologue data. Panel A reports negative binomial maximum likelihood estimates; Panel B reports analogous LS specifications using the log of one plus the number of all civil conflict events as the dependent variable. Columns (1)-(6) include all linguistics groups whereas in columns (7)-(12) we focus on groups whose centroid's distance to the national boundary is less than the median distance (86 kilometers). The coefficient on the ethnic partitioning index is positive and highly significant across all permutations. The most conservative estimate in Panel A implies that conflict intensity is approximately 30% ($exp(0.267) - 1 = 0.30$) higher in the contemporary homelands of partitioned groups.

6 Conclusion

This study examines the consequences of a neglected aspect of colonization, the artificial drawing of political boundaries among European powers in the end of the 19th century, which in the eve of African independence led to the partitioning of several ethnicities across the newly created African states.

In the first part of our paper we formally show the artificial nature of African political boundaries. Utilizing information on the spatial distribution of ethnicities at the time of colonization, we associate ethnic partitioning with various geographic and ethnic-specific pre-colonial characteristics. With the sole exceptions of the size of the historical homeland and water bodies, there are no other significant differences between partitioned and non-partitioned ethnicities. These results offer support to the African historiography on the accidental and artificial drawing of colonial and consequently national borders.

Second, we examine in a quasi-natural experimental setting the effect of ethnic partitioning on civil conflict, as this has been hypothesized to be the major consequence of the scramble for Africa. Our analysis is based on regional data spanning the universe of ethnic areas across Africa. We exploit a new rich geocoded dataset that reports information on more than 43,000 conflict events over the period $1997 - 2010$. The database is quite useful in examining the effect of ethnic partitioning on various aspects of conflict, as it reports both the precise location of battles between government forces, militias, and rebel groups, as well as the incidents involving violence against civilians (such as rapes, terrorism, abductions, fires, village

burning) and episodes of territorial change of control. Our regional focus enables us to solely examine within-country variation and as such account for all country-level features that may affect warfare.

We find that partitioned ethnicities have suffered disproportionately more from civil conflict compared to non-split ones. Battles between armed groups, as well as violence against the civilian population are concentrated in the homelands of partitioned ethnicities. Moreover, these territories are intensely contested among different conflict actors having a significantly higher likelihood of a change in territorial control. These results are robust to different estimation techniques, alternative classifications of partitioned ethnicities, accounting for ethnic family features and geographical characteristics, and more.

The uncovered differences in the probability and intensity of civil conflict, battles and violence against civilians between partitioned and non-partitioned groups becomes more dramatic when viewed in light of the fact that these two groups were socially, culturally and economically similar in the eve of colonization. Our work thus suggests that the scramble for Africa, by partitioning ethnicities in different countries, laid the seeds of a violent legacy of civil conflict and political violence. Our work suggests that future research should examine the effects of ethnic partitioning on economic and institutional development. Moreover, our study calls for future work to uncover the mechanisms via which the scramble for Africa has affected long-run economic performance. Finally, since border artificiality has been argued to be present in other parts of the world, such as the Middle East and Eastern Europe, subsequent works could also study the effects of partitioning in those regions.

7 Data Appendix

Partitioning Index (SPLIT): Indicator variable that equals 1 if at least 10% of the historical homeland of an ethnic group is partitioned into different countries. We also construct and alternative partitioning index that equals 1 if at least 5% 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 2000. Appendix Table 1 reports partitioned ethnicities.*

Continuous Measure of Partitioning (FRAC): The index reflects the probability that a square kilometer of an ethnic area falls to a different country than the rest of the historical ethnic homeland. Computed similarly to the Herfindahl index. *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 2000.*

All Civil Conflict Incidents: Sum of all civil conflict incidents. There are 8 event types. (1) Battles without change of control; (2) Battles where rebel groups gain control of the location; (3) Battles where the government regains control of a location; (4) Headquarter of base establishments, where rebel groups establish (via violent or non-violent means) their base; (5) Non-violent conflict events where rebel groups, militias or government forces proceed in non-violent actions (without active fighting) that are however within the context of an ongoing civil conflict and dispute (e.g. recruitment drives, incursions or rallies); (6) Riots and protests; (7) Violence again civilians, where armed groups (rebels, militias or government forces) attack unarmed civilians; (8) Non-Violent transfer of control. Depending on the unit of analysis we aggregate the data either at the ethnic homeland or at the country-ethnic homeland level. See Section 2 for details. *Source: ACLED.*

Battles: Total number of battles between two violent armed groups at the ethnic homeland (in each country for partitioned ethnicities). Battles include armed conflict where a control of the contested location does not change and conflict events resulting in a territorial change of control. We aggregate the data at the ethnic homeland level and at the country-ethnic homeland level. See Section 2 for details. *Source: ACLED.*

Violence against Civilians: Total number of violent events against civilians at the ethnic homeland (in each country for partitioned ethnicities). Violence against civilians occurs when any armed/violent group attacks unarmed civilians. Rebels, governments, militias, rioters can all commit violence against civilians.We aggregate the data at the ethnic homeland level and at the country-ethnic homeland level. See Section 2 for details. *Source: ACLED.*

Territorial Change of Control: Indicator that takes on the value of one if a battle

resulting in change of territorial control takes place at the historical homeland on an ethnic group / country-ethnicity. *Source: ACLED.*

Population at Independence: Log of population as recorded in the first post independence census. *Source: UNESCO (1987). Available at: http://na.unep.net/datasets/datalist.php.*

Land Area: Log surface area of the historical homeland of each ethnic group in $1000s$ 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 value (and standard deviation) of elevation in kilometers. *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 value (and standard deviation) of land quality for cultivation. The index is the product of two components reflecting the climatic and soil suitability for cultivation. *Source: Michalopoulos (2011); 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. We use the average value for each ethnic homeland (and for each country-ethnic region). *Source: Kiszewski, Mellinger, Spielman, Malaney, Sachs, and Sachs (2004)*

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

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

Distance to the Sea: The geodesic distance of the centroid of the historical homeland of each ethnic group from the nearest coastline, measured in $1000s$ 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 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 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/

Capital Indicator: Dummy variable that takes on the value one when a capital city is located in an ethnic historical homeland (in a country for partitioned ethnicities) and zero otherwise.

Latitude: Latitude of the centroid of each ethnic group. *Source: Constructed using ArcGIS Software.*

Longitude: Longitude of the centroid of each ethnic group. *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 per 1000 of square kilometers. *Source: Nunn (2008) and Nunn and Wantchekon (2011).*

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).*

Distance to Explorer's Routes: The geodesic distance of the centroid of each group to the nearest route of the principal European explorers. *Source: The "Century Atlas, Africa" digitized by Nunn (2012).*

Settlement Pattern: Ordered variable ranging from 0 to 7 quantifying "settlement pattern of each group". 0 indicates fully nomadic (migratory) groups, 1 indicates semi-nomadic, 2 indicates semi-sedentary, 3 identifies groups that live in compact and impermanent settlements, 4 indicates societies those in neighborhoods of dispersed family homes, 5 indicates for groups in separated hamlets forming a single community, 6 indicates societies living in compact and relatively permanent settlements, and 7 denotes the groups residing in complex settlements. *Source: Murdock (1967); variable code in the Ethnographic Atlas v30.*

Political Centralization: The binary index is constructed using Murdock's (1967) 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).*

Class Stratification: Ordered variable ranging from 0 to 4 quantifying "*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 "*the presence of wealth distinctions, based on possession or distribution of property, which however have not crystallized into distinct and hereditary social classes.*" A score of 2 indicates "*elite stratification in which an elite class derives its superior status from control over scarce resources, particularly land, and is thereby differentiated from a propertyless proletariat or serf class*". A score of 3 indicates a "*dual stratification into a hereditary aristocracy and a lower class of ordinary commoners or freemen, where traditionally ascribed noble status is at least as decisive as control over scarce resources.* A score of 4 indicates "*complex stratification into social classes correlated in large measure with extensive differentiation of occupational statuses.*" Source: Murdock (1967); variable code in the Entholinguistic Atlas v67.*

Dependence on Agriculture: $0-10$ scale index reflecting the dependence of each ethnicity in agriculture at the time of colonization. *Source: Murdock (1967).*

Animal Husbandry: $0-10$ scale index reflecting the percentage of subsistence coming from animal husbandry for each ethnicity at the time of colonization. *Source: Murdock (1967).*

Number of Civil Wars: $0-5$ index that counts the number of civil wars that have affected each ethnic homeland over the period $1970-2005$. *Source: Uppsala Conflict Data Program (UCDP)/International Peace Research Institute, Oslo (PRIO) Armed Conflict Dataset, Version 4-2006; and Raleigh, Cunningham, Wilhelmsen, and Gleditsch (2006).*

Number of Civil War Zones: $0-8$ index that counts the number of civil war zones that have affected each ethnic homeland over the period $1970-2005$. *Source: Uppsala Conflict Data Program (UCDP)/International Peace Research Institute, Oslo (PRIO) Armed Conflict Dataset, Version 4-2006; and Raleigh, Cunningham, Wilhelmsen, and Gleditsch (2006).*

References

ACEMOGLU, D., S. JOHNSON, AND J. A. ROBINSON (2001): "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91(5), 1369–1401.

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

——— (2005): "Institutions as a Fundamental Cause of Long-Run Growth," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. N. Durlauf, pp. 109–139. Elsevier North-Holland, Amsterdam, The Netherlands.

AGHION, P., A. ALESINA, AND F. TREBBI (2004): "Endogenous Political Institutions," *Quarterly Journal of Economics*, 119(2), 565–612.

ALESINA, A., R. BAQIR, AND W. EASTERLY (1999): "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 114, 1243–1284.

ALESINA, A., A. DEVLEESCHAUWER, W. EASTERLY, S. KURLAT, AND R. WACZIARG (2003): "Fractionalization," *Journal of Economic Growth*, 8, 155–194.

ALESINA, A., W. EASTERLY, AND J. MATUSZESKI (2011): "Artificial States," *Journal of the European Economic Association*, 9(2), 246–277.

ALESINA, A., AND E. L. FERRARA (2005): "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 43, 762–800.

ALESINA, A., S. MICHALOPOULOS, AND E. PAPAIOANNOU (2012): "Ethnic Inequality," mimeo Harvard University.

ALESINA, A., AND E. SPOLAORE (2003): *The Size of Nations*. MIT Press.

ALESINA, A., AND E. ZHURAVSKAYA (2011): "Segregation and the Quality of Government in a Cross-Section of Countries," *American Economic Review, forthcoming*.

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

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

BELLOWS, J., AND E. MIGUEL (2009): "War and Local Collective Action in Sierra Leone," *Journal of Public Economics*, 93, 1144–1157.

BESLEY, T., AND M. REYNAL-QUEROL (2012): "The Legacy of Historical Conflicts. Evidence from Africa," mimeo, London School of Economics and UPF.

BLATTMAN, C., AND E. MIGUEL (2010): "Civil War," *Journal of Economic Literature*, 48(1), 3–57.

BUHAUG, H., AND J. K. ROD (2006): "Local Determinants of African Civil Wars," *Political Geography*, 25(1), 315–335.

CAMERON, C. A., J. GELBACH, AND D. MILLER (2011): "Robust Inference with Multi-Way Clustering," *Journal of Business and Economic Statistics*, 28(2), 238–249.

CASELLI, F., AND W. J. COLEMAN (2012): "On the Theory of Ethnic Conflict," mimeo London School of Economics.

CASELLI, F., M. MORELLI, AND D. ROHNER (2012): "The Geography of Inter-State Wars," mimeo London School of Economics.

CERVELLATI, M., U. SUNDE, AND S. VALMORI (2011): "Disease Environment and Civil Conflict," IZA Discussion Paper No. 5614.

CHANDLER, T. (1987): *Four Thousand Years of Urban Growth: An Historical Census*. Edwin Mellon Press, Lewiston, NY.

COLLIER, P., AND A. HOEFFLER (2007): "Civil War," in *Handbook of Defense Economics, Volume 2, Defense in a Globalized World,*, ed. by T. Sandler, and K. Hartley, pp. 711–740. Elsevier, North Holland, Amsterdam and Oxford.

COLLIER, P., AND N. SAMBANIS (2005): *Understanding Civil War. Volume 1: Africa*. World Bank, Washington, DC.

COLLIER, P., AND A. VENABLES (2008): "Trade and Economic Performance: Does Africa's Fragmentation Matter?," Paper prepared for the ABCDE Conference.

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

CONLEY, T. G. (1999): "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 92(1), 1–45.

DELL, M. (2010): "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 78(6), 839–857.

DOWDEN, R. (2008): *Africa: Altered States, Ordinary Miracles.* Portobello Books Ltd, London, UK.

EASTERLY, W., AND R. LEVINE (2009): "The European Origins of Economic Development," Working Paper, mimeo Brown University.

ENGLEBERT, P. (2009): *Africa, Unity, Sovereignty and Sorrow.* Lynne Rienner Publishers, Inc., Boulder, Colorado.

ENGLEBERT, P., S. TARANGO, AND M. CARTER (2002): "Dismemberment and Suffocation: A Contribution to the Debate on African Boundaries," *Comparative Political Studies*, 35(10), 1093–1118.

ESTEBAN, J., L. MAYORAL, AND D. RAY (2012): "Ethnicity and Conflict: An Empirical Investigation," mimeo NYU Department of Economics, forthcoming American Economic Review.

ETHNOLOGUE (2005): Languages of the World, SIL International; Fifteenth edition.

FEARON, J., AND D. LAITIN (2003): "Ethnicity, Insurgency and Civil War," *American Political Science Review*, 97, 75–90.

FENSKE, J. (2009): "Does Land Abundance Explain African Institutions," mimeo Yale University.

GENNAIOLI, N., AND I. RAINER (2006): "Precolonial Centralization and Institutional Quality in Africa," in *Institutions and Norms in Economic Development*, ed. by M. Gradstein, and K. Konrad. MIT Press.

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

GLAESER, E. L., R. LAPORTA, F. L. DE SILANES, AND A. SHLEIFER (2004): "Do Institutions Cause Growth?," *Journal of Economic Growth*, 9(3), 271–303.

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

HERBST, J. (2000): *States and Power in Africa.* Princeton University Press, Princeton, NJ.

KISZEWSKI, A., A. MELLINGER, A. SPIELMAN, P. MALANEY, S. E. SACHS, AND J. SACHS (2004): "A Global Index of the Stability of Malaria Transmission," *American Journal of Tropical Medicine and Hygiene*, 70(5), 486–498.

LAPORTA, R., F. L. DE SILANES, A. SHLEIFER, AND R. VISHNY (1999): "The Quality of Government," *Journal of Law Economics and Organization*, 15(1), 222–279.

MATUSZESKI, J., AND F. SCHNEIDER (2006): "Patterns of Ethnic Group Segregation and Civil Conflict," Working Paper, mimeo Harvard University.

MEREDITH, M. (2005): *The Fate of Africa*. Public Affairs, New York, NY.

MICHALOPOULOS, S. (2011): "The Origins of Ethnolinguistic Diversity," American Economic Review, forthcoming.

MICHALOPOULOS, S., AND E. PAPAIOANNOU (2012a): "National Institutions and African Development: Evidence from Partitioned Ethnicities," mimeo, Brown University and Dartmouth College.

——— (2012b): "Precolonial Ethnic Institutions and Contemporary African Development," mimeo, Brown University and Dartmouth College.

MIGUEL, E., AND M. KREMER (2004): "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 72(1), 159–217.

MONTALVO, J. G., AND M. REYNAL-QUEROL (2005): "Ethnic Polarization, Potential Conflict and Civil War," *American Economic Review*, 95(3), 796–816.

MURDOCK, G. P. (1959): *Africa: Its Peoples and Their Culture History*. McGraw-Hill Book Company, New York.

——— (1967): *Ethnographic Atlas*. University of Pittsburgh Press, Pittsburgh, PA.

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

——— (2012): "Culture and the Historical Process," *Economic History of Developing Regions*, 27(S1), 108–126.

NUNN, N., AND L. WANTCHEKON (2011): "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, 101(7), 3221–3252.

PRUNIER, G. (2009): *Africa's World War. Congo, the Rwandan Genocide, and the Making of a Continental Catastrophe.* Oxford University Press, Oxford, UK.

RALEIGH, C., D. CUNNINGHAM, L. WILHEMSEN, AND N. P. GLEDITSCH (2006): "Conflict Sites Codebook," Centre for the Study of Civil War (PRIO).

RALEIGH, C., A. LINKE, AND C. DOWD (2012): "Armed Conflict Location and Event Dataset (ACLED) Codebook Version 2," Working Paper, Centre for the Study of Civil War, International Peace Research Institute, Oslo (PRIO).

RALEIGH, C., A. LINKE, H. HEGRE, AND J. KARLSEN (2010): "Introducing ACLED-Armed Conflict Location and Event Data," *Journal of Peace Research*, 47(5), 1–10.

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

ROSS, M. L. (2006): "A Closer Look on Oil, Diamonds, and Civil War," *Anuual Review of Political Science*, 9(1), 265–300.

SILVA, J. S., AND S. TENREYRO (2006): "The Log of Gravity," *Review of Economics and Statistics*, 88(4), 641–658.

SILVA, J. S., S. TENREYRO, AND F. WINDMEIJER (2010): "Is it Different for Zeros? Discriminating Between Models for Non-Negative Data with Many Zeros," CeMMAP working papers CWP20/10.

SPOLAORE, E., AND R. WACZIARG (2009): "The Diffusion of Development," *Quarterly Journal of Economics*, 124(2), 469–529.

S.SAMATAR, S. (1985): "The Somali Dilemma: Nation in Search for State," in *Partitioned Africans*, ed. by A. Asiwaju, pp. 155–194. St. Martin's Press,, New York, NM.

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

WIMMER, A., L.-E. CEDERMAN, AND B. MIN (2009): "Ethnic politics and Armed Conflict. A Configurational Analysis of a New Global Dataset," *American Sociological Review*, 74, 316–337.

WLMS (2006): World Language Mapping System, Version 3.2 Available on-line at http://www.gmi.org/wlms/.

WOOLDRIDGE, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Table 1: Test of Means and Medians for Main Civil Conflict Measures

Panel A: All Ethnic Homelands

	Indicator Likelihood	Number of Incidents		Excluding Capitals		Excluding Outliers (top 1%)	
	mean	mean	median	mean	median	mean	median
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
All Types of Conflict Events							
all ethnic homelands (N=826)	0.762	52.386	4.000	35.202	4.000	36.584	4.000
non-partitioned ethnic groups (N=596)	0.732	47.616	3.000	30.495	3.000	30.460	3.000
partitioned ethnic groups (N=230)	0.839	64.748	13.000	47.872	10.000	52.297	13.000
difference	**0.11**	**17.13**	**10.00**	**17.38**	**7.00**	**21.84**	**10.00**
difference (p-value)	*(0.00)*	*(0.28)*	*(0.00)*	*(0.02)*	*(0.00)*	*(0.00)*	*(0.00)*
Battles							
all ethnic homelands (N=826)	0.593	22.645	1.000	16.781	1.000	14.920	1.000
non-partitioned ethnic groups (N=596)	0.577	19.837	1.000	14.551	1.000	13.074	1.000
partitioned ethnic groups (N=230)	0.635	29.926	2.500	22.787	2.000	21.123	2.000
difference	**0.06**	**10.09**	**1.50**	**8.24**	**1.00**	**8.05**	**1.00**
difference (p-value)	*(0.13)*	*(0.20)*	*(0.01)*	*(0.05)*	*(0.01)*	*(0.01)*	*(0.01)*
Battles with Change in Territory							
all ethnic homelands (N=826)	0.265	2.810	0.000	2.460	0.000	2.096	0.000
non-partitioned ethnic groups (N=596)	0.227	2.183	0.000	1.752	0.000	1.558	0.000
partitioned ethnic groups (N=230)	0.365	4.448	0.000	4.355	0.000	3.498	0.000
difference	**0.14**	**2.26**	**0.00**	**2.60**	**0.00**	**1.94**	**0.00**
difference (p-value)	*(0.00)*	*(0.01)*	*(0.00)*	*(0.00)*	*(0.00)*	*(0.00)*	*(0.00)*
Violence against Civilians							
all ethnic homelands (N=826)	0.608	19.180	1.000	12.196	1.000	12.227	1.000
non-partitioned ethnic groups (N=596)	0.572	17.648	1.000	10.502	1.000	10.381	1.000
partitioned ethnic groups (N=230)	0.700	23.157	3.000	16.758	2.000	17.004	3.000
difference	**0.13**	**5.51**	**2.00**	**6.26**	**1.00**	**6.62**	**2.00**
difference (p-value)	*(0.00)*	*(0.43)*	*(0.00)*	*(0.03)*	*(0.00)*	*(0.01)*	*(0.00)*

Table 1: Test of Means and Medians for Main Civil Conflict Measures

Panel B: Ethnic Homelands close to the National Border

	Indicator Likelihood	Number of Incidents		Excluding Capitals		Excluding Outliers	
	mean	mean	median	mean	median	mean	median
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
All Types of Conflict Events							
all ethnic homelands (N=413)	0.741	49.660	4.000	38.280	3.000	36.643	4.000
non-partitioned ethnic groups (N=200)	0.645	32.300	1.000	29.270	1.000	24.704	1.000
partitioned ethnic groups (N=213)	0.831	65.967	12.000	47.249	10.000	47.957	12.000
difference	**0.19**	**33.67**	**11.00**	**17.98**	**9.00**	**23.25**	**11.00**
difference (p-value)	*(0.00)*	*(0.05)*	*(0.00)*	*(0.10)*	*(0.00)*	*(0.00)*	*(0.00)*
Battles							
all ethnic homelands (N=413)	0.545	23.194	1.000	18.031	1.000	16.186	1.000
non-partitioned ethnic groups (N=200)	0.465	15.700	0.000	13.867	0.000	11.402	0.000
partitioned ethnic groups (N=213)	0.620	30.235	3.000	22.178	2.000	20.724	2.500
difference	**0.15**	**14.53**	**3.00**	**8.31**	**2.00**	**9.32**	**2.50**
difference (p-value)	*(0.00)*	*(0.10)*	*(0.00)*	*(0.17)*	*(0.00)*	*(0.02)*	*(0.00)*
Battles with Change in Territory							
all ethnic homelands (N=413)	0.266	3.281	0.000	2.911	0.000	2.496	0.000
non-partitioned ethnic groups (N=200)	0.170	2.135	0.000	1.633	0.000	1.303	0.000
partitioned ethnic groups (N=213)	0.357	4.357	0.000	4.183	0.000	3.616	0.000
difference	**0.19**	**2.22**	**0.00**	**2.55**	**0.00**	**2.31**	**0.00**
difference (p-value)	*(0.00)*	*(0.04)*	*(0.00)*	*(0.01)*	*(0.00)*	*(0.02)*	*(0.00)*
Violent Indcidents							
all ethnic homelands	0.574	17.910	1.000	13.860	1.000	12.616	1.000
non-partitioned ethnic groups	0.450	11.390	0.000	10.786	0.000	7.525	0.000
partitioned ethnic groups	0.690	24.033	3.000	16.919	2.000	17.393	3.000
difference	**0.24**	**12.64**	**3.00**	**6.13**	**2.00**	**9.87**	**3.00**
difference (p-value)	*(0.00)*	*(0.07)*	*(0.00)*	*(0.14)*	*(0.00)*	*(0.00)*	*(0.00)*

The table reports summary statistics and test of means and medians for the ACLED civil conflict (outcome) variables employed in the empirical analysis at the ethnic homeland level. Panel A reports test of means/medians at the full sample. Panel B reports test of means/medians across ethnic homelands close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland to the national border; 102 kilometers). Column (1) reports the likelihood that a conflict (all conflict incidents, battles, battles resulting in a territorial change, and violence against the civilian population) affects an ethnic homeland. Columns (2)-(3) report the mean and the median value for each type of conflict, respectively. Columns (4)-(5) report the mean and the median value for each type of conflict excluding ethnic regions where capital cities fall. Columns (6)-(7) report the mean and the median value for each type of conflict excluding ethnic regions where the respective variable exceeds the 99th percentile. For each variable the table reports the mean/median value using all ethnic homelands, partitioned ethnicities and non-partitioned ethnicities. The table also reports the mean and median difference and the p-value of mean-median equality between the group of partitioned and non-partitioned ethnicities. The Data Appendix gives detailed variable definitions and data sources.

Table 2 - Border Artificiality

Panel A: Geographical, Ecological and Natural Resources Features

	SPLIT (1)	FRAC (2)	SPLIT (3)	FRAC (4)	SPLIT (5)	FRAC (6)	SPLIT (7)	FRAC (8)
Land Area under Water	0.3219*** (0.0955)	0.0613*** (0.0125)	0.3891*** (0.1012)	0.0683*** (0.0117)	0.3449*** (0.0953)	0.0670*** (0.0122)	0.3352*** (0.0998)	0.0623*** (0.0134)
Land Area	0.0869 (0.0567)	0.0149* (0.0080)	0.1051* (0.0600)	0.0171** (0.0086)	0.0938 (0.0583)	0.0167** (0.0084)	0.0697 (0.0542)	0.0131* (0.0071)
Elevation			-0.0623 (0.1834)	-0.0209 (0.0293)				
St. Dev. Elevation			-0.0001 (0.0006)	0.0000 (0.0001)				
Suitability for Agriculture			0.4494 (0.3328)	0.0621 (0.0491)				
St. Dev. Suit. Agricult.			0.8556 (0.7386)	0.0672 (0.0859)				
Malaria Stability Index					0.1250 (0.2297)	0.0292 (0.0409)		
Distance to the Coast					-0.0001 (0.0002)	0.0000 (0.0000)		
Diamond Mine Indicator							0.1626 (0.1802)	0.018 (0.0287)
Oil Indicator							0.0081 (0.1696)	0.0026 (0.0351)
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.050	__	0.057	__	0.051	__	0.051	__
Adjusted R-squared	__	0.082	__	0.090	__	0.089	__	0.083
Observations	826	826	826	826	826	826	826	826

Table 2 - Panel A reports probit marginal effects (in odd-numbered columns) and OLS estimates (in even-numbered columns) associating ethnic partitioning with geographical, ecological and natural resource variables. In odd-numbered specifications, the dependent variable is an indicator that equals one when at least 10% of the historical ethnic homeland (as portrayed in Murdock's (1959) Ethnolinguistic map) falls to more than one contemporary countries. In even-numbered columns, the dependent variable is a continuous index of ethnic partitioning that reflects the probability that a randomly chosen pixel of the historical homeland of an ethnic group falls into a different country. All specifications 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 2 - Border Artificiality

Panel B: Historical Features

	SPLIT (1)	FRAC (2)	SPLIT (3)	FRAC (4)	SPLIT (5)	FRAC (6)	SPLIT (7)	FRAC (8)
Land Area under Water	0.3298***	0.0624***	0.3210***	0.0613***	0.3162***	0.0605***	0.2829**	0.0514***
	(0.0957)	(0.0123)	(0.0959)	(0.0124)	(0.0980)	(0.0130)	(0.1183)	(0.0162)
Land Area	0.0835	0.0143*	0.0858	0.0149*	0.0817	0.0142*	0.1569**	0.0252***
	(0.0559)	(0.0079)	(0.0564)	(0.0079)	(0.0571)	(0.0081)	(0.0731)	(0.0090)
Slave Exports	0.0225	0.0033						
	(0.0244)	(0.0036)						
Major City in 1400AD			0.0547	-0.0042				
			(0.2080)	(0.0344)				
Distance to Explorer's Routes					-0.0004	0.0000		
					(0.0003)	(0.0000)		
Pre-colonial Settlement Patterns							0.0203	0.0038
							(0.0411)	(0.0055)
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.051	__	0.050	__	0.053	__	0.068	__
Adjusted R-squared	__	0.085	__	0.082	__	0.085	__	0.105
Observations	826	826	826	826	826	826	451	451

Table 2 - Panel B reports probit marginal effects (in odd-numbered columns) and OLS estimates (in even-numbered columns) associating ethnic partitioning with historical variables. In odd-numbered specifications, the dependent variable is an indicator that equals one when at least 10% of the historical ethnic homeland (as portrayed in Murdock's (1959) Ethnolinguistic map) falls to more than one contemporary country. In even-numbered columns, the dependent variable is a continuous index of ethnic partitioning that reflects the probability that a randomly chosen pixel of the historical homeland of an ethnic group falls into a different country. All specifications 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 2 - Border Artificiality

Panel C: Pre-colonial Ethnic Features

	SPLIT (1)	FRAC (2)	SPLIT (3)	FRAC (4)	SPLIT (5)	FRAC (6)	SPLIT (7)	FRAC (8)
Land Area under Water	0.2528**	0.0475***	0.3372***	0.0613***	0.2928**	0.0558***	0.2612**	0.0508***
	(0.1155)	(0.0159)	(0.1109)	(0.0163)	(0.1176)	(0.0162)	(0.1209)	(0.0158)
Land Area	0.1628**	0.0257***	0.1727**	0.0238**	0.1677**	0.0254***	0.1609**	0.0244***
	(0.0721)	(0.0094)	(0.0797)	(0.0097)	(0.0676)	(0.0084)	(0.0681)	(0.0084)
Political Centralization	-0.1965	-0.0321						
	(0.1354)	(0.0208)						
Class Stratification			-0.0242	-0.0038				
			(0.0569)	(0.0069)				
Share of Agriculture					0.0322	0.0048		
					(0.0297)	(0.0055)		
Animal Husbandry							0.0000	0.0004
							(0.0374)	(0.0056)
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.068	__	0.085	__	0.068	__	0.067	__
adjusted R-squared	__	0.106	__	0.128	__	0.107	__	0.090
Observations	437	437	394	394	487	487	487	487

Table 2- Panel C reports probit marginal effects (in odd-numbered columns) and OLS estimates (in even-numbered columns) associating ethnic partitioning with pre-colonial ethnic variables (using data from Murdock (1967)). In odd-numbered specifications, the dependent variable is an indicator that equals one when at least 10% of the historical ethnic homeland (as portrayed in Murdock's (1959) Ethnolinguistic map) falls to more than one contemporary country. In even-numbered columns, the dependent variable is a continuous index of ethnic partitioning that reflects the probability that a randomly chosen pixel of the historical homeland of an ethnic group falls into a different country. All specifications 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: Ethnic Partitioning and Geographic Characteristics within Countries

	Log Land Area	Log Area under Water	Mean Elevation	Land Suitability	Malaria Stability	Distance to the Sea	Diamond Indicator	Oil Indicator	Major City in 1400
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
SPLIT (Partitioning)	-0.1748	-0.0190	0.0179	0.0001	-0.0091	0.0203	-0.0216	-0.0126	-0.0003
Double-clustered s.e.	(0.1092)	(0.0255)	(0.0369)	(0.0194)	(0.0221)	(0.0324)	(0.0224)	(0.0190)	(0.0097)
adjusted R-squared	0.223	0.255	0.575	0.418	0.640	0.620	0.340	0.187	0.063
Mean Dependent Variable	2.0620	0.2706	0.6168	0.4079	0.7209	0.6042	0.0854	0.0947	0.0270
Observations	1182	1182	1182	1182	1182	1182	1182	1182	1182
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Dependent variable is:

Table 3 reports OLS estimates associating various geographical, ecological, and natural resource characteristics with ethnic partitioning within countries. The unit of analysis is an ethnic territory in a country (ethnicity-country). SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. All specifications include country fixed effects (constants not reported). The dependent variable in column (1) is the log of a country-ethnicity's region surface area; in column (2) is the log of (1 + surface area under water); column (3) is average elevation; in column (4) is an index capturing land's (soil) suitability (quality) for agriculture; in column (5) is the average value of a malaria stability index; in columns (6) is the distance of the centroid of each country-ethnic region to the nearest sea coast; in column (7) is a binary index that takes on the value of one if a diamond mine is present; in column (8) is a binary index that takes on the value of one if an oil/petroleum field is present; and in column (9) a binary index that takes on the value of one if a major city was present before European's arrival in Africa (in 1400). The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 4: Ethnic Partitioning and Civil Conflict
Cross-Sectional Estimates

	All Ethnic Homelands						Ethnic Homelands Close to the National Border					
	All Observations				Excl. Outliers	Excl. Capitals	All Observations				Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SPLIT (Partitioning)	0.7590***	0.7250***	0.6740***	0.6207***	0.7093***	0.6813***	0.7076***	0.7136***	0.9580***	0.8492**	0.9430***	0.8812**
Double-clustered s.e.	(0.2928)	(0.2735)	(0.2338)	(0.2234)	(0.2150)	(0.2317)	(0.2473)	(0.1798)	(0.2979)	(0.3316)	(0.2970)	(0.3608)
Log Likelihood	-3221.79	-3204.03	-3159.39	-3151.76	-3095.09	-2847.00	-1593.4	-1578.59	-1560.3	-1552.09	-1498.35	-1422.98
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	826	826	826	826	821	779	413	413	413	413	408	393

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the ethnic homeland level. The dependent variable is the total number of civil conflict incidents at each ethnic homeland over the period 1997-2010. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. The specifications in columns (2)-(6) and (7)-(12) include a set of region fixed effects (constants not reported). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population approximately in 1960. The set of location controls includes the distance of the centroid of each ethnic homeland from the capital, from the sea coast, from the national border, and an indicator that takes on the value one if a capital city falls in the historical homeland. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The specifications in columns (5) and (11) exclude ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (6) and (12) exclude ethnic homelands where capital cities fall. The specifications in columns (7)-(12) focus on ethnic areas close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland to the national border; 102 kilometers). The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 5: Ethnic Partitioning and Civil Conflict
Country-Fixed-Effects Estimates

	All Ethnicity-Country Homelands						Ethnicity-Country Homelands Close to the National Border					
	All Observations				Excl. Outliers	Excl. Capitals	All Observations				Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SPLIT (Partitioning)	0.4929***	0.4748***	0.6731***	0.6185***	0.6284***	0.6171***	1.0208***	0.8465***	0.9258***	0.8502***	0.8502***	0.8388***
Double-clustered s.e.	(0.1559)	(0.2073)	(0.1977)	(0.1876)	(0.1876)	(0.1829)	(0.1767)	(0.3061)	(0.3143)	(0.3145)	(0.3140)	(0.3154)
Log Likelihood	-3942.45	-3708.6	-3615.94	-3603.19	-3498.16	-3340.02	-1556.06	-1419.28	-1393.15	-1384.94	-1377.37	-1322.22
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	1182	1182	1182	1182	1170	1139	576	576	576	576	575	565

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnicity homeland level. The dependent variable is the total number of civil conflict incidents at each ethnic homeland within a country over the period 1997-2010. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. The specifications in columns (2)-(6) and (7)-(12) include country fixed effects (constants not reported). The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country homeland to the national border; 62 kilometers). The specifications in columns (5) and (11) exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (6) and (12) exclude country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 6: Ethnic Partitioning and Civil Conflict
Country Fixed Effects Estimates with Various Measures of Civil Conflict

	All Ethnicity-Country Homelands						Ethnicity-Country Homelands Close to the National Border					
	Battles		Violence		Territorial Change		Battles		Violence		Territorial Change	
	All Obs (1)	No Capitals (2)	All Obs (3)	No Capitals (4)	All Obs (5)	No Capitals (6)	All Obs (7)	No Capitals (8)	All Obs (9)	No Capitals (10)	All Obs (11)	No Capitals (12)
SPLIT (Partitioning)	0.6051***	0.5909***	0.4943***	0.5029***	0.0376*	0.0396*	0.8442***	0.8500***	0.8895***	0.8937***	0.0507*	0.0502*
Double-clustered s.e.	(0.1420)	(0.1426)	(0.1621)	(0.1743)	(0.0199)	(0.0209)	(0.2510)	(0.2537)	(0.2439)	(0.2488)	(0.0280)	(0.0293)
Log Likelihood	-2630.406	-2432.384	-2502.748	-2299.752	—	—	-995.072	-946.214	-938.915	-891.698	—	—
Adjusted R-squared	—	—	—	—	0.455	0.446	—	—	—	—	0.466	0.457
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1182	1139	1182	1139	1182	1139	576	565	576	565	590	565

The table reports Negative Binomial Maximum Likelihood (ML) estimates (in columns (1)-(4) and (7)-(10) and linear probability (LS) estimates associating various aspects of civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in columns (1), (2), (7) and (8) is the total number of battles between government forces, rebel groups, and militias; the dependent variable in columns (3), (4), (9), and (10) is the number of violent events against civilian populations. The dependent variable in columns (5), (6), (11) and (12) is a dummy variable that equals one if a battle resulting in territorial change of control has taken place and zero otherwise. All specifications include country fixed effects (constants not reported). The specifications in columns (7)-(12) focus on ethnicity-country areas close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland within a country to the national border; 62 kilometers). The specifications in even-numbered columns exclude country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis. OLS Specifications

| | All Ethnicity-Country Homelands | | | | | All Ethnicity-Country Homelands Close to the National Border | | | | |
| | OLS ln(1+All Civil Conflict Events) | | | OLS Linear Probabily | | OLS ln(1+All Civil Conflict Events) | | | OLS Linear Probabily | |
	All Obs (1)	No Capitals (2)	No Outliers (3)	All Obs (4)	All Obs (5)	All Obs (6)	No Capitals (7)	No Outliers (8)	All Obs (9)	All Obs (10)
SPLIT (Partitioning)	0.1912**	0.1984**	0.1869**	0.0824***	0.0993***	0.3087***	0.3131***	0.3281***	0.0801*	0.1016*
Double-clustered s.e.	(0.0953)	(0.0916)	(0.0908)	(0.0314)	(0.0336)	(0.1179)	(0.1158)	(0.1197)	(0.0459)	(0.0553)
Adjusted R-squared	0.617	0.593	0.576	0.439	0.447	0.603	0.595	0.574	0.465	0.451
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1182	1170	1139	1182	1182	590	575	565	576	576

The table reports OLS estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnicity level. The dependent variable in columns (1)-(3) and (6)-(9) is the log of one plus the total number of civil conflict incidents in an ethnic region within a country over the period 1997-2010. The dependent variable in columns (4) and (8) is a dummy variable that takes on the value of zero if no conflict has taken place during 1997-2010. The dependent variable in columns (5) and (10) is an indicator that equals zero if there have been at most two conflict incidents (2 is the median value of all conflict incidents over the 1997-2010) in an ethnicity-country. The specifications in columns (6)-(10) focus on country-ethnicity areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country to the national border; 62 kilometers). The specifications in columns (2) and (7) exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (3) and (8) exclude country-ethnic homelands where capital cities fall. All specifications include country fixed effects (constants not reported). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 8: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis. Alternative Measures of Ethnic Partitioning

	All Ethnicity-Country Homelands						All Ethnicity-Country Homelands Close to the Border					
	All Obs (1)	No Outliers (2)	No Capitals (3)	All Obs (4)	No Outliers (5)	No Capitals (6)	All Obs (7)	No Outliers (8)	No Capitals (9)	All Obs (10)	No Outliers (11)	No Capitals (12)
FRAC	0.7463**	0.6824*	0.7422**				0.8122*	0.687	0.8115**			
Double-clustered s.e.	(0.3707)	(0.3883)	(0.3724)				(0.4340)	(0.4237)	(0.4337)			
SPLIT-ALT				0.4974***	0.5119***	0.5178***				0.8614***	0.8412***	0.8612***
Double-clustered s.e.				(0.1621)	(0.1728)	(0.1637)				(0.2377)	(0.2377)	(0.2374)
Log Likelihood	-3612.47	-3349.05	-3507.77	-3608.11	-3344.23	-3502.91	-1392.54	-1329.67	-1384.94	-1387.22	-1324.50	-1379.63
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1182	1139	1170	1182	1139	1170	576	565	575	576	565	575

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnic homeland level. The dependent variable is the total number of civil conflict incidents at each ethnic homeland in each country over the period 1997-2010. FRAC is a continuous measure of ethnic partitioning that reflects the probability that a randomly chosen pixel (area) of the historical homeland of an ethnic group falls into a different country. SPLIT-ALT is an indicator variable that identifies partitioned ethnicities as those with at least 5% of the historical homeland falling into more than one contemporary country. All specifications include country fixed effects (constants not reported). The specifications in columns (7)-(12) focus on country-ethnic areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country to the national border; 62 kilometers). The specifications in columns (2), (5), (8) and (11) exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (3), (6), (9), and (12) exclude country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 9A: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis. Accounting for Location

	Latitude-Longitude Polynomial				Ethnic Family Fixed Effects				Latitude-Longitude Polynomial & Ethnic Family Fixed Effects			
	All Ethnicity-Country Areas		Border Areas		All Ethnicity-Country Areas		Border Areas		All Ethnicity-Country Areas		Border Areas	
	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SPLIT (Partitioning)	0.6298***	0.6182***	0.7750**	0.7692**	0.5568***	0.5562**	0.7434**	0.7200*	0.5354**	0.5110**	0.6207*	0.5770*
Double-clustered s.e.	(0.1923)	(0.1941)	(0.3046)	(0.3068)	(0.2182)	(0.2335)	(0.3792)	(0.3814)	(0.1984)	(0.2074)	(0.3346)	(0.3335)
Log Likelihood	-3582	-3317.33	-1354.67	-1293.51	-3467.48	-3209.94	-1279.21	-1219.5	-3451.45	-3195.16	-1269.92	-1210.46
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1182	1139	576	565	1182	1139	576	565	1182	1139	576	565

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnic homeland level. The dependent variable is the total number of civil conflict incidents in each ethnic homeland in each country over the period 1997-2010. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. All specifications include country fixed effects (constants not reported). The specifications in columns (3), (4), (7), (8), (11), and (12) focus on areas close to the national border (using as a cutoff the median distance from the centroid of each country-ethnic homeland to the national border; 62 kilometers). The specifications in even-numbered columns exclude country-ethnic homelands where capital cities fall. The specifications in (1)-(4) and (9)-(12) include a cubic polynomial in latitude and longitude. The specifications in (5)-(12) include a set of ethnic family fixed effects (constants not reported). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 9B: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis. Excluding Each Time a Different African Region

	Excluding									
	North		South		West		East		Central	
	All	Border	All	Border	All	Border	All	Border	All	Border
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SPLIT (Partitioning)	0.5593*	0.7990*	0.6595***	0.8090**	0.5548***	1.2937***	0.7961***	0.8755**	0.4311**	0.4296**
Double-clustered s.e.	(0.3213)	(0.3844)	(0.1183)	(0.3297)	(0.1974)	(0.2514)	(0.1507)	(0.4904)	(0.1371)	(0.2000)
Log Likelihood	-3416.154	-1354.598	-3143.062	-1237.219	-2617.48	-894.705	-2558.178	-1026.955	-2635.134	-993.766
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1104	554	1040	508	808	350	893	456	883	436

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnic homeland level. The dependent variable is the total number of civil conflict incidents in an ethnic homeland within a country over the period 1997-2010. In columns (1)-(2) we exclude ethnicity-country observations that fall in North Africa. In columns (3)-(4) we exclude observations that fall in South Africa. In columns (5)-(6) exclude observations that fall in West Africa. In columns (7)-(8) we exclude observations that fall in East Africa. In column (9)-(10) we exclude observations that fall in Central Africa The regional classification follows Nunn (2007). All specifications include country fixed effects (constants not reported). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 10: Ethnic Partitioning and Civil Wars (1970-2005)
PRIO Dataset

	Civil War Incidents		Civil War Main Locations	
	(1)	(2)	(3)	(4)
SPLIT - Partitioning	0.2189***	0.1579***	0.2434***	0.1616***
Double-clustered s.e.	(0.0557)	(0.0526)	(0.0615)	(0.0560)
Log Likelihood	-846.11	-809.689	-958.417	-907.662
Simple Controls	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Polynomial Latitude & Longitude	Yes	Yes	Yes	Yes
Ethnic Family Fixed Effects	Yes	Yes	Yes	Yes
Region Fixed Effects	Yes	No	Yes	No
Country Fixed Effects	No	Yes	No	Yes
Observations	826	822	826	822

The table reports Poisson Maximum Likelihood (ML) estimates associating various measures of civil war with ethnic partitioning. The dependent variable in columns (1)-(2) is the number of civil wars that have taken place in the historical homeland of an ethnic group between 1970 and 2005. The dependent variable in columns (3)-(4) is the number of conflict zones associated with civil wars that have affected the historical homeland of an ethnic group during the period 1970-2005. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland belonging to more than one contemporary country. The specifications in odd-numbered columns include a set of region fixed effects (constants not reported); the specifications in even-numbered columns include a set of country fixed effects (constants not reported). The assignment of the country fixed effects is based on the centroid of each ethnic homeland.

All specifications include a cubic polynomial in latitude and longitude of the centroid of each ethnic group. All specifications include a rich set of conditioning variables, namely log land area, log land area under water (lakes, rivers, and other streams), log population around independence, the distance of each ethnic homeland to the national border, the distance to the capital city, the distance to the closest sea coast, land suitability for agriculture, mean 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 indicator. 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 11: Ethnic Partitioning and Civil Conflict
Results with Ethnologue's Mapping of Ethnic Groups

	All Language-Country Homelands						Language-Country Homelands Close to the National Border					
	All Observations				Excl. Outliers	Excl. Capitals	All Observations				Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: NB ML Estimates												
SPLIT (Partitioning)	0.7170***	0.4918**	0.4421***	0.3603***	0.3383***	0.3516***	0.5262***	0.3525***	0.3619***	0.3191***	0.2672***	0.3157***
Double-clustered s.e.	(0.1288)	(0.0968)	(0.0836)	(0.0899)	(0.0923)	(0.0905)	(0.1490)	(0.1230)	(0.0967)	(0.0948)	(0.1029)	(0.0961)
Log Likelihood	-4850.15	-4489.32	-4414.53	-4365.37	-4149.16	-4022.85	-2338.24	-2081.06	-2061.75	-2040.4	-1970.1	-1904.15
Panel B: LS Estimates												
SPLIT (Partitioning)	0.2467***	0.1496**	0.0747**	0.0903**	0.0803**	0.0856**	0.1729***	0.0843**	0.0794*	0.1097***	0.0933**	0.1148***
Double-clustered s.e.	(0.0586)	(0.0664)	(0.0377)	(0.0374)	(0.0334)	(0.0364)	(0.0172)	(0.0390)	(0.0440)	(0.0388)	(0.0397)	(0.0385)
Adjusted R-squared	0.381	0.546	0.597	0.609	0.569	0.54	0.283	0.537	0.587	0.602	0.579	0.542
Country FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	2405	2405	2405	2405	2379	2353	1202	1202	1202	1202	1194	1182

The table reports Negative Binomial Maximum Likelihood (ML) estimates (in Panel A) and OLS estimates (in Panel B), associating civil conflict incidents with ethnic partitioning at the country-language homeland level using the spatial distribution of languages across Africa according to the Ethnologue database. The dependent variable in Panel A is the total number of civil conflict incidents in an ethnic homeland within a country over the period 1997-2010. The dependent variable in Panel B is the log of one plus the total number of civil conflict incidents at each ethnic homeland in each country over the period 1997-2010. SPLIT is an indicator variable that identifies partitioned groups as those that Ethnologue maps into more than one contemporary country. The specifications in columns (2)-(6) and (7)-(12) include country fixed effects (constants not reported). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 2000. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 1: Correlation Structure of the Main Dependent Variables

Panel A: Ethnic Homeland Level

All types of ethnic conflict	1			
Battles	0.8733*	1		
Violent events against civilians	0.9049*	0.6068*	1	
Battles resulting is change of territory (Indicator)	0.2529*	0.3164*	0.1422*	1

Panel B: Country-Ethnicity Homeland Level

All types of ethnic conflict	1			
Battles	0.8268*	1		
Violent events against civilians	0.8839*	0.4965*	1	
Battles resulting is change of territory (Indicator)	0.2462*	0.3208*	0.1209*	1

The table reports the correlation structure between all ACLED civil conflict measures. Panel A reports the correlogram at the ethnic homeland level (*N=826*); Panel B reports the correlogram at the country-ethnicity level (*N=1182*). * indicates statistical significance at the 95% confidence level. The Data Appendix gives detailed variable definitions and data sources.

Appendix Table 2: Summary Statistics

	Obs.	mean	st. dev.	min	p25	median	p75	p99	max
				Panel A: Ethnic Homeland Level					
All Civil Conflict Incidents	826	52.39	195.69	0.00	1.00	4.00	34.00	580.00	2916.00
Battles	826	22.65	94.44	0.00	0.00	1.00	13.00	313.00	1608.00
Battles with Territorial Change	826	19.18	96.20	0.00	0.00	1.00	9.00	231.00	2009.00
Violent Events against Civilians	826	0.27	0.44	0.00	0.00	0.00	1.00	1.00	1.00
Population at Independence	826	328,991	1,063,156	57	41,810	118,160	303,659	2,912,382	25,700,000
Land Area	826	34.17	59.18	0.24	6.16	14.48	36.07	286.33	604.90
Land Area under Water	826	0.86	2.26	0.00	0.01	0.17	0.68	10.87	27.66
Catital City Indicator	826	0.06	0.23	0.00	0.00	0.00	0.00	1.00	1.00
Distance to the Capital City	826	141.15	126.00	0.00	44.87	102.25	204.98	536.73	636.87
Distance to the Sea Coast	826	499.98	371.34	11.31	255.59	391.81	629.83	1673.47	1846.93
Distance to the National Border	826	597.49	432.10	0.22	208.79	554.35	918.12	1609.77	1721.30
Mean Elevation	826	0.62	0.44	0.00	0.30	0.49	0.94	1.84	2.17
Land Suitability for Agriculture	826	0.41	0.24	0.00	0.26	0.42	0.57	0.93	0.98
Malaria Stability Index	826	0.75	0.36	0.00	0.58	0.98	1.00	1.00	1.00
Oil Indicator	826	0.13	0.40	0.00	0.00	0.00	0.00	2.00	4.00
Diamond Mine Indicator	826	0.12	0.33	0.00	0.00	0.00	0.00	1.00	1.00
				Panel B: Country-Ethnic Homeland Level					
All Civil Conflict Incidents	1182	33.63	139.63	0.00	0.00	2.00	18.00	438.00	2888.00
Battles	1182	14.54	67.09	0.00	0.00	0.00	6.00	238.00	1608.00
Battles with Territorial Change	1182	1.94	7.88	0.00	0.00	0.00	0.00	1.00	100.00
Violent Events against Civilians	1182	12.25	72.43	0.00	0.00	1.00	5.00	182.00	2009.00
Population at Independence	1182	222,093	880,884	17	15,528	62,931	186,570	2,143,565	25,600,000
Land Area	1182	22.41	40.68	0.11	2.81	8.33	23.29	216.23	493.82
Land Area under Water	1182	0.51	1.39	0.00	0.00	0.10	0.40	6.18	21.63
Catital City Indicator	1182	0.04	0.19	0.00	0.00	0.00	0.00	1.00	1.00
Distance to the Capital City	1182	0.51	0.36	0.01	0.26	0.41	0.63	1.70	1.88
Distance to the Sea Coast	1182	0.60	0.43	0.00	0.23	0.56	0.93	1.63	1.74
Distance to the National Border	1182	0.11	0.12	0.00	0.02	0.06	0.16	0.52	0.64
Mean Elevation	1182	0.62	0.44	0.00	0.29	0.49	0.95	1.91	2.18
Land Suitability for Agriculture	1182	0.41	0.24	0.00	0.25	0.42	0.57	0.94	0.98
Malaria Stability Index	1182	0.72	0.34	0.00	0.50	0.89	1.00	1.00	1.00
Oil Indicator	1182	0.09	0.28	0.00	0.00	0.00	0.00	1.00	1.00
Diamond Mine Indicator	1182	0.09	0.29	0.00	0.00	0.00	0.00	1.00	1.00

The table reports descriptive statistics for all variables employed in the empirical analysis. Panel A reports summary statistics for all control variables at the ethnic homelad level. Panel B reports summary statistics for all control variables at the country-ethnicity sample. The Data Appendix gives detailed variable definitions and data sources. The Data Appendix gives detailed variable definitions and data sources.

Appendix Table 3: Pre-colonial Ethnic Features and Ethnic Partitioning

	Specification A Unconditional Relationship		Specification B Conditional Relationship	
	Additional Variable (1)	Obs. (2)	Additional Variable (3)	Obs. (4)
Gathering	0.0128 (0.0524)	487	-0.0682 (0.0486)	487
Hunting	0.0415 (0.0483)	487	0.0388 (0.0452)	487
Fishing	0.0234 (0.0556)	487	-0.004 (0.0518)	487
Milking	-0.0116 (0.0552)	452	0.0118 (0.0377)	452
Agriculture	-0.0109 -0.0108	487	0.0104 -0.0102	487
Alternative Argiculture Dep.	-0.0381 -0.0283	452	-0.0231 -0.0322	452
Polygyny	-0.0174 (0.0453)	478	-0.0051 (0.0503)	478
Clan Communities	0.0197 (0.0609)	396	0.0385 (0.0614)	396
Complex Settlements	0.0209 (0.0607)	451	0.017 (0.0528)	451
Binary Class Stratification	-0.0158 (0.0495)	394	-0.0754 (0.0457)	394
Jurisdictional Hierarchy of Local Community	0.0221 -0.0395	440	0.0061 -0.0319	440
Elections	-0.0663 (0.0822)	344	-0.0542 (0.0903)	344
Inheritance Rule for Property	-0.0821 (0.0927)	374	0.0211 (0.0898)	374

The table reports OLS (linear probability model) estimates associating ethnic partitioning with pre-colonial ethnic-specific variables (using data from Murdock (1967)). In all specifications the dependent variable is an indicator that equals one when at least 10% of the historical ethnic homeland (as portrayed in Murdock's (1959) Ethnolinguistic map) falls to more than one contemporary country. Specifications A are simple unconditional models. Specifications B include a set of region fixed effects (constants not reported), log land area under water, and log land area. 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.

Variable Definitions:
Gathering: Binary index that indicates the reliance of the economy on "the collection of wild plans and small land fauna." The index equals zero when the dependence is between 0% and 5%; the index equals one when dependence is greater than 5% dependence. Source: Murdock (1967); variable code in the Ethnographic Atlas v1.
Hunting: Binary index that indicates the intensity in hunting (including trapping and fowling). The index equals zero when the dependence is between 0% and 5%; the index equals one when dependence is greater than 5%. Source: Murdock (1967); variable code in the Ethnographic Atlas v2.
Fishing: Binary index that indicates the intensity in fishing (including shell fishing and the pursuit of large aquatic animals). The index equals zero when the dependence is between 0% and 5%; the index equals one when dependence is greater than 5%. Source: Murdock (1967); variable code in the Ethnographic Atlas v3.
Milking: Binary index that equals zero when "domestic animals are milked more often that sporadically" and zero when "little or no milking". Source: Murdock (1967); variable code in the Ethnographic Atlas v41.
Agriculture: 0-9 scale index reflecting the intensity of agriculture. "It includes penetration of the soil, planting, tending the growing crops, and harvesting but not subsequent food preparation". The index equals 0 when there 0%-5% dependence; 1 when there is 6%-15% dependence; 2 when there is 16%-25% dependence; 3 when there is 26%-35% dependence; 4 when there is 36%-45% dependence; 5 when there is 46%-55% dependence; 6 when there is 56%-65% dependence; 7 when there is 66%-75% dependence; 8 when there is 76%-85% dependence; and 9 when there is 86%-100% dependence. Source: Murdock (1967); variable code in the Ethnographic Atlas v5.
Agriculture Type: 0-4 scale index reflecting the type of agriculture. The index equals 0 when there is "no agriculture"; 1 when there is "causal agriculture"; 2 when there is "extensive or shifting agriculture"; 3 when there is "intensive agriculture"; and 4 when there is "intensive irrigated agriculture." Source: Murdock (1967); variable code in the Ethnographic Atlas v28.
Polygyny: Indicator that equals one when polygyny is practised and zero otherwise. The indicator equals one when the original variable indicates that polygyny is common and when large extended families are present. Source: Murdock (1967); variable code in the Ethnographic Atlas v8.
Binary Class Stratification:The dummy stratification index equals zero when Murdock's variable equals zero indicating "absence of significant class distinctions among freemen, ignoring variations in individual repute achieved through skill, valor, piety, or wisdom," and one when Murdock's class stratification measure equals 1, 2, 3, or 4. The construction of this variable follows Gennaioli and Rainer (2006, 2007). Source: Murdock (1967); variable code in the Ethnographic Atlas v67.
Clan Communities: Indicator that equals one when Murdock's community marriage organization variable (v15) equals 6 ("clan communities or clan barrios") and zero otherwise. Source: Murdock (1967); variable code in the Ethnographic Atlas v15.
Complex Settlements: Indicator that equals one for ethnicities living in compact and relatively permanent settlements (v30=7) or in complex settlements (v30=8), and zero otherwise. Source: Murdock (1967); variable code in the Ethnographic Atlas v30.
Jurisdictional Hierarchy of Local Community: Ordered variable ranging from 0 to 2 reflecting the hierarchy of local community organization. A zero score indicates the theoretical minimum of two (e.g., family and band), while a score of 2 indicates the theoretical maximum of four levels (e.g., nuclear family, extended family, clan barrio, village levels). Source: Murdock (1967); variable code in the Ethnographic Atlas v32.
Elections: Indicator that equals 1 when succession to the office of the local headman is conducted via "election or other formal consensus, nonhereditary" and zero otherwise. Source: Murdock (1967); variable code in the Ethnographic Atlas v72.
Inheritance Rule for Property: Indicator that equals one when some form of inheritance rule of real property (land) is present; the binary indicator equals zero when there is "absence of individual property rights". Source: Murdock (1967); variable code in the Ethnographic Atlas v74; the construction of the index follows Fenske (2009).

Appendix Table 4: Ethnic Partitioning and Alternative Measures of Civil Conflict
Cross-Sectional Estimates

	All Ethnic Homelands						Ethnic Homelands Close to the National Border					
	All Observations				Excl. Outliers	Excl. Capitals	All Observations				Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Battles												
SPLIT (Partitioning)	0.8545***	0.8543***	0.8536***	0.6579***	0.7706***	0.7288***	0.6687**	0.6884***	0.7843**	0.5764	0.6919*	0.6408
Double-clustered s.e.	(0.2970)	(0.2516)	(0.2305)	(0.2240)	(0.1991)	(0.2363)	(0.2788)	(0.1984)	(0.3383)	(0.3880)	(0.3456)	(0.4308)
Log Likelihood	-2483.17	-2452.05	-2418.24	-2406.04	-2349.76	-2171.71	-1213.03	-1193.44	-1176.18	-1165.08	-1113.3	-1060.28
Panel B: Violence against Civilians												
SPLIT (Partitioning)	0.6025***	0.5750***	0.5300**	0.5327**	0.6435***	0.6119**	0.6184**	0.6389***	0.9951***	0.9688***	1.1552***	0.9986***
Double-clustered s.e.	(0.2580)	(0.2489)	(0.2592)	(0.2507)	(0.2659)	(0.2549)	(0.2182)	(0.1913)	(0.2818)	(0.3062)	(0.2950)	(0.3186)
Log Likelihood	-2342.88	-2316.08	-2296.55	-2294.53	-2241.91	-2049.53	-1159.94	-1144.5	-1137.28	-1133.69	-1081.03	-1030.17
Panel C: Battles that Resulted in a Change of Territory												
SPLIT (Partitioning)	0.1037**	0.1165***	0.1181***	0.1191***	__	0.1364***	0.1156**	0.1272***		0.0905**	__	0.0989**
Double-clustered s.e.	(0.0490)	(0.0331)	(0.0414)	(0.0374)		(0.0395)	(0.0468)	(0.0390)	(0.0474)	(0.0365)		(0.0398)
Adjusted R-squared	0.091	0.225	0.257	0.278		0.281	0.1	0.199	0.235	0.283		0.282
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	826	826	826	826	821	779	413	413	413	413	408	393

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the ethnic homeland level. The dependent variable in Panel A, B, C is the total number of battles, violence against civilians, battles resulting in territorial change at each ethnic homeland over the period 1997-2010, respectively. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. The specifications in columns (2)-(6) and (7)-(12) include region fixed effects (constants not reported). The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country homeland to the national border; 102 kilometers). The specifications in columns (5) and (11) exclude ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (6) and (12) exclude ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each ethnic homeland from the capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table A: Partitioned Ethnicities

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

Group	Val	Country	N	Group	Val	Country	N
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 A 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 2.1 gives details on our approach in identifying partitioned ethnicities.