

The Long-Run Effects of the Scramble for Africa*

Stelios Michalopoulos
Brown University and NBER

Elias Papaioannou
London Business School, NBER and CEPR

December 15, 2012

Abstract

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 apart from the land mass and water area of an ethnicity's historical homeland, no other geographic, economic, and historical trait, including proxies of pre-colonial conflict, predicts partitioning by the national borders. Second, we exploit a detailed geo-referenced database that records various types of conflict across African regions and show that civil conflict is concentrated in the historical homeland of partitioned ethnicities. We also document that violence against civilians (child soldiering, village burning, abductions, rapes) and territorial changes between rebel groups, militias, and government forces are more prevalent in the homelands of split groups. These results are robust to a rich set of local controls, the inclusion of country fixed effects and ethnic-family fixed effects. The uncovered evidence brings in the foreground the violent repercussions of an important aspect of European colonization, that of ethnic partitioning.

Keywords: Africa, Borders, Ethnicities, Conflict, Development,
JEL classification Numbers: O10, N17, N97, Z10.

*We thank Alberto Alesina, Maarten Bosker, Francesco Caselli, Jeremiah Dittmar, James Fenske, Federico Finan, Horacio Larreguy Arbesú, Stathis Kalyvas, Janina Matuszeski, Ted Miguel, Nathan Nunn, Gregorios Siourounis, Neils Weidman, Andreas Wimmer, and participants at the AEA meetings in Denver, Yale, UC Berkeley, Princeton, ALBA, UC Irvine, UC Merced, American University, Brown University, Institute for Economic Analysis and Autonomia University, George Washington University, CERGE-EI, the NBER Political Economy meetings, the NBER Summer Institute Meetings on the Development of the American Economy and Income Distribution and Macroeconomics for useful comments and suggestions. We are also thankful to three referees of this journal for their constructive suggestions and comments. All errors are our sole responsibility.

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 -if any- 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 are partitioned by the national border.¹

A considerable body of work (e.g., Asiwaju (1985); Dowden (2008); Wesseling (1996)) argues that the main channel of Europeans' influence on African development was not colonization per se, but the improper border design. As Herbst (2000) summarizes "*for the first time in Africa's history [at independence], territorial boundaries acquired salience...The boundaries were, in many ways, the most consequential part of the colonial state.*" According to the African historiography, the artificial borders have led to struggles, patronage politics, violence, and conflict, mostly by partitioning ethnic groups across more than one contemporary state. Ethnic partitioning has led to irredentism and has helped create an ideology of secession and nationalism (Horowitz (1985)). Besides partitioning, Africa's border design produced some of the largest and most heterogeneous countries in the world, such as the Democratic Republic of Congo and Sudan. Moreover, the Scramble for Africa created many hinterland countries and states with challenging geographies that limit the national governments' ability to broadcast power effectively (Herbst (2000), Collier (2007)).

Yet, there is little work formally examining the ramifications of ethnic partitioning in the context of the Scramble for Africa. Some promising cross-country studies have touched upon this issue, showing, for example, that the likelihood of conflict increases when there is an ethnic war in adjacent countries (Bosker and de Ree (2010)) and that countries with straight borders where a large share of the population belongs to ethnicities that are also present in nearby nations perform economically worse (Alesina, Easterly, and Matuszeski (2011)). Nevertheless, to the best of our knowledge there is no empirical work exploring the consequences of ethnic partitioning for groups in Africa where the idiosyncratic border design offers the

¹Asiwaju (1985) identifies 177 partitioned ethnic groups that span all African borders. Englebert, Tarango, and Carter (2002) estimate that partitioned ethnic groups constitute on average 40% of the total population and 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%)).

opportunity to identify causal effects. This study is a step in this direction.

While there is little disagreement among historians that colonial (and hence post-independence) borders were drawn with little respect to the local political geography, we start our analysis establishing formally this thesis. We do so exploring whether partitioned ethnicities differ systematically from non-split groups across a variety of geographic, ecological, and pre-colonial social, economic, and cultural traits. With the sole exceptions of the size of the historical ethnic 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), the presence of natural resources (diamond mines and oil fields), and measures of early contact with colonizers. We further show that there are no systematic differences, on average, 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, proxies of urbanization (Murdock (1967)), etc. Most importantly, we show that there is no association between ethnic partitioning and two distinct proxy measures of pre-colonial conflict, one based on pre-colonial wars mostly between African kingdoms and the other capturing conflict during the slave trades. Our large-scale econometric results thus support the claim of the African historiography that in the overwhelming majority of cases, Europeans did not take into account local political and geographic conditions while designing the borders.

We then employ the scramble for Africa as a quasi-natural experiment to assess the 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 homelands of partitioned ethnicities. Our regional focus allows us to account for the numerous country-wide factors that interact with civil conflict, such as national institutions and politics, the type of colonization, fractionalization, etc. Furthermore, by moving the analysis from the national level to the ethnic homeland, we are able to condition on ethnic-family fixed effects and thus take into consideration cultural, genetic, social, and economic differences across the quite heterogeneous African ethnicities. We obtain similar results when we restrict estimation to ethnic homelands close to national borders. Our most conservative estimates suggest that civil conflict intensity is approximately 30% 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 roughly 40% higher in regions where split groups reside. The evidence thus uncovers the on-going violent repercussions of the colonial border design.

1.1 Historical Background

The "Scramble for Africa" starts in the 1860s when the French and the British begin the systematic exploration of Western Africa and sign bilateral agreements assigning spheres of influence. In the next 40 years, Europeans 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.² 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 alive. As a result, in the overwhelming majority of cases, European powers drew borders without taking into account local conditions. African leaders were not invited and had no say.³ In many cases, European leaders were in such a rush that they didn't wait for the information arriving from explorers, geographers, and missionaries. 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 trod; 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.*" 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 idiosyncratic border design. First, at the time Europeans had very little knowledge of local geographic conditions, as with the exception of very few coastal areas, the continent was largely unexplored. Second, Europeans were not drawing borders of prospective states or -in many cases- even colonies (in most cases Europeans ruled Africa indirectly via protectorates where they had minimal presence). 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.⁴ Fourth, while in most cases the treaties indicated that the exact boundaries would be set by special

²Three 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.

³Asiwaju (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*".

⁴An illustrative example is the annexation of Katanga in Congo Free State that turned out to be its richest province. King Leopold got Katanga in exchange for the Niari-Kwilu area that the French insisted on getting themselves. Wesseling (1996) writes "*what impelled him [Leopold] was a general imperialist surge, the desire for compensation for the Niari-Kwilu, and the*

commissions, demarcation was poor. Fifth, Europeans were not willing to sacrifice their commitment not to go to war for any part of Africa and hence were reluctant to change colonial borders.⁵ 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, as little changed on the ground. 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.*" Wesseling (1996) summarizes: "*The partition of Africa was recorded by the Europeans on their maps, but the matter rested there for the time being....In Europe conquests preceded the drawing of maps; in Africa the map was drawn, and then it was decided what was going to happen. These maps did not therefore reflect reality but helped to create it.*"

The other major event in recent African history, namely the wave of independence, occurred at a speed that not even the key protagonists expected (Herbst (2000)). 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 borders, African leaders and leaving Europeans did not touch this issue. The leaders of African independence believed that nation building and industrialization would sideline ethnic divisions. Moreover, national leaders feared that a border realignment would threaten their position whereas Europeans' main objective was to maintain the special rights and corporate deals with their former colonies, and, as such, they were also reluctant to open the border issue.⁶

1.2 Channels and Case Studies

Irredentism, secession, and autonomy The literature has mostly stressed the impact of ethnic partitioning on generating irredentist demands, as split ethnicities usually want to unify with their peers across the border.⁷ Somali tribes, for example, 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

objective of making the new state as large as possible and filling as much of the Congo basin as possible."

⁵Wesseling (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.*"

⁶Almost 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.

⁷Horowitz (1985) notes "*a quick tour d'horizon reveals the rich range of possibilities (for conflict and irredentism). The Ghana-Togo border divides the Ewe, as the Nigeria-Benin border divides the Yoruba. There are Hausa in Nigeria and Hausa in Niger. There are Fulani across a wide belt of West and Central Africa, Beteke in Gabon and Congo (Brazzaville), and Fang in Cameroon, Gabon, and Equatorial Guinea. The Bakongo are divided among, Zaire, Congo (Brazzaville) and Angola; the Lunda among Zaire, Zambia, and Angola. There are Somalis in Somalia, Ethiopia,, Kenya, and Djibouti. There are Wolof in Mauritania, in Gambia, and in Senegal, Kakwa in Sudan and in Uganda. And various Berber groups are distributed among more than one North African state.*"

by Somali tribes (Somalia, North Kenya, Southern Ethiopia, Djibouti, and Eritrea); at least three long-lasting wars have been (partly at least) driven by the desire of Somalis in Ethiopia, Djibouti, and Kenya to become part of Somalia (Meredith (2005)). 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 clans are located (while only 7% of Ethiopia’s total population resides in these regions). Similarly, in the initial years after independence Kenya experienced substantial conflict in the Northern Frontier District as Somalis were fighting (*shifita*) for annexation to Somalia (Touval (1967)).

In line with this argument, Wimmer, Cederman, and Min (2009) estimate that 20% of all civil wars in Africa have a secessionist demand. Unlike non-split groups, partitioned ethnicities have the advantage of getting assistance from their peers on the other side of the border as well as find a safe heaven when in need to retreat and regroup. 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 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. 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, although only 5% of Senegal’s population resides there.⁸

Even when partitioning does not lead to secessionist demands, it may still spur conflict via shaping ethnic politics. As Horowitz (1985) notes ethnic partitioning creates an ideology that local ethnic-specific regional parties exploit in pursuit of their special interest. This is particularly pronounced for partitioned groups, as 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.

⁸The conflict in the Casamance region illustrates the usefulness of the ACLED data. While this conflict has resulted in 3,000 – 5,000 deaths and 50,000 (internally and externally displaced) refugees over the period 1980 – 2000, this event is not-classified as a war in most civil war datasets as deaths spread over a long period of time and the casualties threshold is not met in any particular year (Humphreys and Mohamed (2005)).

Suffocation and Cross-Border Migration Ethnicities in Africa are often targets of the national government. What is different between partitioned and non-split groups is that the latter can seek shelter at their homeland at the other side of the border. Not only this leads directly to conflict, but the population displacement and the refugee flows may spread conflict in nearby counties. The conflict in the Alur-land offers an illustration of this channel. 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). When Mobutu Sese Seko initiated the subjugation of many minority groups in Zaire, a large portion of the Alur were pushed to their homeland in Uganda. This in turn generated opposition from the Buganda in Southern Uganda leading to conflict. In line with this case, Fearon and Laitin (2011) report that 31% of all civil wars (and 57% of all ethnic wars) involve "*members of a regional ethnic group that considers itself to be the indigenous sons-of-the-soil and recent migrants from other parts of the country*".

Other Aspects of the Scramble for Africa The artificial border design has contributed to underdevelopment and 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 heterogeneity, polarization, land size, access to the sea that affect development. 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 power (see also Michalopoulos and Papaioannou (2012a)). Collier (2007) discusses how the border design resulted in Africa having the largest proportion of landlocked countries limiting their growth potential. While our analysis focuses on a single 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 absorb all time-invariant, country-specific characteristics. Nevertheless, in the last section, we explore how these different nationwide by-products of border drawing interact with ethnic partitioning in shaping local conflict.

1.3 Related Literature

Our paper belongs to the genre of studies that investigate the historical origins of comparative development. The literature has mainly focused on the impact of colonization mainly via the formation of early institutions (e.g., Acemoglu, Johnson, and Robinson (2005)), infrastructure (e.g., Huillery (2009)), and human capital (e.g., Glaeser, LaPorta, de Silanes, and Shleifer (2004) and Easterly and Levine (2012)). Complementing these works, we emphasize 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, under-perform economically compared to more "organic" countries.

Moreover, our work fits a growing strand of literature that examines the roots of African underde-

velopment and conflict. Nunn (2008), Nunn and Puga (2012), and Nunn and Wantchekon (2011) show that the slave trades have crucially shaped African development mostly by spurring ethnic conflict and by lowering trust. Gennaioli and Rainer (2006, 2007) and Michalopoulos and Papaioannou (2012b) show that deeply-rooted pre-colonial ethnic institutions correlate significantly with contemporary economic development. A closely related study is that of Besley and Reynal-Querol (2012), who use geo-referenced data of pre-colonial conflict and show that the latter is associated with a greater prevalence of contemporary conflict and lower levels of trust.

Our work also 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 are works that link a country's ethnic composition to civil war. 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 association 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. Englebert, Tarango, and Carter (2002) show a positive cross-country correlation between proxy measures of suffocation and dismemberment and violence, secession attempts, border disputes, and warfare.⁹

The correlations found in studies linking cross-country variation in border features and the distribution of ethnicities to development proxies (income or conflict) are informative; yet they cannot be easily interpreted (see Blattman and Miguel (2010) for a discussion). The main concern is that the process of border drawing is usually an outcome of state formation that naturally affects directly economic performance and conflict and is related to voluntary and forced peoples' movements as well. Moreover, as the recent literature on state capacity shows, nation building, development, and conflict are inter-linked and driven by hard-to-account-for factors related to societal structure and geography (see Besley and Persson (2011)). Thus, issues of selection, reverse causation, and omitted variables are non-negligible in cross-country works. Likewise, due to measurement error in the main independent variables, multi-collinearity, and the limited degrees of freedom, the correlations documented in cross-country studies are often quite sensitive even to small permutations and data revisions (see Hegre and Sambanis (2006) and especially Ciccone and Jarocinski (2010), who show that the identification of the correlates of cross-country growth are fragile even to minimal data revisions).

⁹Our paper also makes contact with the voluminous literature in political science and security studies focusing on nationalism. Using mostly case-studies and narrative arguments this body of research examines the origins and consequences of ethnic identification, nationalism, and irredentism (e.g. Fearon and Laitin (2000), Brancati (2006)), as well as the pros and cons of regional separation as a solution to ethnic conflict (e.g. Horowitz (1985), Kaufman (1996, 1998), Fearon (2004), Englebert (2009)). See also Sambanis (2000), Chapman and Roeder (2007), and Sambanis and Schulhofer-Wohl (2009) for cross-country regression evidence examining whether the separation of regions in response to civil war, like the one in Sudan or the secession of Eritrea from Ethiopia result in long-lasting peace.

By exploiting variation across ethnic homelands, we account for many of the shortcomings of existing cross-country works. First, by showing that there are no systematic differences on average in geographic, economic, institutional, and cultural characteristics between partitioned and non-split ethnicities, our analysis offers large-scale econometric evidence supportive to the African historiography on the accidental nature of most African borders. Moreover, our results that on average there is no association between ethnic partitioning and various proxies of pre-colonial conflict and the plethora of case studies documenting the idiosyncratic nature of African borders enhances the plausibility of our research design. 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. Since borders were drawn by Europeans, we are able to focus on cases where country boundaries were not the result of political, economic, or military developments (and thus selection and reverse causation issues are secondary -if not absent). Third, our analysis at the ethnic homeland is conceptually appealing in the context of Africa, where ethnic identification is especially strong and conflict has a strong ethnic component. In their synthesis of the case-study evidence on conflict in Africa and the results of cross-country regressions, Collier and Sambanis (2005) note "*the country-year is not the appropriate unit of observation to study such wars. Instead it would be more appropriate to focus on the ethnic group or we should analyze patterns of violence in a geographical region that does not necessarily correspond to predefined national boundaries. With current data limitations, however, it may not be feasible to adjust the unit of analysis problem.*"

Structure In the next section we first discuss how we identify partitioned ethnic groups and present the geo-referenced civil conflict data. We then report descriptive statistics illustrating the significant differences in the likelihood and intensity of armed conflict between partitioned and non-split ethnic homelands. 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 that may independently affect conflict. Section 4 reports our baseline estimates on the effect of ethnic partitioning on various aspects of civil conflict. In Section 5 we report the results of our sensitivity analysis and explore heterogeneous effects. 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).¹⁰ 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 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 state (Figure 1b). Appendix Table A lists partitioned ethnicities. When we use a broader threshold of 5% we identify 267 partitioned groups. 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.¹¹

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 Asiwaaju (1985), who provides the only (to our knowledge) codification of partitioned ethnicities in

¹⁰We also drop resulting from the intersection of the two maps regions of less than 100 square kilometers, as such tiny areas are due to projection error and the lack of precision in the underlying mapping of ethnicities.

¹¹Since in our empirical analysis we primarily explore within-country variation, in many specifications we lose observations in countries with either one ethnic group or without variability in partitioning, namely Burundi, Djibouti, Swaziland, Comoros, Madagascar, and Western Sahara.

Africa. Our strategy identifies almost all ethnic groups that Asiwaju (1985) lists as partitioned.



Figure 1a

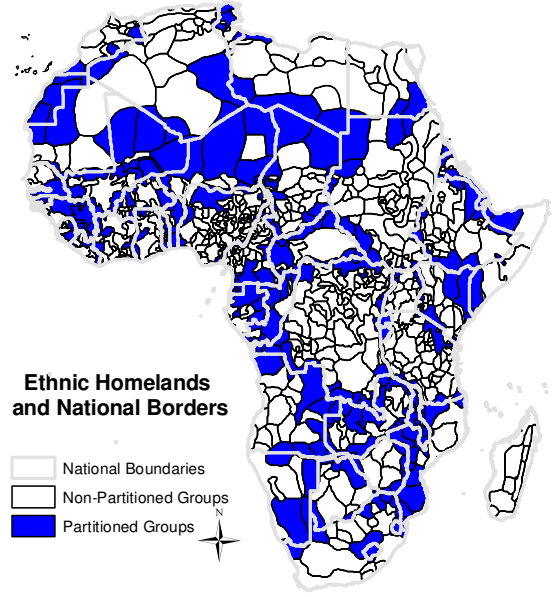


Figure 1b

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.¹²

2.2 Civil Conflict Data

Our data for conflict come from the Armed Conflict Location and Event Dataset (Raleigh, Linke, and Dowd (2012b)) that provides geo-referenced information on the location and some key characteristics of violent events across all African countries since 1997.¹³ A nice feature of this dataset (compared to the standard Correlates of War and PRIO data on civil wars) is the reporting of incidents of violence against the civilian population. This is quite important as Africa is plagued by prolonged civil warfare that the standard data sources of civil war miss (as they focus on battles). Not only violence against the civilian

¹²We 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 why conflict propensity 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 the Sensitivity Analysis Section we report specifications with the continuous index.

¹³In contrast the PRIO dataset just reports a centroid and an approximate radius of where most battles took place (Raleigh, Cunningham, Wilhelmsen, and Gleditsch (2006)). In the previous draft of the paper, we used this data finding similar results (see Supplementary Appendix Table 10).

population, such as child soldiering raids, rapes, abductions, and burning is recurrent, but these incidents are deadly, economically harmful, and devastating for the victims and the local community.

The ACLED database is quite rich as there are 43,271 incidents of political violence over the period 1997 – 2010. 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 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 on-going civil conflict (e.g., recruitment drives and incursions); (6) Riots and protests; (7) Violence against civilians, where armed groups (rebels, militias or government forces) attack unarmed civilians; (8) Non-Violent transfer of control.

The data are based on a quite diverse set of sources.¹⁴ For almost all countries data come from more than ten different sources, while for most war-prone countries we have data from around twenty different sources. This diversity assuages concerns of systematic biases in reporting from government controlled actors and media. The data are mostly based on respectable international sources, such as the BBC (close to 10,000 incidents), Reuters (close to 5,000 incidents), the Associated Press (around 1,500 incidents), and the Agence France Press (around 3,500 incidents). A considerable fraction of the data (around 10%) comes from international newspapers and media outlets from the United Kingdom, Portugal, Canada, the United States, and Australia. ACLED also relies on reports from international and local NGOs, such as *Human Rights Watch* and *Amnesty International*, and the United Nations.¹⁵ Even in cases of data coming from local sources (around 25% of the sample), most incidents come from pan-African news agencies, such as the *All Africa* network and independent –rather than government run/controlled- newspapers.

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 of territorial control, and (iii) violent events against civilians. Battles among armed forces account for 43.2% of all incidents; out of a total of 18,705 battles, 2,324 resulted in a change of territorial control. The dataset records information on 15,844 cases of violence against civilians (36.6% of all incidents).

Examples of battles between armed actors include the numerous fights between the Acholi-based

¹⁴In parallel work Rohner, Thoenig, and Zilibotti (2012), Besley and Reynal-Querol (2012), and Harari and Ferrara (2012) also use the ACLED to study other than ethnic partitioning aspects of civil conflict.

¹⁵Going over the documentation it seems that the data are based on verified information and not simply the reproduction of government statements and state press releases. For example, in Zimbabwe, besides data coming from BBC and Reuters a lot of data come from Zimbabwe Human Rights NGO Forum, a coalition of nineteen human rights NGOs in Zimbabwe that get the data from their representatives on the ground. Likewise, in Somalia a lot of data come from international NGO CARE via its Security and Preparedness Project that “*aims to reduce the risks posed to programme personnel and assets of NGOs operating in Somalia.*” Similarly, in Kenya many incidents are based on reports from the local Peace and Development Network Trust, a local NGO, co-founded by Oxfam that monitors peace.

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 in Rwanda and in Eastern Congo; and the (relatively small scale) battles between Kikuyu rebel groups against Maasai militias. Battles result usually in 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 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 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.

Violent events against civilians include the raids of the Janjaweed militias against civilian population in the Darfur region in Western Sudan; the assaults and tortures 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 militias in the Democratic Republic of Congo. Violent events include also the burning of churches, hostage-taking and child-soldiering raids by rebels in Nigeria and in Sierra Leone. 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 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. Violence is also pervasive in Somalia, Ethiopia, and Zimbabwe. 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. Likewise, most of the conflict in Senegal is concentrated on the Southern region of Casamance, while most conflict in Angola is close to the northern border with Congo and in the Cabinda enclave.

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

distribution of all civil conflict incidents at the ethnic-homeland level.

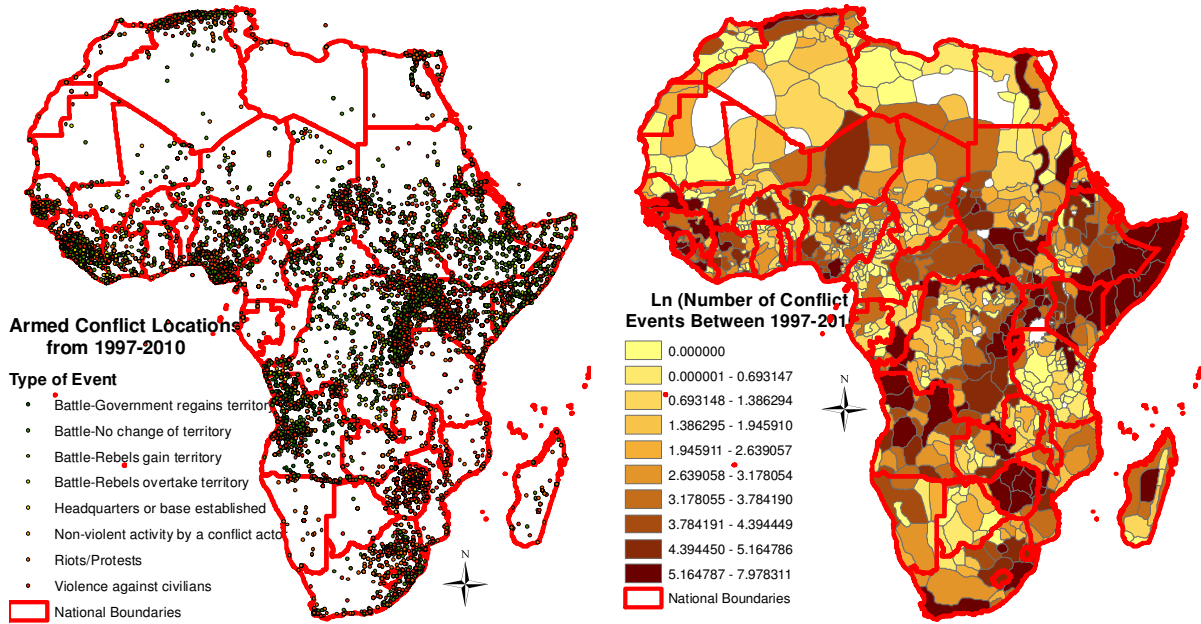


Figure 2a

Figure 2b

We examine separately 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 violence against civilians. The correlation between battles and violence against civilians is high, but far from perfect (0.60). For example, in Zimbabwe we observe many violent events against civilians with very few battles between military actors. Conversely in Ethiopia and Sudan we predominantly observe conflict between the government and rebel groups rather than unilateral violent events against civilians. In the Democratic Republic of Congo and in Uganda the numbers of battles go hand in hand with civilian violence. 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 main components.

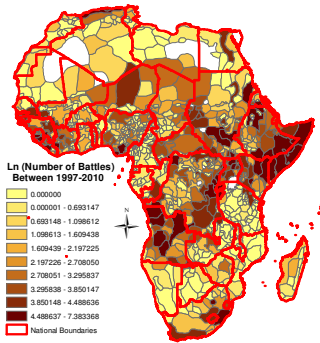


Figure 3a

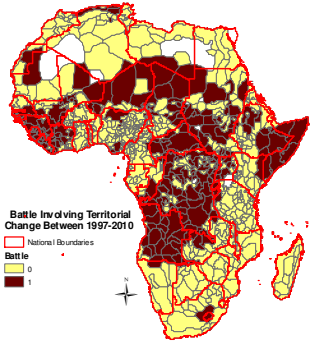


Figure 3b

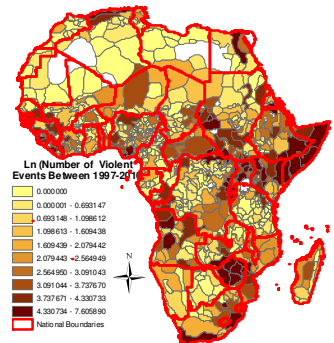


Figure 3c

2.2.2 Summary Statistics

Table 1 reports descriptive/summary statistics for the main outcome variables across the 826 ethnic homelands.¹⁶ 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 from the centroid of a group; 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 partitioned ethnicities 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, we observe similar differences; on average 65% of non-partitioned ethnic homelands experienced a conflict, while 83% of split groups suffered from 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 are recorded. This difference is not statistically significant because there are some extreme cases both across partitioned and non-split ethnic groups (see Appendix Table 2). To account for outliers we exclude ethnic homelands where capitals fall (in (4) and (5)) and homelands where the number of armed conflicts exceeds the 99th percentile (in (6)-(7)). The differences are now statistically significant. There are also large differences between partitioned (median=13) and non-split ethnicities across the median value of conflict events (13 compared to 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).¹⁷

¹⁶In 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*). Supplementary Appendix Table 3 reports the number of conflict incidents for each country.

¹⁷The results are similar if we use a narrower threshold of distance to the national border to identify ethnic homelands close

Battles On average 59% of all ethnic homelands have experienced some 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 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 regions where capitals fall, the difference becomes significant. 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; moreover, these incidents may reflect more accurately irredentist demands and state collapse. 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.

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 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 groups (3 versus 1) independently on whether we examine all ethnic homelands or we limit our attention to those close to the national border.

3 Border Artificiality

As we briefly discussed above, the African historiography provides ample evidence that, in the overwhelming majority of cases, Europeans did not consider ethnic politics and local geography in the design of colonial borders - that were mostly drawn before Europeans settled in Africa. In a few instances, nevertheless, Europeans did try taking into account political geography, as, for example, in Swaziland, Burundi, and in some parts of Uganda. Moreover, in two cases (Cameroon-Nigeria; Ghana-Togo) there were referenda on

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).

the redrawing of borders at independence. Yet what is key for establishing causality is not that all borders are idiosyncratically drawn (though many were); what is needed for inference is that there are no major differences between partitioned and non-split ethnic homelands across factors that may independently affect conflict. Thus in this section we examine in detail whether there are -on average- significant differences between the two sets of ethnicities across observable factors that may shape the propensity and intensity of conflict.

We search for potential correlates of ethnic partitioning estimating models of the following form:

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

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 state. For robustness, we also show results using the continuous measure of partitioning, $FRAC_i$. X_i is a vector of geographic, ecological, natural resource variables and proxies of pre-colonial conflict and development 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* that are available for a subset of African ethnicities. Appendix Table 2 reports 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.

3.1 Geographical, Ecological, and Natural Resource Measures

Table 2 reports probit (maximum-likelihood) marginal effects, where the dependent variable is the benchmark binary partitioning index ($SPLIT$), and LS estimates with the continuous measure of partitioning ($FRAC_i$) as the dependent variable.¹⁸

In Panel *A* we explore the role of geographic, ecological, and natural resources. Specifications (1)-(2) show that ethnic groups spanning large territories in the pre-colonial period were more likely to be partitioned. This is consistent with the narratives describing the idiosyncratic border drawing. The estimates further show that ethnicities residing in areas with larger water bodies were more likely to find themselves split by the political boundaries. This result is again in accord with the historical evidence that Europeans attempted to use natural barriers while delineating spheres of 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 to proxy for ruggedness and the variance of land quality. Examining the association between partitioning and land characteristics is important because cross-country works and regional studies show that conflict is more prevalent in mountainous or malarious terrains (e.g., Fearon and Laitin (2003), Buhaug and Rod (2006), Cervellati, Sunde, and Valmori (2011)). All four geographic features enter with insignificant estimates.¹⁹

¹⁸The results are similar if we estimate linear probability or logit specifications or when we estimate Tobit models that account for truncation (at zero) of the continuous partitioning index.

¹⁹In some specifications mean land suitability for agriculture enters with a (weakly) significant estimate. We further explored

In columns (5)-(6) we examine the role of ecological conditions, augmenting the empirical model with a malaria stability index and distance to the coast. Since Europeans settled almost exclusively in coastal areas and regions where malaria was less pervasive, these models shed also light on whether contact with colonizers affected partitioning. Both indicators enter with small and statistically insignificant coefficients.

In columns (7)-(8) we include indicators identifying 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 that are correlated with conflict both across and within countries (see Ross (2012) for a thorough review). There are no systematic differences between the two groups of ethnic homelands.

3.2 Pre-colonial Conflict²⁰

While at the time of border design in the late 19th century, Europeans had limited or absolutely no knowledge of local political geography, it is necessary to examine the association between ethnic partitioning and pre-colonial conflict, as recent cross-country works (Fearon and Laitin (2012)) and cross-regional studies in Africa show a significant legacy of conflict from the pre-colonial period to contemporary times (Besley and Reynal-Querol (2012), Nunn and Wantchekon (2011)). In Table 2 -Panel *B* we thus examine the association between ethnic partitioning and two proxies of pre-colonial conflict, one based on wars between large pre-colonial African kingdoms and one reflecting conflict during the slave trade period.

Using historical data from Brecke (1999) and Cioffi-Revilla (1996), Besley and Reynal-Querol (2012) show that contemporary conflict is higher in regions that suffered from pre-colonial wars (such as the Songai-Gourma conflict in contemporary Mali in the end of the 15th century or the conflict between the powerful Banyoro and Buganda kingdoms around 1600 in contemporary Uganda). Specifications (1)-(2) show the lack of a systematic association between ethnic partitioning and pre-colonial violence, as reflected by an indicator that takes the value one for ethnic homelands that experienced conflict over the period 1400 – 1700. Similarly, specifications (3)-(4) show that ethnic partitioning is not related to distance to the centroid of the closest pre-colonial conflict (the results are similar with log distance). In line with the historical accords, these simple specifications reveal that Europeans did not take into account pre-existing animosity between African ethnicities when designing colonial borders. More importantly for our work, these results show that ethnic partitioning captures a distinct source of contemporary conflict to that emphasized by Besley and Reynal-Querol (2012). And while there is some precision error on the exact

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. 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.

²⁰We thank Francesco Caselli for proposing exploring the link between the legacy of pre-colonial conflict and ethnic partitioning.

locations of pre-colonial conflict, the data have signal, as Besley and Reynal-Querol (2012) do find a robust positive association between pre-colonial conflict and civil wars after African independence.

As detailed in Nunn (2008) and previous historical studies, Africa experienced conflict during the slave trades, as the most common method of enslavement was “*through raids and kidnapping by members of one ethnicity of another or even between members of the same ethnicity*” (Nunn and Puga (2012)). In line with this Nunn and Wantchekon (2011) document a strong negative within-country correlation between historical enslavement and proxies of social capital, arguing that the long-run effect of the slave trades works via spurring animosity between African ethnicities. Similarly Djankov and Reynal-Querol (2010) present cross-country evidence of a significant positive association between enslavement and civil war post-independence. In columns (5)-(6) we regress ethnic partitioning on an indicator that takes on the value of one for ethnicities that were affected by the slave trades and zero otherwise (“extensive” margin) while in (7)-(8) we follow Nunn (2008) and use the log of 1 plus the number of slaves normalized by the area of each homeland. In all specifications the coefficient on slave trades is quantitatively small and statistically indistinguishable from zero. This further assuages concerns that the ethnic partitioning index captures pre-colonial levels of violence.

3.3 Pre-colonial Political and Economic Development

In Panel *C* we explore whether ethnic partitioning is related to pre-colonial political and economic development. Both anecdotal evidence from the African historiography (Herbst (2000)) and the growing literature on state capacity (Besley and Persson (2011)), suggests that pre-colonial conflict involved large kingdoms with a high degree of political centralization (see also Acemoglu and Robinson (2012)). Thus we associate ethnic partitioning with the homeland falling within the boundaries or being close to a large pre-colonial kingdom, using data from Besley and Reynal-Querol (2012) and O’Brien (1999). There is no systematic association between ethnic partitioning and the homeland being part of a large kingdom (columns (1)-(2)). Likewise, models (3)-(4) show that there is no relationship between ethnic partitioning and distance to the centroid of the closest pre-colonial kingdom.

In columns (6)-(7) 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 and zero otherwise (using data from Chandler (1987)). There is no evidence that ethnicities with historical urban centers were differentially treated during the early stage of colonization when borders were drawn in European capitals.

In columns (7)-(8) we regress the partitioning measures on the average distance of each ethnic group to the main European exploration routes. This helps shed light on whether Europeans used the (limited) knowledge they had on local political geography. The coefficient is small and statistically indistinguishable from zero.

3.4 Ethnic-Specific Pre-colonial Traits

In Table 3 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. While due to the drop in the sample size, we lose efficiency, examining the role of various ethnic-specific pre-colonial features on partitioning sheds light on the (absence of) considerations of Europeans when they drew colonial borders in the late 19th century.

In columns (1)-(2) of Panel A - Table 3 we use an ordered variable, ranging from 0 to 7 with higher numbers indicating more complex and more densely populated local communities. There is no association between partitioning and pre-colonial settlement pattern. Building on the argument of the African historiography and the results of the unified growth theory (Galor (2011)), that pre-colonial economic development was higher in areas with intensive use of agriculture (Fenske (2012b) provides empirical support), in (3)-(4) we proxy development at the time of colonization with a 0 – 9 index measuring the importance of agriculture, failing again to detect a systematic association. In columns (5)-(8) we further investigate whether Europeans took into account the degree of political centralization of the African ethnicities while designing the borders. Following Gennaioli and Rainer (2006, 2007), in (5) and (6) 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 Tiv); and becomes 1 when the ethnicity is part of either a "*large paramount chiefdom*" or a "*large state*" (e.g., Ganda and Zulu). Since institutional and economic development in pre-industrial societies goes in tandem with class stratification (see Diamond (1997)), in (7) and (8) we examine the association of ethnic partitioning and a stratification index that ranges from zero, for societies without any class distinctions, to four, for groups with significant class and wealth distinctions. There are no systematic differences between partitioned and non-split ethnicities along these proxies of political centralization.

In Panel B of 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 inheritance, the role of clans, and other traits. There are no significant differences between split and non-partitioned ethnicities. And while these variables do contain noise (and thus attenuation is a valid concern), measurement error cannot fully explain the lack of significance, as many works document robust associations between several societal traits in Murdock's map and contemporary outcomes (e.g., Fenske (2012c), Gennaioli and Rainer (2007), Fenske (2012a), Alesina, Giuliano, and Nunn (2012)).

3.5 Country-Fixed-Effects Estimates

Since we mostly explore within-country variation, one would also like to know whether post-independence there are systematic differences between partitioned ethnicities and non-split groups in the same country. In Table 4 we report 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 to the corresponding country. The evidence suggests that the two sets of ethnic groups are comparable. Interestingly, the positive correlation between surface and water area and partitioning turns now insignificant. This is because after partitioning both the surface area and the area under water of split groups within a country are similar to those of non-partitioned ethnicities.

3.6 Summary

The results in Tables 2 – 4 are broadly consistent with the historical narrative on the arbitrary design of African borders. Our results, however, do not imply that all African borders were randomly designed; this is clearly not the case. What our large-scale econometric evidence shows is that on average there are no systematic differences between partitioned ethnic homelands and non-split ethnic regions across observable characteristics that may affect conflict independently.²¹

4 Partitioning and Civil Conflict

4.1 Econometric Specification

We estimate the long-run effect of ethnic partitioning on contemporary civil conflict running variants of the following empirical specification:

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

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 group i 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 Lobi’s homeland in Burkina Faso is assigned to Burkina Faso. The coefficient γ on $SPLIT$ captures the direct (local) effect of ethnic partitioning on civil conflict. Given the lack of systemic correlation between the partitioning index and various historical, ecological, and geographical variables that previous works show that correlate with conflict the coefficient captures the local average treatment effect of ethnic partitioning. Vector $X'_{i,c}$ includes geographical, ecological, and other controls at the country-ethnicity level.

As the dependent variable (all incidents, number of battles, violent events against civilians) is a count, we estimate negative binomial models with maximum likelihood (Wooldridge (2002)).²² The negative

²¹It is worth noting that the explanatory power of the models in Tables 2–4 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 230 partitions with the benchmark index ($SPLIT$).

²²Due to overdispersion in the dependent variable, specification tests reject the Poisson, favoring the negative binomial model.

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. For robustness we also report log-linear LS specifications taking the log of one plus the respective civil conflict measure as the dependent variable.²³ 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 adjust/cluster standard errors at the country level and at the ethnic-family level using method of Cameron, Gelbach, and Miller (2011) that accounts for arbitrary residual correlation within each country and within each ethnic family and spatial correlation.²⁴

4.2 Cross-Sectional Estimates

We start our analysis estimating the relationship between partitioning and civil conflict across the 826 ethnic homelands, without conditioning on country fixed effects. Table 5 reports the results. While these estimates have many shortcomings, it is useful examining the cross-sectional patterns before moving to the country fixed effect results. In column (1) we simply control for log population using the first post-independence census (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 Tables 2 – 3.²⁵ 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 (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 homelands where capitals fall. Overall, distance to the sea enters with a positive and significant estimate suggesting that there is less conflict in areas closer to the coast. 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 always 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.

²³The non-linear estimator is more appealing than the log-linear model, because it preserves the higher moments of the distribution (see Silva and Tenreyro (2006) and Silva, Tenreyro, and Windmeijer (2010)). Standardizing the dependent variable with land area or population yields similar results.

²⁴Cameron, 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. In the Sensitivity Analysis Section we also report spatial lag specifications, finding similar results. We also estimated standard errors using the method of Conley (1999) to account for spatial dependence of an unknown form, finding similar (and if anything less conservative) errors.

²⁵Since civil conflict may affect population dynamics, following the recommendation of Angrist and Pischke (2008) we control for (log) population using the earliest date. As we show in Appendix Table 3 the results are very similar when we use population estimates in 2000.

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). 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. 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 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 groups.²⁶ 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 (which however by itself could reflect the impact of ethnic split).

4.3 Within-Country Analysis

Baseline Estimates In Table 6 we report our baseline 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 whose centroid is 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 (2) implies that on average civil conflict intensity is higher in homelands of partitioned groups by approximately 60% ($\exp(0.47) - 1 = 0.60$). 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 (around 0.85) and highly significant. Although the coefficient in areas close to the national border

²⁶When we restrict estimation to homelands where the centroid falls within the median value of distance to the national border, we do not use information from 17 (out of the 230) split groups. This is because in these cases the centroid of the partitioned ethnic area is more than 103 kilometers from the border. The results are intact if we (manually) include these groups in the estimation close to the national border.

is somewhat larger compared to the estimate in the full sample, a Hausman-Chow type test that compares the estimates reveals that these differences are not statistically significant.²⁷

While the ACLED does not report an official number of casualties, going over the description of each event we can classify deadly incidents. As it is interesting to investigate whether ethnic partitioning is linked to casualties, we re-run our baseline specification focusing on deadly incidents. For brevity, we report these results in Appendix Table 4. Not only partitioning is associated with a higher incidence of conflict, but those violent events (mostly battles and violence against civilians) were also deadly.

Ethnic Partitioning and Type of Civil Conflict In an effort to shed some light on the mechanism at work, in Table 7 we examine the effect of ethnic partitioning on the different types of conflict.²⁸ We start our analysis focusing on battles between armed groups. The coefficients on ethnic partitioning in (1)-(2) imply that fighting between government forces, militias, and rebel groups is more pervasive in the historical homelands of partitioned groups; the estimate suggests 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.

A useful feature of the ACLED is the reporting of violence against the civilian population, a socially and economically devastating aspect of conflict that the commonly employed civil war data miss. In columns (3), (4), (9), and (10) we examine the role of ethnic partitioning on abductions, child soldiering raids, village burning, looting, and other incidents of violence against civilians. Ethnic partitioning is systematically linked with civilian violence. The estimate implies that there are 65% ($\exp(0.50) - 1 = 0.65$) more violent incidents against civilians in the homelands of partitioned ethnicities. Restricting estimation to ethnic regions close to the national border yields similar -and if anything somewhat higher estimates (although the difference in magnitudes is not statistically significant).

Another interesting dimension of the ACLED data is the reporting of incidents, mostly battles, between government forces, rebel groups, and militias resulting in territorial changes. Examining instances of territorial changes is useful as they reflect -almost by definition- the government's lack of monopoly of violence. 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 (6) and (12) implies that there is a 4% - 5% higher likelihood that a battle resulting in a change of territorial control occurs in the homeland of a partitioned ethnicity. This effect is far from being small, as in the country-ethnic homeland sample, the overall likelihood that a territorial change takes place is around 20%. The results in Table 7 are in accord with the arguments put forward by the African historiography suggesting that areas of partitioned ethnic groups are tightly contested between the government, ethnic

²⁷The main reason for the lack of systematic differences in the two estimates is that there are some outliers in conflict incidence (both far and close to the national border and both in the homelands of partitioned and non-split groups).

²⁸Appendix Table 5 reports analogous cross-sectional specifications at the ethnic homeland level.

militias, and rebel groups.²⁹

LS Specifications In Table 8, columns (1)-(3) and (6)-(8) we report LS specifications using the natural log of one plus the total number of civil conflict incidents as the dependent variable. The estimate on the partitioning index is positive and highly significant both in the full sample and when we restrict estimation to areas close to the border.³⁰ The estimates in the full sample imply that conflict is approximately 20% higher in the homelands of partitioned groups, as compared to groups that were not split by 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, we do not exploit the richness of the data, we further account for the non-linear nature of the dependent variable. Moreover, these specifications shed light on the margin that ethnic partitioning operates on civil 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 cut-off the median number of conflicts; as the median is 2, in these specifications we just assign regions that experienced just one incident in the group of ethnicities with no conflict. The coefficient on the ethnic partitioning index in (5) and (10) retains its economic and statistical significance, implying that compared to non-split groups, the likelihood that a partitioned ethnic homeland will experience more than one conflict incident is 10% higher.

Counterfactual We also performed an out-of-sample exercise to get a rough estimate of the overall impact of partitioning in Africa, under the heroic assumptions that borders do not split any group and there are no spillovers. We first estimate the baseline model (without and with country fixed effects) in the group of non-split groups (680 country-ethnicity observations). Second, we obtained out-of-sample predictions of conflict for the group of partitioned ethnicities (502 country-ethnicity observations). Third, we compared the predictions with the actual number of incidents. The forecast model -that does not take into account the role of partitioning- predicts approximately 4,300 conflict incidents in the simple specification and 6,265 incidents when we include country fixed effects. Yet in the country-ethnic homelands of split groups we observe 11,731 incidents, implying that at least 5,500 (if not 7,500) incidents may be attributed to ethnic partitioning. Since overall we observe 43,271 incidents, this simple (and based on heroic assumptions) out-of-sample calculation suggests that conflict would be 12.5% to 17.5% lower in the absence of ethnic partitioning. When we perform the out-of-sample exercise with the linear probability specification, the model predicts that 55% of partitioned ethnic groups would have experienced at least one conflict. In practice, however, 66.5% of country-ethnicity homelands experienced conflict. Thus ethnic partitioning

²⁹In line with the arguments put forward by the African historiography that stress the role of ethnic partitioning in spurring civil conflict, we do not detect a significant association between partitioning and riots and protests (Appendix Table 6).

³⁰We also examined whether the coefficient estimates on the ethnic partitioning index in the full sample differs systematically from the coefficient in the sample of ethnic homelands close to the border. The χ^2 (with one degree of freedom) is in the range of 1.5 – 2.0 indicating that the estimates are statistically indistinguishable from each other.

explains approximately 10% of overall conflict incidence.

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 results. Let us start from Tanzania, a country with little conflict; in the 69 ethnic regions of Tanzania there have been 175 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. While in the Rundi homeland only 1.3% of Tanzania’s total population lives (in both 1960 and in 2000), the share of conflict is 27.5%. Conflict also appears in the nearby homeland of the (ethnically similar) Ha, where militias based in nearby Rwanda and Burundi raid against the civilian population. The share of total conflict occurring in the Ha homeland is 27.5% (19 incidents) while the share of Tanzania’s population living in this region is 4.5%. Interestingly, there is zero conflict in the non-split homelands of the Bende and the Fipa, 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 serves as the natural border of the two countries. Focusing now on the northern border of Tanzania with Kenya, there is recurring conflict (in total 10 incidents, 15%) in the homeland of the partitioned Maasai, while the fraction of Tanzania’s population living in the land of Maasai is 3%. For example, ACLED documents a fight resulting in the death of 30 farmers by Maasai militia on December 8th 2000. In contrast, there is no conflict in the Eastern part of the Tanzania-Kenya border where the non-split Pare reside.

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 observe 946 incidents (more than 20% of all conflict), while the share of population residing in the homelands of the Hunde, the Konjo, and the Rwanda is around 6%; and in the two adjacent non-split, but ethnically similar Rwanda groups of the Hunde and Toro we have 409 and 27 events, respectively. So while the share of Zaire’s population living in the Hunde homeland is around 2.2%, we observe close to 10% of all conflict. 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.³¹ In contrast there is no conflict at the (non-split) Holoholo ethnic homeland on the Western (Zairian) bank of the Tanganyika lake; yet in the adjacent to the Holoholo homeland in the South, in the Tabwa region we observed 91 conflict events (2%), while the share of Zaire’s population is 0.8%. This is not surprising since, although

³¹For example, in early 2007 in just one event, FDLR groups raided two villages killing 17 civilians and wounding 19. In the next day FDLR militias looted 18 houses in a nearby town. Also in a 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.

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.

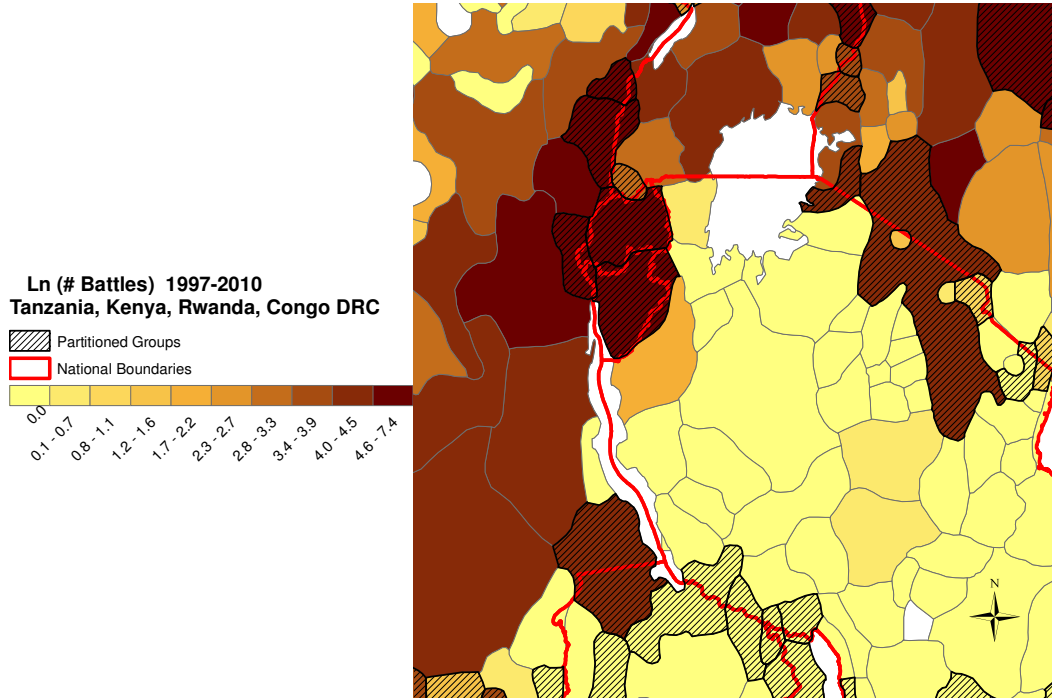


Figure 4

5 Sensitivity Analysis and Heterogeneous Effects

5.1 Main Robustness Checks

We perturbed the empirical model in various ways to explore the robustness of our findings.

Unobservables First, to account for unobservable characteristics that vary smoothly in space in Table 9A, columns (1)-(4) we augment the specification with a cubic polynomial in latitude and longitude of the centroid of an ethnic group in each country (see Dell (2010)).³² 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-family factors. This robustness check is particularly interesting, as recent works show that ethnic-specific factors, related to history, culture, the type of political and economic organization, and genetics affect crucially development (see Nunn (2012) and Spolaore and Wacziarg (forthcoming)). 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,

³²Letting x denote latitude and y denote longitude the polynomial reads: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

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 Second, we estimated models dropping iteratively ethnic homelands from each of the five African regions to investigate whether the results are driven by a particular part of the continent. Table 9B reports the results. 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. Moreover, since most large pre-colonial African states were present in Ethiopia and in other parts of Eastern Africa (Besley and Reynal-Querol (2012)), we further account for the role of pre-colonial conflict and political development. The results show that the strong positive effect of ethnic partitioning on civil conflict is not driven by a particular region.

Spatial Effects Third, we estimated spatial models that explicitly account for spill-overs and for factors in nearby ethnic regions. Table 10 reports spatial lag specifications (estimated by maximum-likelihood). Since spatial lag models are somewhat sensitive to the details of the specification, we report estimates using two different types of (inverse) weighting matrixes (linear in distance and quadratic in distance) and three different types of models. Accounting for spatial spill-overs in conflict as well as the effect of the independent variables in the homelands of neighboring groups (in the Durbin and Generalized Spatial models) has little effect on the coefficient of ethnic partitioning that retains its economic and statistical significance.³³

Data Sources Fourth, to minimize concerns related to the underlying data source, we classified each of the approximately 43,000 events based on the original source (international newspaper, global news agency, NGOs, United Nations, academic bulletins, local sources) and then re-estimated the empirical

³³While in the presence of externalities our estimates will be lower bounds of the effect of partitioning, in Appendix Table 7 we further explore the impact of spatial spill-overs running specifications linking conflict with ethnic partitioning conditioning on the total number of conflicts in the same country (netting out conflict in each ethnic homeland) and conditioning on total conflict of each ethnolinguistic family (netting out conflict of each ethnic homeland). The cross-sectional estimates reveal that conflict is higher when an ethnic homeland falls in a conflict-prone country and when there is a lot of conflict in groups from the same family. These results are quite interesting as they show that conflict and violence against civilians spread within country and along ethnic family lines. Yet, this has little impact on the ethnic partitioning index that retains its economic and statistical significance.

specifications using events reported solely by international sources, such as global news networks (e.g., Associated Press, Reuters, BBC), non-African newspapers (e.g., the Guardian, New York Times, 24 Horas), international NGOs (e.g., the Human Rights Watch, Amnesty International), and the United Nations. Table 11-Panel *A* reports the results. In all specifications the coefficient on the ethnic partitioning indicator is positive and statistically significant at the 99% confidence level. We also estimated models using only those events reported by NGOs and the various reports of the United Nations teams. Panel *B* reports the results. While the number of events drops dramatically (as we use just 15% of the data) and we lose countries which are not covered either by NGO's or UN reports, this minimizes concerns of any media-specific bias. In spite of the efficiency loss, we obtain similar results.

Accounting for Pre-colonial Conflict Fifth, we estimated empirical models linking contemporary conflict on ethnic partitioning, controlling for the historical legacy of violence. Table 12 reports the results. Following Besley and Reynal-Querol (2012), in Panel *A* we augment the empirical specification with the distance of the centroid of each ethnic homeland to the center of the closest pre-colonial conflict. Given the lack of association between pre-colonial conflict and ethnic partitioning, it should come at no surprise that the estimate on the partitioning index retains its economic and statistical significance across all permutations. Since many wars before colonization took place in the territories of large centralized kingdoms, in Panel *B* we include in the specification a dummy variable that takes on the value of one if an ethnic homeland was part of a large pre-colonial kingdom. The coefficient on the ethnic partitioning index is positive and highly significant.³⁴ In Panel *C*, we control for the impact of conflict during the slave trades augmenting the specification with the estimates of Nunn (2008) and Nunn and Wantchekon (2011) on enslavement at the ethnicity level. The coefficient on the partitioning index remains intact, further reassuring that it does not capture conflict during the pre-colonial period.

Alternative Measures of Partitioning Sixth, we repeated estimation using alternative measures of ethnic partitioning. Table 13 reports the results. Columns (1)-(3) and (7)-(9) present results associating civil conflict with the continuous index of partitioning (*FRAC*). 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.

³⁴To further minimize concerns that ethnic partitioning captures the legacy of pre-colonial conflict, in Appendix Table 8 we report estimates excluding all ethnic homelands that experienced significant pre-colonial conflict. The results are intact. Likewise, in Appendix Table 9 we drop from the estimation all ethnic areas that were part of large pre-colonial states, kingdoms, and empires. In spite of the significant drop in the number of observations, the coefficient on the ethnic partitioning index retains its economic and statistical significance.

Alternative Mapping of Ethnic Homelands and Ethnic Partitioning Seventh, we repeated the analysis using *Ethnologue's* database that reports the spatial distribution of linguistic groups in the early/mid 1990s. *Ethnologue* explicitly maps linguistic homelands within each country making the identification of partitioned 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, local violence, colonization, as well as numerous country-level characteristics that affect conflict. *Ethnologue* includes information on 2405 linguistic groups in Africa out of which 821 are mapped in more than one countries. 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 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 groups located in a single country have experienced on average 10 conflict events. All differences are statistically significant at conventional levels. Table 14 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 implies that conflict intensity is approximately 30% ($\exp(0.267) - 1 = 0.30$) higher in the contemporary homelands of partitioned groups.

5.2 Heterogeneous Effects

Our results suggest that ethnic partitioning has had a strong positive effect on civil conflict, on average. Yet it may be the case that ethnic partitioning is especially harmful in certain environments. Based on previous works on the origins of African development and the historical narratives of the influence of the Scramble for Africa on conflict, in Table 15 we explore potential heterogeneous effects of partitioning. Searching for heterogeneous effects is also helpful as it sheds light on how and under which circumstances ethnic partitioning spurs conflict.

Ethnic Composition, Size, and Landlocked Countries The Scramble for Africa has led to the creation of some very large and heterogeneous states. Moreover, the border design resulted in Africa having the most landlocked countries across all continents. In Table 15-Panel *A* we examine whether these aspects of the Scramble for Africa interact with ethnic partitioning. In columns (1)-(3) we investigate whether the impact of ethnic partitioning differs with respect to the degree of ethnic, linguistic, or religious fractionalization; we do so by augmenting the baseline specification with an interaction term between the *SPLIT* index and a dummy identifying countries scoring above the median on the Alesina, Devleeschauwer,

Easterly, Kurlat, and Wacziarg (2003) fractionalization measures. The estimates show that the impact of ethnic partitioning is quite homogeneous and does not depend on the degree of fractionalization. The results are similar when we group countries based on the degree of cultural fractionalization (using data from Fearon (2003)), ethnic polarization (using the index of Montalvo and Reynal-Querol (2005)), or ethnic segregation (see Alesina and Zhuravskaya (2011) and Matuszeski and Schneider (2006)).

In column (7) we augment the specification with an interaction between the partitioning index (*SPLIT*) and a dummy variable that identifies large in terms of size countries (using the median value of land area as a cutoff). The coefficient on the interaction term is positive, implying that the impact of partitioning is more pronounced in large countries, such as Zaire, Sudan, and Angola. Yet the estimate is not statistically significant at standard confidence levels. We obtain similar results when we use other thresholds of land area to identify large countries. In column (8) we add an interaction between the partitioning index and an indicator for landlocked countries (e.g., Chad, Burkina Faso, Rwanda, Mali, the Central African Republic). The results suggest that ethnic partitioning is particularly deleterious for groups in landlocked countries.

The Role of National Institutions In Table 15-Panel *B* we examine whether the negative consequences of ethnic participating are mitigated in relatively more (institutionally and economically) advanced countries. In column (1) we use a composite rule of law index that reflects the quality of property rights and legal institutions (using data from World Bank’s Governance Matters Project). The interaction term enters with a negative coefficient, hinting that sound national institutions may attenuate the impact of partitioning; yet the estimate is not significant at standard confidence levels (the results are similar with other proxies of legal quality). In column (2) we allow the effect of ethnic partitioning to differ for democratic countries, using a binary democracy variable from Przeworski, Alvarez, Cheibub, and Limongi (2000). The interaction enters with a significantly negative coefficient suggesting that democratic regimes are better at dealing with the adverse consequences of ethnic partitioning. The results are similar when we use alternative sources to identify democratic countries (Freedom House, Polity Project). In column (3) we further explore the role of democracy interacting the partitioning index with a measure of democratic capital (from Persson and Tabellini (2009)) that reflects counties’ experience with representative rule (rather than the degree of democracy in a given year). The interaction is negative and significant, further showing that democratic rule tends to reduce the negative impact of ethnic partitioning.

While these results are supportive of the idea that national institutions and democracy, in particular, may be beneficial by lowering the propensity of conflict emerging from ethnic partitioning, they cannot be casually interpreted because conflict may be both a cause and a consequence of institutional quality. Moreover, both civil war and institutional quality are a function of economic development, as well as other factors (related to geography, culture, colonization, etc.). In line with this, specification (4) shows that the positive effect of partitioning on conflict is lower in more economically advanced countries. In columns (5)-

(6) we jointly include in the specification the interaction between ethnic partitioning and democracy and the interaction of ethnic partitioning with economic development. Both interactions enter with significantly negative coefficients, implying that the attenuating effect of democracy on conflict works on top of the role of economic development. In columns (7)-(8), following Acemoglu, Johnson, Robinson, and Yared (2008), we use executive constraints at independence (in the 1960s) and income per capita in 1961 to identify institutionally and economically advanced countries at the time of decolonization. The results suggest that partitioned groups found in countries that at independence were relatively more institutionally advanced experience today less conflict, implying that inclusive institutions may be helpful in lowering conflict caused by ethnic partitioning. If civil conflict and institutional development are inter-linked, then these results suggest ethnic partitioning may have contributed to the emergence and persistence of a conflict-driven poverty trap, where partitioning has particularly strong effects in under-developed countries, which in turn impedes the consolidation of inclusive institutions.

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 explore the 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. We also show that on average there are no significant differences between partitioned and non-split groups on proxies of pre-colonial conflict, early development, and numerous other dimensions that affect the propensity and intensity of civil conflict (such as natural resources and ecology). Our results offer support to the African historiography on the accidental drawing of colonial and consequently national borders in the overwhelming majority of cases.

Second, we examine 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 long-run impact of ethnic partitioning, 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. This is key as due to data limitations most previous works had not examined this quite damaging and economically important aspect of conflict. Our regional focus enables us to solely examine within-country variation and as such account for all country-level features that may affect warfare. Moreover, moving the analysis from the country to the ethnic homeland level is appropriate as in Africa ethnicity is particularly salient and

most conflicts and violent events have an ethnic dimension.

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 more. We also show that the impact of ethnic partitioning is especially strong in landlocked countries whereas the negative impact of partitioning seems to be mitigated in countries with inclusive, democratic institutions. The uncovered differences in the probability and intensity of civil conflict, battles and violence against civilians between partitioned and non-split 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 may examine the impact of ethnic partitioning on other aspects of economic and institutional development. Moreover, our study calls for future work to uncover the exact mechanisms via which the Scramble for Africa has affected long-run economic performance. Finally, since border artificiality and ethnic partitioning in particular are not an exclusive African phenomenon subsequent works could also study their effect in other regions, such as the Middle East.

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 an 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 against civilians, where armed groups (rebels, militias or government forces) attack unarmed civilians; (8) Non-Violent transfer of control. In the cross-sectional specifications we aggregate the data at the ethnic homeland level and in the within-country specifications we aggregate the data 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.*

Pre-colonial Conflict: Using data on the centroid (latitude-longitude) of 91 major conflict incidents in Africa in the pre-colonial period (between 1400 – 1700) we define a dummy variable that takes on the value of one for ethnic homelands that experienced such a conflict in their territory, as defined by Murdock’s map. Following Besley and Reynal-Querol we also use the distance of each ethnic homeland to the centroid of the closest pre-colonial conflict. *Source: Besley and Reynal-Querol (2012); original source: Brecke (1999) and Cioffi-Revilla (1996).*

Pre-colonial Kingdoms and Empires: We define a dummy variable that takes on the value of one for ethnic homelands that were part of a large pre-colonial kingdom and empire. We also use the distance of each ethnic homeland to the centroid of the closest pre-colonial kingdom/empire. *Source: Besley and Reynal-Querol (2012); original source: O’Brien (1999).*

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 and Wantchekon (2011).*

Population at Independence: Log of population as recorded in the first post-independence census (in the 1960s for most countries). *Source: UNESCO (1987).*

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, Terrain-Base, release 1.0 (CD-ROM), Boulder, Colorado.*

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 (2012); Original Source: Atlas of the Biosphere.*

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.

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.*

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.*

Petroleum: Indicator variable that takes on the value of one if an on-shore oil field and gas deposit is in the historical homeland of an ethnic group and zero otherwise. *Source: The Petroleum Dataset v.1.1*

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. PRIO.*

Capital City 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, constructed using ArcGIS Software.

Longitude: Longitude of the centroid of each ethnic group, 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 and a dummy variable that takes on the value of one for ethnic groups that were affected by the slave trades (and zero otherwise). *Source: Nunn (2008) and Nunn and Wantchekon (2011).*

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 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 v67.

Dependence on 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 v5.

Gathering: Binary index that indicates the reliance of the economy on "the collection of wild plants 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 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 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 v3.

Milking: Binary index that equals zero when "domestic animals are milked more often than sporadically" and zero when "little or no milking". Source: Murdock (1967); variable code v41.

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 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 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 reputes 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 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 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 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 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 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 v74.*

Land Area: Log surface area. *Source: Nunn and Puga (2012).*

Income level: Log of per capita GDP at PPP (Chain Index) in 2000 and in 1961. *Source: Penn World Tables, Edition 7. Heston, Summers, and Aten (2011).*

Rule of Law: The rule of law index is "capturing perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence." The standardized index, which corresponds in 2000, ranges from -2.5 to $+2.5$ with higher values indicating better functioning institutions. *Source: World Bank Governance Matters Indicators Database (Kaufmann, Kraay, and Mastruzzi (2005)).*

Ethnic/Linguistic/Religious Fractionalization: Index of ethnic/linguistic/religious heterogeneity, constructed as one minus the Herfindahl index of the share of the largest groups. It reflects the probability that two randomly selected individuals follow different ethnolinguistic/religious groups. *Source: Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003).*

Ethnolinguistic Polarization: Index of ethnolinguistic polarization that achieves a maximum score when a country is occupied by two groups of the same population. *Source: Montalvo and Reynal-*

Querol (2005).

Ethnic Segregation: Index ranging from zero to one capturing ethnic/linguistic/religious segregation (clustering) within countries. If each region is comprised of a separate group, then the index is equal to 1, and this is the case of full segregation. If every region has the same fraction of each group as the country as a whole, the index is equal to 0, this is the case of no segregation. The index is increasing in the square deviation of regional-level fractions of groups relative to the national average. The index gives higher weight to the deviation of group composition from the national average in bigger regions than in smaller regions. *Source: Alesina and Zhuravskaya (2011).*

Cultural Heterogeneity: Index of ethnolinguistic fractionalization that accounts for the degree of similarity between linguistic groups using the *Ethnologue* linguistic tree. *Source: Fearon (2003).*

Democracy: Binary index of political regime (autocracy and democracy), based on the conduct of contested elections in 2000. *Source: Przeworski, Alvarez, Cheibub, and Limongi (2000); updated series available: <http://politics.as.nyu.edu/object/przeworskilinks.html>.*

Democratic Capital: Index reflecting the stock of democratic capital in 2000. Democratic capital is build in years of democracy and depreciates geometrically in years of autocracy. We use the index that is constructed assuming a depreciation rate of 0.99. *Source: Persson and Tabellini (2009); Source: Polity IV Project (Marshall, Jagers, and Gur (2011)).*

Executive Constraints at Independence: Standardized (ranging from 0-1) index that reflects the extent of institutionalized constraints on the decision-making powers of chief executives, whether individuals or collectivities, in the initial ten years since independence. The construction of the index follows Acemoglu, Johnson, Robinson, and Yared (2008). *Source: Polity IV Project (Marshall, Jagers, and Gur (2011)).*

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.
- ACEMOGLU, D., S. JOHNSON, J. A. ROBINSON, AND P. YARED (2008): “Income and Democracy,” *American Economic Review*, 98(3), 808–842.
- ACEMOGLU, D., AND J. A. ROBINSON (2012): *Why Nations Fail? The Origins of Power, Prosperity, and Poverty*. Crown Publishers, New York, NY.
- ALESINA, A., A. DEVLEESCHAUWER, W. EASTERLY, S. KURLAT, AND R. WACZIARG (2003): “Fractionalization,” *Journal of Economic Growth*, 8(2), 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. ZHURAVSKAYA (2011): “Segregation and the Quality of Government in a Cross-Section of Countries,” *American Economic Review*, 101(5), 1872–1911.
- ALESINA, A. F., P. GIULIANO, AND N. NUNN (2012): “On the Origins of Gender Roles: Women and the Plough,” unpublished manuscript, Harvard University.
- ANGRIST, J., AND J.-S. PISCHKE (2008): *Mostly Harmless Econometrics*. Princeton University Press, Princeton, NJ.
- ASIWAJU, A. (1985): “The Conceptual Framework,” in *Partitioned Africans*, pp. 1–18. St. Martin Press, New York.
- BESLEY, T., AND T. PERSSON (2011): *Pillars of Prosperity. The Political Economics of Development Clusters*. Princeton University Press, Princeton, NJ.
- 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.

- BOSKER, M., AND J. DE REE (2010): “Ethnicity and the Spread of Civil War,” CEPR Discussion Paper No. 8055.
- BRANCATI, D. M. (2006): “Decentralization: Fueling the Fire or Dampening the Flames of Ethnic Conflict and Secessionism?,” *International Organization*, 60(2), 651–685.
- BRECKE, P. (1999): “Violent Conflicts 1400 A.D. to the Present in Different Regions of the World,” Paper prepared for the 1999 Meeting of the Peace Science Society in Ann Arbor, Michigan.
- 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.
- 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.
- CHAPMAN, T., AND P. G. ROEDER (2007): “Partition as a Solution to Wars of Nationalism: The Importance of Institutions,” *American Political Science Review*, 101(4), 677–691.
- CICCONE, A., AND M. JAROCINSKI (2010): “Determinants of Economic Growth: Will Data Tell?,” *American Economic Journal Macroeconomics*, 2(2), 222–246.
- CIOFFI-REVILLA, C. (1996): “Origins and Evolution of War and Politics,” *International Studies Quarterly*, 40(1), 1–22.
- COLLIER, P. (2007): *The Bottom Billion*. Oxford University Press, Oxford, United Kingdom.
- 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.
- 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.

- DIAMOND, J. (1997): *Guns, Germs, and Steel. The Fates of Human Societies*. W.W. Norton Co, New York, NY.
- DJANKOV, S., AND M. REYNAL-QUEROL (2010): “The Causes of Civil Wars,” mimeo UPF.
- DOWDEN, R. (2008): *Africa: Altered States, Ordinary Miracles*. Portobello Books Ltd, London, UK.
- EASTERLY, W., AND R. LEVINE (2012): “The European Origins of Economic Development,” NBER Working Paper No. 18162.
- 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 Study,” *American Economic Review*, 102(4), 1310–1342.
- ETHNOLOGUE (2005): *Languages of the World*, SIL International; Fifteenth edition.
- FEARON, J. D. (2003): “Ethnic Structure and Cultural Diversity by Country,” *Journal of Economic Growth*, 8, 195–222.
- (2004): “Separatist Wars, Partition, and World Order,” *Security Studies*, 13(4), 394–415.
- FEARON, J. D., AND D. D. LAITIN (2000): “Violence and the Social Construction of Ethnic Identity,” *International Organization*, 54(4), 845–877.
- (2003): “Ethnicity, Insurgency and Civil War,” *American Political Science Review*, 97, 75–90.
- (2011): “Sons of the Soil, Migrants, and Civil War,” *World Development*, 39(2), 199–211.
- (2012): “How Persistent Is Armed Conflict?,” mimeo, Stanford University.
- FENSKE, J. (2012a): “African polygamy: Past and present,” mimeo, Oxford University.
- (2012b): “Does Land Abundance Explain African Institutions?,” mimeo Oxford University.
- (2012c): “Ecology, Trade, and States in Pre-colonial Africa,” mimeo, Oxford University.
- GALOR, O. (2011): *Unified Growth Theory*. Princeton University Press, Princeton, NJ.
- 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.
- HARARI, M., AND E. L. FERRARA (2012): “Conflict, Climate and Cells: A Disaggregated Analysis,” mimeo MIT and Bocconi University.
- HEGRE, H., AND N. SAMBANIS (2006): “Sensitivity Analysis of Empirical Results on Civil War Onset,” *Journal of Conflict Resolution*, 50(4), 508–535.
- HERBST, J. (2000): *States and Power in Africa*. Princeton University Press, Princeton, NJ.
- HESTON, A., R. SUMMERS, AND B. ATEN (2011): “Penn World Table Version 7.0,” Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- HOROWITZ, D. L. (1985): *Ethnic Groups in Conflict*. University of California Press, Berkeley and Los Angeles, California.
- HUILLERY, E. (2009): “History Matters: The Long Term Impact of Colonial Public Investments in French West Africa,” *American Economic Journal - Applied Economics*, 1(2), 176–215.
- HUMPHREYS, M., AND H. A. MOHAMED (2005): “Senegal and Mali: A Comparative Study of Rebellions in West Africa,” in *Understanding Civil War: Evidence and Analysis Vol 1 Africa*, ed. by P. Collier, and N. Sambanis. The World Bank, Washington DC, USA.
- KAUFMAN, C. (1996): “Possible and Impossible Solutions to Ethnic Civil Wars,” *International Security*, 20(4), 133–175.
- (1998): “When All Else Fails: Separation as a Remedy for Ethnic Conflicts, Ethnic Partitions and Population Transfers in the Twentieth Century,” *International Security*, 23(2), 120–156.
- KAUFMANN, D., A. KRAAY, AND M. MASTRUZZI (2005): “Governance Matters IV : Governance Indicators for 1996-2004,” *Policy Research Working Paper Series*, *The World Bank*.
- 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.
- MARSHALL, M. G., K. JAGGERS, AND T. R. GUR (2011): *Polity IV Project. Political Regime Characteristics and Transitions, 1800-2010. User’s Manual*. Center for Systemic Peace.

- MATUSZESKI, J., AND F. SCHNEIDER (2006): “Patterns of Ethnic Group Segregation and Civil Conflict,” Working Paper, mimeo Harvard University.
- MEREDITH, M. (2005): *The State of Africa: A History of the Continent since Independence*. The Free Press, London, United Kingdom.
- MICHALOPOULOS, S. (2012): “The Origins of Ethnolinguistic Diversity,” *American Economic Review*, 102(4), 1508–1539.
- MICHALOPOULOS, S., AND E. PAPAIOANNOU (2012a): “National Institutions and African Development: Evidence from Partitioned Ethnicities,” NBER Working Paper 18275.
- (2012b): “Pre-colonial Ethnic Institutions and Contemporary African Development,” *Econometrica*, forthcoming.
- MONTALVO, J.-G., AND M. REYNAL-QUEROL (2005): “Ethnic Polarization, Potential Conflict, and Civil Wars,” *American Economic Review*, 95(3), 796–816.
- MURDOCK, G. P. (1959): *Africa: Its Peoples and their Culture History*. McGraw-Hill, New York, NY.
- (1967): *Ethnographic Atlas: A Summary*. The 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 D. PUGA (2012): “Ruggedness: The Blessing of Bad Geography in Africa,” *Review of Economics and Statistics*, 94(1), 20–36.
- NUNN, N., AND L. WANTCHEKON (2011): “The Slave Trade and the Origins of Mistrust in Africa,” *American Economic Review*, 101(7), 3221–3252.
- O’BRIEN, P. (1999): *Atlas of World History*. Oxford University Press, Oxford, United Kingdom.
- PERSSON, T., AND G. TABELLINI (2009): “Democratic Capital: The Nexus of Political and Economic Change,” *American Economic Journal: Macroeconomics*, 1(1), 88–126.
- PRUNIER, G. (2009): *Africa’s World War. Congo, the Rwandan Genocide, and the Making of a Continental Catastrophe*. Oxford University Press, Oxford, United Kingdom.
- PRZEWORSKI, A., M. ALVAREZ, J. CHEIBUB, AND F. LIMONGI (2000): *Democracy and Development: Political Institutions and Well-Being in the World 1950-1900*. Cambridge University Press, Cambridge, United Kingdom.

- RALEIGH, C., D. CUNNINGHAM, L. WILHELMSSEN, AND N. P. GLEDITSCH (2006): “PRIO Conflict Sites Codebook, 1946-2005,” Centre for the Study of Civil War (PRIO).
- RALEIGH, C., A. LINKE, AND C. DOWD (2012a): “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).
- (2012b): “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).
- ROHNER, D., M. THOENIG, AND F. ZILIBOTTI (2012): “Seeds of Distrust: Conflict in Uganda,” CEPR DP 8741.
- ROSS, M. L. (2012): *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton University Press, Princeton, NJ.
- SAMBANIS, N. (2000): “Partition as a Solution to Ethnic War: An Empirical Critique of the Theoretical Literature,” *World Politics*, 52(2), 437–483.
- SAMBANIS, N., AND J. SCHULHOFER-WOHL (2009): “What’s In A Line? Is Partition the Solution to Civil War?,” *International Security*, 34(2), 82–118.
- SILVA, J. M. S., AND S. TENREYRO (2006): “The Log of Gravity,” *Review of Economics and Statistics*, 88(4), 641–658.
- SILVA, J. M. S., S. TENREYRO, AND F. WINDMEIJER (2010): “Is it Different for Zeros? Discriminating Between Models for Non-Negative Data with Many Zeros,” CeMMAP Working Paper CWP20/10.
- SPOLAORE, E., AND R. WACZIARG (2009): “The Diffusion of Development,” *Quarterly Journal of Economics*, 124(2), 469–529.
- (forthcoming): “How Deep Are the Roots of Economic Development?,” *Journal of Economic Literature*.
- TOUVAL, S. (1967): “The Organization of African Unity and African Borders,” *International Organization*, 21(1), 102–127.
- WESSELING, H. L. (1996): *Divide and Rule: The Partition of Africa, 1880-1914*. Praeger, Amsterdam, Netherlands.
- 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(2), 316–337.

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

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

Panel A: All Ethnic Homelands

	Indicator	Number of Incidents		Excluding Capitals		Excluding Outliers (top 1%)	
	Likelihood	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)	<i>(0.00)</i>	<i>(0.28)</i>	<i>(0.00)</i>	<i>(0.02)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>
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)	<i>(0.13)</i>	<i>(0.20)</i>	<i>(0.01)</i>	<i>(0.05)</i>	<i>(0.01)</i>	<i>(0.01)</i>	<i>(0.01)</i>
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)	<i>(0.00)</i>	<i>(0.01)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>
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)	<i>(0.00)</i>	<i>(0.43)</i>	<i>(0.00)</i>	<i>(0.03)</i>	<i>(0.00)</i>	<i>(0.01)</i>	<i>(0.00)</i>

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

Panel B: Ethnic Homelands close to the National Border

	Indicator	Number of Incidents		Excluding Capitals		Excluding Outliers	
	Likelihood	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)	<i>(0.00)</i>	<i>(0.05)</i>	<i>(0.00)</i>	<i>(0.10)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>
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)	<i>(0.00)</i>	<i>(0.10)</i>	<i>(0.00)</i>	<i>(0.17)</i>	<i>(0.00)</i>	<i>(0.02)</i>	<i>(0.00)</i>
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)	<i>(0.00)</i>	<i>(0.04)</i>	<i>(0.00)</i>	<i>(0.01)</i>	<i>(0.00)</i>	<i>(0.02)</i>	<i>(0.00)</i>
Violent Incidents							
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)	<i>(0.00)</i>	<i>(0.07)</i>	<i>(0.00)</i>	<i>(0.14)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>

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

	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(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: Pre-colonial Conflict**

	ML	OLS	ML	OLS	ML	OLS	ML	OLS
	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>
	(1)	(3)	(3)	(4)	(5)	(6)	(7)	(8)
Pre-colonial Conflict Indicator	-0.2320 (0.2531)	-0.0386 (0.0264)						
Distance to Pre-colonial Conflict			-0.1414 (0.2044)	-0.0290 (0.0321)				
Slave Trades Indicator					0.0410 (0.1131)	0.0061 (0.0162)		
Log Number of Slaves (normalized by land area)							0.0225 (0.0244)	0.0079 (0.0080)
Land Area under Water	0.3282*** (0.0937)	0.0621*** (0.0125)	0.3217*** (0.0972)	0.0611*** (0.0130)	0.3246*** (0.0963)	0.0616*** (0.0126)	0.3298*** (0.0957)	0.1194*** (0.0250)
Land Area	0.0934 (0.0157)	0.0160* (0.0084)	0.0883 (0.0560)	0.0151* (0.0080)	0.084 (0.0559)	0.0144* (0.0080)	0.0835 (0.0559)	0.0264 (0.0162)
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.051	—	0.05	—	0.05	—	0.051	—
Adjusted R-squared	—	0.075	—	0.085	—	0.083	—	0.064
Observations	826	826	826	826	826	826	826	826

Table 2 - Panel B reports probit marginal effects (in odd-numbered columns) and OLS estimates (in even-numbered columns) associating ethnic partitioning with proxy measures of pre-colonial conflict. 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). In columns (1)-(4) we use data on pre-colonial conflict (in the period 1400-1700) from Besley and Reynal-Querol (2012) and Brecke (1999). In columns (5)-(8) we use data on enslavement during the African slave trades from Nunn (2008) and Nunn and Watcekon (2011). 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 Statehood & Development

	ML	OLS	ML	OLS	ML	OLS	ML	OLS
	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pre-colonial Kingdom Indicator	0.1233 (0.1317)	0.0178 (0.0193)						
Distance to Pre-colonial Kingdom			-0.0043 (0.3176)	-0.0092 (0.0409)				
Major City in 1400AD					0.0547 (0.2080)	-0.0042 (0.0344)		
Distance to Explorer's Routes							-0.0004 (0.0003)	0.0000 (0.0000)
Land Area under Water	0.3053*** (0.0988)	0.0590*** (0.0126)	0.3217*** (0.0949)	0.0608*** (0.0120)	0.3210*** (0.0959)	0.0613*** (0.0124)	0.3162*** (0.0980)	0.0605*** (0.0130)
Land Area	0.0788 (0.0582)	0.0135 (0.0086)	0.0868 (0.0580)	0.0147* (0.0085)	0.0858 (0.0564)	0.0149* (0.0079)	0.0817 (0.0571)	0.0142* (0.0081)
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.051	—	0.05	—	0.050	—	0.053	—
Adjusted R-squared	—	0.084	—	0.083	—	0.082	—	0.085
Observations	826	826	826	826	826	826	826	826

Table 2 - Panel C 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). Data on pre-colonial statehood (large kingdoms and empires) are taken from Besley and Reynal-Querol (2012) and O' Brian (1999). Data on major cities before the slave trades (in 1400) come from Chandler (1987). Data on the principal European explorer's routes come from Nunn (2009). 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 as a Function of Economic, Social, Cultural, and Political Ethnic Traits
Using Data from Murdock's Ethnographic Atlas

Panel A: Baseline Estimates. Pre-colonial Economic & Political Development

	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>	<u>SPLIT</u>	<u>FRAC</u>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Land Area under Water	0.2829** (0.1183)	0.0514*** (0.0162)	0.2528** (0.1155)	0.0475*** (0.0159)	0.3372*** (0.1109)	0.0613*** (0.0163)	0.2928** (0.1176)	0.0558*** (0.0162)
Land Area	0.1569** (0.0731)	0.0252*** (0.0090)	0.1628** (0.0721)	0.0257*** (0.0094)	0.1727** (0.0797)	0.0238** (0.0097)	0.1677** (0.0676)	0.0254*** (0.0084)
Settlement Patterns	0.0203 (0.0411)	0.0038 (0.0055)						
Share of Agriculture			0.0322 (0.0297)	0.0048 (0.0055)				
Political Centralization					-0.1965 (0.1354)	-0.0321 (0.0208)		
Class Stratification							-0.0242 (0.0569)	-0.0038 (0.0069)
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.068	—	0.068	—	0.085	—	0.068	—
adjusted R-squared	—	0.105	—	0.106	—	0.128	—	0.107
Observations	451	451	437	437	394	394	487	487

Table 3- Panel A 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)) reflecting early economic and political development. 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 as a Function of Economic, Social, Cultural, and Political Ethnic Traits
Using Data from Murdock's Ethnographic Atlas**

Panel B: Further Evidence

	Specification A		Specification B	
	Unconditional Relationship		Conditional Relationship	
	<u>Additional Variable</u>	<u>Obs.</u>	<u>Additional Variable</u>	<u>Obs.</u>
	(1)	(2)	(3)	(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
Alternative Agriculture 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

Table 3-Panel B 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.

Table 4: Ethnic Partitioning and Geographic Characteristics within Countries

	Dependent variable is:								
	Log Land Area	Log Area under Water	Mean Elevation	Land Suitability	Malaria Stability	Distance to the Sea	Diamond Indicator	Petroleum 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

The table 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 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 5: 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 6: 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 7: 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	
	<u>All Obs</u>	<u>No Capitals</u>	<u>All Obs</u>	<u>No Capitals</u>	<u>All Obs</u>	<u>No Capitals</u>	<u>All Obs</u>	<u>No Capitals</u>	<u>All Obs</u>	<u>No Capitals</u>	<u>All Obs</u>	<u>No Capitals</u>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(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	-2630.41	-2432.38	-2502.75	-2299.75	—	—	-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 8: Ethnic Partitioning and Civil Conflict
OLS Specifications**

	All Ethnicity-Country Homelands					All Ethnicity-Country Homelands Close to the National Border				
	OLS			OLS		OLS			OLS	
	ln(1+All Civil Conflict Events)			Linear Probability		ln(1+All Civil Conflict Events)			Linear Probability	
	<u>All Obs</u>	<u>No Capitals</u>	<u>No Outliers</u>	<u>All Obs</u>	<u>All Obs</u>	<u>All Obs</u>	<u>No Capitals</u>	<u>No Outliers</u>	<u>All Obs</u>	<u>All Obs</u>
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(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 9A: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis. Accounting for Unobservables**

	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	
	<u>All Obs</u>	<u>No Capitals</u>	<u>All Obs</u>	<u>No Capitals</u>	<u>All Obs</u>	<u>No Capitals</u>	<u>All Obs</u>	<u>No Capitals</u>	<u>All Obs</u>	<u>No Capitals</u>	<u>All Obs</u>	<u>No Capitals</u>
	(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	-1219.501	-3451.45	-3195.16	-1269.921	-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	
	<u>All</u>	<u>Border</u>	<u>All</u>	<u>Border</u>	<u>All</u>	<u>Border</u>	<u>All</u>	<u>Border</u>	<u>All</u>	<u>Border</u>
	(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	-3416.15	-1354.6	-3143.06	-1237.22	-2617.48	-894.705	-2558.18	-1026.96	-2635.13	-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 Conflict
Sensitivity Analysis: Accounting for Spillovers with Spatial Models**

Weighting Matrix	Linear in Euclidian Distance			Quadratic in Euclidian Distance		
	<u>Simple</u>	<u>Durbin</u>	<u>Generalized</u>	<u>Simple</u>	<u>Durbin</u>	<u>Generalized</u>
<u>Spatial Model Type</u>	(1)	(2)	(3)	(4)	(5)	(6)
SPLIT Indicator (Partitioning)	0.1891** (0.0765)	0.1754** (0.0757)	0.1822** (0.0765)	0.1844** (0.0756)	0.1582** (0.0777)	0.1685** (0.0781)
Log Likelihood	-1750.65	-1670.10	-1744.80	-1740.92	-1671.16	-1736.70
Observations	1182	1182	1182	1182	1182	1182
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

The table reports spatial auto-regressive model maximum-likelihood estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnicity level. The dependent variable 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. Columns (1) and (4) report simple spatial lag models that control for conflict in neighbouring regions. Columns (2) and (5) report Durbin spatial models that include as additional controls the vector of independent variables in neighbouring regions. Columns (3) and (6) report generalized spatial lag models that control both for conflicts in neighbouring regions and for the effect of the independent variables in nearby regions. In columns (1)-(3) we use a linear in Euclidian distance to the centroid of each country-ethnic region weighting matrix. In columns (4)-(6) we use a quadratic in Euclidian distance to the centroid of each country-ethnic region weighting matrix. 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 standard errors accounting for heteroskedasticity and spatial correlation. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 11: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Distinguishing between Conflict Incident Media Sources**

	All Observations				Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Data from All International Sources						
SPLIT Indicator (Partitioning)	0.4234*** (0.1657)	0.4439** (0.1852)	0.6224*** (0.1859)	0.5565*** (0.1732)	0.5995*** (0.1664)	0.5531*** (0.1760)
Log Likelihood	-3382.90	-3183.73	-3107.80	-3095.23	-2991.64	-2854.01
Observations	1182	1182	1182	1182	1170	1139
Panel B: Data only from NGOs and United Nation's Reports						
SPLIT Indicator (Partitioning)	0.9258*** (0.2995)	0.5588*** (0.2149)	0.7778*** (0.2779)	0.7943*** (0.2636)	0.7830*** (0.2757)	0.8183*** (0.2864)
Log Likelihood	-1330.075	-1164.565	-1125.269	-1120.363	-1028.719	-995.075
Observations	1038	1038	1038	1038	1026	1005
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in all columns is the total number of conflict incidents at the country-ethnic homeland level. In Panel A we use conflict events reported by international media, news agencies, NGOs, and the United Nations. In Panel B we use conflict events reported by NGOs and the United Nations. In columns (1)-(4) we estimate the specifications in the full sample. The specification in column (5) excludes country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specification in column (6) excludes 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. All specifications include country fixed effects (constants not reported). 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 12: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Controlling for Pre-colonial Conflict (1400-1700)**

	All Ethnicity-Country Homelands			Ethnicity-Country Homelands Close to the National Border		
	<u>ALL</u>	<u>Battles</u>	<u>Violence</u>	<u>ALL</u>	<u>Battles</u>	<u>Violence</u>
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Controlling fo Pre-colonial Conflict (1400-1700)						
SPLIT Indicator (Partitioning)	0.6309*** (0.1857)	0.6184*** (0.1753)	0.4958*** (0.1632)	0.8234*** (0.3079)	0.7975** (0.3649)	0.8486*** (0.2530)
Log Likelihood	-3600.55	-2628.39	-2500.89	-1426.05	-999.88	-979.25
Panel B: Controlling for Pre-colonial Kingdoms/Empires						
SPLIT Indicator (Partitioning)	0.5198*** (0.1813)	0.5257*** (0.1652)	0.4316*** (0.1585)	0.7633*** (0.2907)	0.7910** (0.3486)	0.8117*** (0.2299)
Log Likelihood	-3599.87	-2629.526	-2501.75	-1428.978	-1003.138	-981.882
Panel C: Controlling for Slave Trades						
SPLIT Indicator (Partitioning)	0.5198*** (0.1927)	0.5257*** (0.1779)	0.4316*** (0.1654)	0.7633*** (0.2980)	0.7910*** (0.3350)	0.8117*** (0.2402)
Log Likelihood	-3600.55	-2629.22	-2501.68	-1428.85	-1003.76	-980.99
Observations	1182	1182	1182	590	590	590
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating various aspects of civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in columns (1) and (4) is the total number of conflict incidents; the dependent variable in columns (2) and (5) is the total number of battles between government forces, rebel groups, and militias; the dependent variable in columns (3) and (6) is the number of violent events against civilian populations. In Panel A we estimate the specifications in the full sample. In columns (4)-(6) we restrict estimation 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). In Panel A we control for the distance of each ethnic homeland to the centroid of the closest war during the period 1400-1700, using data from Besley and Reynal-Quaerol (2012). In Panel B we control an indicator variable that takes the value of one when the historical homeland falls within the boundaries of a large pre-colonial kingdom and empire using data from Besley and Reynal-Quaerol (2012). In Panel C we control for the log of one plus the number of slaves at the ethnicity level normalized by the surface area of each homeland, using data from Nunn (2008) and Nunn and Watchekon (2011).

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. All specifications include country fixed effects (constants not reported). 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 13: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Alternative Measures of Ethnic Partitioning**

	All Ethnicity-Country Homelands						All Ethnicity-Country Homelands Close to the Border					
	<u>All Obs</u>	<u>No</u>	<u>No</u>	<u>All Obs</u>	<u>No</u>	<u>No</u>	<u>All Obs</u>	<u>No</u>	<u>No</u>	<u>All Obs</u>	<u>No</u>	<u>No</u>
	(1)	Outliers (2)	Capitals (3)	(4)	Outliers (5)	Capitals (6)	(7)	Outliers (8)	Capitals (9)	(10)	Outliers (11)	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 14: Ethnic Partitioning and Civil Conflict
Results with Ethnologue's Mapping of Linguistic 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-ethnic homeland level. 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.

Table 15: Ethnic Partitioning and Civil Conflict. Heterogeneous Effects

Panel A: Fragmentation, Size, and Landlocked

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT (Partitioning)	0.4614 (0.2915)	0.4283 (0.3073)	0.6553** (0.2690)	0.4526 (0.3023)	0.7451*** (0.2190)	0.6319** (0.2756)	0.4583** (0.2347)	0.3627 (0.2499)
SPLIT (Partitioning) Interacted with an Indicator of:								
High Ethnic Fractionaliza	0.3137 (0.8944)							
High Linguistic Fractionalization		0.3772 (1.0332)						
High Religious Fractionalization			-0.076 (0.1242)					
High Cultural Heterogeneity				0.3262 (0.9244)				
High Ethnic Polarization					-0.2928 (0.3209)			
High Ethnic Segregation						-0.0192 (0.2390)		
High Land Area (Size)							0.3897 (1.0509)	
Landlocked Indicator								0.6598* (0.3798)
Log Likelihood	-3602.165	-3601.703	-3603.129	-3602.073	-3602.318	-3603.188	-3601.676	-3599
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location & Geo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1182	1182	1182	1182	1182	1182	1182	1182

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating civil conflict with ethnic partitioning at the country-ethnic 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. All specifications include interactions terms of the SPLIT index with dummy variables identifying countries with high ethnic fractionalization (in (1)), linguistic fractionalization (in (2)), religious fractionalization (in (3)), cultural heterogeneity (in (4)), ethno-linguistic polarization (in (5)), ethnic segregation (in (6)), large in terms of size, (in (7)), and landlocked countries (in (8)). With all measures we use the median value of the respective variable. 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 15: Ethnic Partitioning and Civil Conflict. Heterogeneous Effects

Panel B: Institutional and Economic Development

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT (Partitioning)	0.8289*** (0.2182)	0.9139*** (0.2503)	0.9720*** (0.2782)	0.9153*** (0.2295)	1.1205*** (0.2954)	1.2715*** (0.3326)	0.8648*** (0.2596)	1.4545*** (0.3684)
SPLIT (Partitioning) Interacted with an Indicator of:								
High Rule of Law	-0.4397 (0.3677)							
Democracy Indicator		-0.8405*** (0.3963)			-0.6958** (0.2727)			
High Democratic Capital			-0.7152*** (0.2956)			-0.7032*** (0.2106)		
High Income				-0.7244*** (0.3449)	-0.6016** (0.3358)	-0.7195*** (0.2885)		
Executive Constraints at independence							-0.6346*** (0.2916)	-0.6467*** (0.2679)
High Income at Independence								-1.0945*** (0.3795)
Log Likelihood	-3601.185	-3596.972	-3597.876	-3597.585	-3593.267	-3592.267	-3599.011	-2914.94
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location & Geo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1182	1182	1182	1182	1182	1182	1182	1023

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating civil conflict with ethnic partitioning at the country-ethnic 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. All specifications include interactions terms of the SPLIT index with dummy variables identifying countries scoring high in a composite rule of law index (in (1)), classified as democracies (in (2) and (5)), scoring high in a democratic capital variable (in (3) and (6)), with high income (in (4)-(6)), with high executive constraints at independence (in (7)-(8)), and with high income in 1961 (in (8)). With all measures we use the median value of the respective variable. 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.

Appendix Table 1: Correlation Structure of the Main Dependent Variables

Panel A: Ethnic Homeland Level

All types of 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 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 homeland 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: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Controlling for Population in 2000**

	All Ethnic Homelands			Ethnic Homelands close to the National Border		
	<u>ALL</u>	<u>Battles</u>	<u>Violence</u>	<u>ALL</u>	<u>Battles</u>	<u>Violence</u>
	(1)	(2)	(3)	(4)	(5)	(6)
SPLIT Indicator (Partitioning)	0.5198*** (0.1731)	0.5257*** (0.1657)	0.4316*** (0.1600)	0.7633*** (0.2984)	0.7910** (0.3471)	0.8117*** (0.2403)
Log Likelihood	-3601.378	-2628.953	-2500.478	-1433.027	-1004.356	-983.873
Observations	1182	1182	1182	590	590	590
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating various aspects of civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in columns (1) and (4) is the total number of conflict incidents; the dependent variable in columns (2) and (5) is the total number of battles between government forces, rebel groups, and militias; the dependent variable in columns (3) and (6) is the number of violent events against civilian populations. In columns (1)-(3) we estimate the specifications in the full sample. In columns (4)-(6) we restrict estimation 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 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. All specifications include country fixed effects (constants not reported). 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 4: Ethnic Partitioning and Deadly Civil Conflict

	All Observations				Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)
SPLIT Indicator (Partitioning)	0.4176** (0.1923)	0.4572*** (0.1736)	0.5672*** (0.1721)	0.5463*** (0.1606)	0.5802*** (0.1533)	0.5762*** (0.1675)
Log Likelihood	-2555.53	-2352.632	-2292.014	-2281.25	-2193.339	-2094.972
Observations	1182	1182	1182	1182	1170	1139
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in all columns is the total number of deadly conflict incidents at the country-ethnic homeland level. To identify deadly incidents we go over the event description searching for keywords "kill, death, died". In columns (1)-(4) we estimate the specifications in the full sample. The specification in column (5) excludes country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specification in column (6) excludes 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. All specifications include country fixed effects (constants not reported). 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 5: 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 6: Ethnic Partitioning and Riots and Protests

	All Observations				Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)
SPLIT Indicator (Partitioning)	-0.1531 (0.2897)	-0.0726 (0.2051)	0.1372 (0.2176)	0.1253 (0.2156)	0.1367 (0.2120)	0.1069 (0.2298)
Log Likelihood	-1878.551	-1727.347	-1669.145	-1659.251	-1567.01	-1443.98
Observations	1182	1182	1182	1182	1170	1139
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating riots and protests with ethnic partitioning at the country-ethnic homeland level. The dependent variable in all columns is the total number of riots and protests at the country-ethnic homeland level. In columns (1)-(4) we estimate the specifications in the full sample. The specification in column (5) excludes country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specification in column (6) excludes 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. All specifications include country fixed effects (constants not reported). 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 7: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Accounting for Spillovers at the Country Level and at the Ethnic Family Level

	All Ethnic Homelands			Ethnic Homelands close to the National Border		
	<u>All Obs</u>	<u>No</u>	<u>No Outliers</u>	<u>All Obs</u>	<u>No</u>	<u>No Outliers</u>
		<u>Capitals</u>			<u>Capitals</u>	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Cross-Sectional Specifications						
SPLIT Indicator (Partitioning)	0.3337*** (0.1182)	0.3518*** (0.1119)	0.3625*** (0.1174)	0.4841*** (0.1186)	0.4617*** (0.1212)	0.4947*** (0.1255)
Log Number of Conflicts Ethnic Family	0.0755** (0.0337)	0.0718** (0.0332)	0.0828** (0.0330)	0.0837** (0.0361)	0.0759** (0.0366)	0.0828** (0.0386)
Log Number of Conflicts Country	0.1716*** (0.0368)	0.1715*** (0.0373)	0.1721*** (0.0379)	0.1740*** (0.0436)	0.1614*** (0.0445)	0.1629*** (0.0425)
adjusted R-squared	0.430	0.406	0.378	0.393	0.378	0.348
Panel B: Within Country Specifications						
SPLIT Indicator (Partitioning)	0.1867** (0.0928)	0.1931** (0.0894)	0.1834** (0.0885)	0.3124*** (0.1197)	0.3151*** (0.1179)	0.3329*** (0.1223)
Log Number of Conflicts at the Ethnic Family	0.0239 (0.0218)	0.0222 (0.0221)	0.0311 (0.0230)	0.0117 (0.0326)	0.0056 (0.0333)	0.0126 (0.0355)
Log Number of Conflicts at the Country	-0.0650** (0.0309)	-0.0599* (0.0309)	-0.0603** (0.0307)	0.0074 (0.0426)	0.0051 (0.0455)	0.0115 (0.0439)
adjusted R-squared	0.620	0.596	0.579	0.603	0.595	0.575
Observations	1182	1170	1139	590	575	565
Simple & Location & Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes

The table reports OLS estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnicity level, accounting for spatial spillovers. Panel A reports cross-sectional estimates. Panel B reports within country estimates (that include country constants). The dependent variable 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 specifications in columns (4)-(6) 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 (5) exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (3) and (6) exclude country-ethnic homelands where capital cities fall. In all specifications we control for the log of one plus the total number of all conflict incidents in each ethnic family and the log of one plus the total number of all conflict incidents in each country minus conflicts in each country-ethnic area. 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.

Appendix Table 8: Ethnic Partitioning and Civil Conflict

Sensitivity Analysis: Excluding Ethnic Homelands that Experienced Pre-colonial Conflict (1400-1700)

	All Ethnic Homelands			Ethnic Homelands close to the National Border		
	<u>ALL</u>	<u>Battles</u>	<u>Violence</u>	<u>ALL</u>	<u>Battles</u>	<u>Violence</u>
	(1)	(2)	(3)	(4)	(5)	(6)
SPLIT Indicator (Partitioning)	0.6001*** (0.1888)	0.5955*** (0.1712)	0.4835*** (0.1725)	0.8339*** (0.3044)	0.8257*** (0.3610)	0.8584*** (0.2490)
Log Likelihood	-3358.92	-2456.361	-2320.049	-1384.865	-986.274	-954.654
Observations	1120	1120	1120	566	566	566
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating various aspects of civil conflict with ethnic partitioning at the country-ethnic homeland level dropping from the estimation ethnic homelands that experienced a major war over the period 1400-1700. The dependent variable in columns (1) and (4) is the total number of conflict incidents; the dependent variable in columns (2) and (5) is the total number of battles between government forces, rebel groups, and militias; the dependent variable in columns (3) and (6) is the number of violent events against civilian populations. In columns (1)-(3) we estimate the specifications in the full sample. In columns (4)-(6) we restrict estimation 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 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. All specifications include country fixed effects (constants not reported). 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 9: Ethnic Partitioning and Civil Conflict

Sensitivity Analysis: Excluding Ethnic Homelands that Were Part of Large Kingdoms and Empires

	All Ethnic Homelands			Ethnic Homelands close to the National Border		
	<u>ALL</u>	<u>Battles</u>	<u>Violence</u>	<u>ALL</u>	<u>Battles</u>	<u>Violence</u>
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: O'Brian (1999) Classification						
SPLIT Indicator (Partitioning)	0.6878*** (0.2216)	0.6657*** (0.1907)	0.4405** (0.2231)	1.0696*** (0.3664)	1.0591*** (0.4144)	0.9429*** (0.2280)
Log Likelihood	-2118.812	-1587.995	-1442.143	-740.609	-541.919	-492.551
Observations	707	707	707	324	324	324
Panel B: Mudock (1967) Classification						
SPLIT Indicator (Partitioning)	0.5431*** (0.2149)	0.5171*** (0.2234)	0.5206*** (0.2298)	0.7070*** (0.2855)	0.7196** (0.3155)	0.9147*** (0.3156)
Log Likelihood	-1826.794	-1355.116	-1287.782	-769.822	-548.136	-536.68
Observations	583	583	583	305	305	305
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating various aspects of civil conflict with ethnic partitioning at the country-ethnic homeland level dropping from the estimation ethnic homelands that were part of large pre-colonial states, kingdoms, and empires using the classification of O'Brian (1999) in Panel A and using the classification of Murdock (1967) in Panel B. The dependent variable in columns (1) and (4) is the total number of conflict incidents; the dependent variable in columns (2) and (5) is the total number of battles between government forces, rebel groups, and militias; the dependent variable in columns (3) and (6) is the number of violent events against civilian populations. In columns (1)-(3) we estimate the specifications in the full sample. In columns (4)-(6) we restrict estimation 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 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. All specifications include country fixed effects (constants not reported). 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 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. Data on civil wars come from the Uppsala Conflict Data Program (UCDP)/International Peace Research Institute, Oslo (PRIO) Armed Conflict Dataset, Version 4-2006; and Raleigh, Cunningham, Wilhelmsen, and Gleditsch (2006).

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.

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
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.