DISCUSSION PAPER SERIES

No. 8676

THE LONG-RUN EFFECTS OF THE SCRAMBLE FOR AFRICA

Stelios Michalopoulos
and Elias Papaioannou

DEVELOPMENT ECONOMICS,
INTERNATIONAL MACROECONOMICS,
and ECONOMIC HISTORY INITIATIVE

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP8676.asp**

ISSN 0265-8003

THE LONG-RUN EFFECTS OF THE SCRAMBLE FOR AFRICA

Stelios Michalopoulos, Tufts University and Brown University
Elias Papaioannou, Dartmouth College, NBER and CEPR

Discussion Paper No. 8676
November 2011

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 8676

November 2011

ABSTRACT

The Long-Run Effects of the Scramble for Africa*

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, using information on the spatial distribution of African ethnicities before colonization, we show that borders were arbitrarily drawn. Apart from the land mass and water area of an ethnicity's historical homeland, no other geographic, ecological, historical, and ethnic-specific traits predict which ethnic groups have been partitioned by the national border. Second, using data on the location of civil conflicts after independence, we show that partitioned ethnic groups have suffered significantly more warfare; moreover, partitioned ethnicities have experienced more prolonged and more devastating civil wars. Third, we identify sizeable spillovers; civil conflict spreads from the homeland of partitioned ethnicities to nearby ethnic regions. These results are robust to a rich set of controls at a fine level and the inclusion of country fixed effects and ethnic-family fixed effects. The uncovered evidence thus identifies a sizable causal impact of the scramble for Africa on warfare.

JEL Classification: N17
Keywords: Africa, borders, conflict, development and ethnicities

Stelios Michalopoulos
Department of Economics
Tufts University
8 Upper Campus Rd, Braker Hall 302
Medford, MA 02155
USA

Email:
Stelios.Michalopoulos@tufts.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=165801

Elias Papaioannou
Dartmouth College
Economics Department
6106 Rockefeller Hall
Hanover, NH 03755
USA

Email:
Elias.Papaioannou@dartmouth.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=156542

* We thank Alberto Alesina, Maarten Bosker, 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, UC Irvine; UC Merced, American University, Brown University, CERGE-EI and the NBER Political Economy, NBER Summer Institute Meetings on the Development of the American Economy and Income Distribution and Macroeconomics for useful comments and suggestions. All errors are our sole responsibility.

Submitted 18 November 2011

1 Introduction

The predominant explanations on the deep roots of contemporary African underdevelopment are centered around the influence of Europeans during the colonial period (Acemoglu *et al.* (2001, 2002, 2005)), but also in the centuries before colonization when close to 20 million slaves were exported from Africa (Nunn (2008)). Yet in the period between the ending of the slave trades and the commencement of the colonial period, another major event took place in European capitals that according to the African historiography had malicious long-lasting consequences. The "Scramble for Africa" starts with the Berlin Conference of $1884-1885$ and is completed by the turn of the 20th century. In this brief period, Europeans partitioned Africa into spheres of influence, protectorates, colonies, and free-trade-areas. The borders were designed in European capitals at a time when Europeans had barely settled in Africa and had little knowledge of the geography and ethnic composition of the areas whose borders were designing. Despite their arbitrariness these boundaries endured after African independence in the 1960s. As a result in most African countries a significant fraction of the population is part of ethnic groups that have been partitioned by the national border.[1] A considerable body of research in African history (e.g. Asiwaju (1985); Dowden (2008); Wesseling (1996)) argues that the main impact of Europeans' influence in Africa was not colonization per se, but the improper border design. Partitioning, the argument goes, has led to ethnic struggles, patronage politics, and spurred civil conflict and underdevelopment. Yet there is little -if any- work that formally examines the impact of ethnic partitioning.[2]

This study is a first step to empirically assess the long-run effects of the scramble for Africa on civil conflict. While there is little disagreement that colonial borders were arbitrarily drawn, we start our empirical analysis by establishing formally their artificiality. We estimate various probabilistic models in an attempt to identify factors predicting which ethnicities have been partitioned. With the sole exceptions of the size of the historical homeland and area under water, we are unable to detect any other significant differences between partitioned and non-partitioned ethnicities with respect to geography (mean elevation, distance to the coast, land suitability for agriculture), the disease environment (malaria), natural resources (diamond mines and oil fields), proxy measures of early development, slavery and pre-colonial contact with Europeans. We further show that there are no systematic differences between split and

[1] Asiwaju (1985) identifies 177 partitioned ethnic groups that span all African borders. Englebert, Tarango, and Carter (2002) estimates that the population of partitioned ethnic groups is on average around 40% of the total population. Alesina, Easterly, and Matuszeski (2011) estimate that in several African countries the percentage of the population that belongs to a partitioned group exceeds 80% (e.g. Guinea-Bissau (80%); Guinea (88.4%); Eritrea (83%); Burundi (97.4%); Malawi (89%); Senegal (91%); Rwanda (100%); Zimbabwe (99%)).

[2] The cross-country studies of Alesina, Easterly, and Matuszeski (2011) and Englebert, Tarango, and Carter (2002) do touch upon this issue. We discuss the relationship of our work with these studies below.

non-partitioned ethnic groups, across several pre-colonial ethnic-specific institutional, cultural, and economic features, such as the degree of political centralization, the presence of property rights, class stratification, local institutions, the type of subsistence, etc. (Murdock (1967)).

We then employ the scramble for Africa as a quasi-natural experiment and assess the impact of partitioning on civil conflict, as this has been theorized to be the main channel of influence. Using detailed data on the location, duration, and total casualties of all civil wars in Africa in the post-independence period ($1970 - 2005$), we show that civil conflict is concentrated in the historical homeland of partitioned ethnicities. The positive effect of partitioning on all aspects of civil conflict retains its economic and statistical significance when we condition on country fixed effects and ethnic-family fixed effects to account for national and broad ethnic characteristics, respectively. Our most conservative estimates suggest that civil conflict intensity, as reflected in casualties and duration, is higher by approximately by 25% in areas where partitioned ethnicities reside (as compared to the homelands of ethnic groups that have not been separated by the national borders). Our analysis also uncovers substantial spillovers. Ethnic groups that were not directly affected by the artificial border design experience more conflict if they happen to be adjacent to partitioned ethnicities.

Historical Background

The "Scramble for Africa" starts in 1860s - 1870s when the French and the British begin exploring systematically Western Africa and sign bilateral agreements assigning spheres of influence. In the next 30 years, European powers signed hundreds of treaties that partitioned the largely unexplored continent into protectorates, free-trade-areas, and colonies. The event that stands for the partitioning of Africa is the conference that Otto von Bismarck organized in Berlin from November 1884 till February 1885. While the Berlin conference discussed only the boundaries of Central Africa (the Congo Free State), it came to symbolize the partitioning, because it laid down the principles that would be used among Europeans to divide the continent.[3] The key consideration of European leaders was to preserve the "status quo" preventing conflict among Europeans for Africa (as the memories of the European wars of the 18th-19th century were still alive). To this objective the Europeans divided areas and drew borders in maps, without taking into account local conditions and the ethnic composition of the areas. African

[3]The three major principles that emerged from the Berlin Conference were: First, the hinterland doctrine, according to which a power claiming the coast had also a right to its interior. Yet, the applicability of this principle became problematic, as it was not clear what exactly constitutes the hinterland. For example, at some point France demanded Nigeria claiming that it was the hinterland of Algeria. Second, the principle of effective possession required that Europeans 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 in the coastline.

leaders were not invited and had no say on the drawing of political boundaries.[4] Moreover, European leaders were in such a rush that they didn't wait for the new information arriving from explorers, geographers, and missionaries.

There is little disagreement among historians that the scramble for the continent was to a great extent artificial (see Asiwaju (1985) and Englebert (2009) for references). As the British prime minister at the time Lord Salisbury put it, "*we have been engaged in drawing lines upon maps where no white man's feet have ever tord; we have been giving away mountains and rivers and lakes to each other, only hindered by the small impediment that we never knew exactly where the mountains and rivers and lakes were.*" Asiwaju (1985) summarizes that "*the study of European archives supports the accidental rather than a conspiratorial theory of the marking of African boundaries.*"[5] In line with the historical evidence, Alesina, Easterly, and Matuszeski (2011) document that eighty percent of African borders follow latitudinal and longitudinal lines, more than in any other part of the world.

Several factors have been proposed to rationalize the arbitrary border design. First, at the time Europeans had limited knowledge of local geographic conditions, as with the exception of some coastal areas, the continent was largely unexplored. Second, Europeans were not drawing borders of prospective states or -in many cases- even colonies. Third, there was a constant imperialist back and forth with European powers swapping pieces of land with limited (at best) idea of what they were worth of. An illustrative example is the annexation of Katanga in Congo Free State that turned out to be the richest province. King Leopold demanded and eventually got Katanga in exchange for the Niari-Kwilu area (so-called after two rivers in Southern Congo) that the French insisted of getting themselves.[6] Fourth, while in most cases the treaties indicated that the exact boundaries would be set by special commissions, demarcation was poor and the commissions did not alter much. Fifth, Europeans were not willing to sacrifice their commitment not to go to war for any part of Africa.[7] In many cases

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

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

[6] Wesseling (1996) writes "*what impelled him [Leopold] was a general imperialist surge, the desire for compensation for the Niari-Kwilu, and the objective of making the new state as large as possible and filling as much of the Congo basin as possible.*"

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

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 ethnic leaders did not oppose the colonial design, as very little changed on the ground.[8]

The other major event in recent African history, the wave of independence, was also rapid. The independence of Northern African countries in the 1950s was soon followed by Ghana's and Guinea' independence in 1957 and in 1958, respectively. By the end of 1966, 40 countries had gained independence. While at the time, many proposed changing the colonial borders, African leaders and leaving Europeans did not touch this issue. The leaders of African independence believed that nation building would sideline ethnic divisions. Europeans' main objective was to maintain their special rights and corporate deals with former colonies, and as such, they were reluctant to open the border issue.[9]

Case Studies - Channels

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

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

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

[8] Asiwaju (1985) cites the Ketu king, saying that "*we regard the boundary (between Benin-Dahomey and Nigeria) separating the English and the French, not the Yoruba.*"

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

Democratic Forces of Casamance" was supported by the neighboring Guinea-Bissau (and to a lesser extent by Gambia), where the Diola exert a significant influence.[10]

Third, African borders are poorly demarcated and delineated due to the imprecise colonial treaties. This has resulted in many border disputes, especially when such poorly demarcated borders cause the partitioning of ethnic groups.[11] The conflict between Mali and Burkina Faso over the Agacher Strip, where the Bobo reside, illustrates the problems caused by the poor demarcation. The escalation of minor conflicts that started after independence resulted in a fully-blown war in 1985.[12] Imprecise colonial treaties seem to have contributed to conflict in Somalia as well (Higham (1985)), while the ambiguity of the tripartite treaty of 1902 between Britain, Italy and Ethiopia has also played a role in the Eritrea-Ethiopia war.

Fourth, Africa is characterized by patronage politics where dominant ethnic groups discriminate against minority groups, leading to the marginalization of certain under-privileged ethnicities (see i Miguel (2007) for a theoretical exposition and Wimmer, Cederman, and Min. (2009) for empirical evidence). Split groups are more likely to be smaller in the respective countries and thus more likely to be affected by such marginalization policies.[13] The central government, for example, tries to suffocate partitioned ethnicities by seizing property and imposing high taxation in the activities of a certain ethnic group (Bates (1981)). In many cases the neighboring country intervenes either to support their peers or to prevent migration and refugee flows. The conflict in the Alur-land offers an illustration of this type of violence. The Alur had been split between the Belgian Congo (Zaire) and the British Protectorate of Uganda during the late phase of the scramble for Africa ($1910 - 1914$). After independence when the regime of Mobutu Sese Seko initiated the subjugation of many minority groups in Congo, a large portion of the Alur in Congo moved to their historical homeland in Uganda. This in turn generated opposition from the Buganda (the main group in Uganda) leading to civil conflict in Uganda. In other instances partitioned ethnicities have reacted to their marginalization by participating in coups. 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, the Akan, and the other main ethnic groups in Ghana leading to civil warfare in the subsequent years.

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

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

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

[13] The median population density in 1960 for partitioned areas within country was 17 compared to 22 for non-partitioned ethnic areas.

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

Sixth, partitioning and border artificiality may lead to armed warfare by interacting with natural resources. If the historical homeland of a partitioned ethnic group is rich in natural resources (oil fields, diamond or gold mines) then the benefit of secession or irredentism increases. For example, armed conflict in the Cabinda enclave that is separated from the rest of Angola by a narrow strip of territory belonging to the Democratic Republic of the Congo is driven by the interaction between the artificial border design, the vast oil fields, and the partitioning of the Bakongo. (We present some evidence on this channel in Section 5.3.)

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

Related Literature

Our paper contributes to two main strands of literature. First, our work relates to studies that aim to uncover the deep roots of African - and more broadly global - development. This literature has mainly focused on the impact of colonization mainly via the formation of early institutions (see Acemoglu, Johnson, and Robinson (2005) for a review) or via human capital (see for example Glaeser, Porta, de Silanes, and Shleifer (2004) and Easterly and Levine (2009)). In contrast to this body of work, Gennaioli and Rainer (2006, 2007) and Nunn (2008) focus on the pre-colonial period. Gennaioli and Rainer (2006, 2007) and Michalopoulos and Papaioan-

nou (2010) show that pre-colonial ethnic institutions correlate significantly with contemporary economic development. Nunn (2008) provides evidence that the slave trades between the 15th-19th century has a significant long-run effect on economic development primarily by destroying social capital and spurring armed conflict (see Nunn and Wantchekon (2011) and Djankov and Reynal-Querol (2010)). Of relevance to our work are also studies showing a significant negative association between ethnic fragmentation/polarization and development (e.g. Easterly and Levine (1997); Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003); Montalvo and Reynal-Querol (2005a); see Alesina and Ferrara (2005) for a review). Ethnic fragmentation tends to lower public goods provision (Alesina, Baqir, and Easterly (1999), La-Porta, de Silanes, Shleifer, and Vishny (1999)), fuel authoritarianism (Aghion, Alesina, and Trebbi (2004)), and increase the likelihood of secession (e.g. Alesina and Spolaore (2003)), especially when ethnicities are segregated (Alesina and Zhuravskaya (2011)).

Our study contributes to this body of research, by emphasizing a somewhat neglected aspect of colonization; the drawing of political boundaries in the end of the 20th century creating partitioned ethnicities. Thus our work is mostly related to Alesina, Easterly, and Matuszeski (2011) who show that "artificial states" with straight borders and where a significant part of the population resides in more than one country, perform economically worse compared to countries with more organic (squiggly) borders. We focus on Africa, as the random design of colonial borders that endured after the African independence allows us to identify the causal effect of partitioning. Moreover, we examine the effects of the scramble on civil conflict, as the literature suggests this to be the main channel. Our regional focus allows us to control at a fine level for geography, the disease environment, natural resources, and other factors that a vast literature has emphasized as key determinants of civil conflict and under-development. Most importantly, with the inclusion of country fixed effects we account for all country-level factors that may affect civil war, such as institutional quality, ethnic fragmentation, foreign aid, national policies, etc.[14] We are also able to condition on ethnic-family fixed effects and thus control for broad cross-ethnicity differences in pre-colonial institutions, economic factors, and cultural traits, something crucial as recent works show a strong effect of cultural traits on long-run development (see Guiso, Sapienza, and Zingales (2006) for a review).

Second, our work contributes to the literature on the origins of civil conflict (see Collier and Hoeffler (2007) and Blattman and Miguel (2010) for reviews and Collier and Sambanis (2005) for case studies in Africa). This literature has examined the effect of many country

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

characteristics, such as income, natural resources and colonization on several aspects of civil conflict (see among others Collier and Hoeffler (1998), Collier, Hoeffler, and Soderbom (2004), and Fearon and Laitin (2003) for cross-country evidence). Other studies have attempted to link a country's ethnic composition to civil war eruption. While the correlation between ethnic fragmentation and civil war is weak, recent studies document interesting cross-country correlations associating various aspects of the societal structure with armed conflict. Montalvo and Reynal-Querol (2005b) and Esteban, Mayoral, and Ray (2010) show a strong negative correlation between ethnolinguistic polarization and conflict. Wimmer, Cederman, and Min. (2009) find that the likelihood of ethnic conflict increases when a large share of the population is excluded from power. Matuszeski and Schneider (2006) document that the likelihood, duration, and intensity of civil wars is much higher in countries where ethnicities are clustered in specific areas within a country. Most closely related to our work is the study of Englebert, Tarango, and Carter (2002) who shows a positive cross-country correlation between proxy measures of suffocation and dismemberment and political violence, secession attempts, border disputes, and civil warfare.

The correlations found in studies exploiting cross-country variation in border design and the distribution of ethnicities are informative; yet they cannot be causally interpreted (see Blattman and Miguel (2010) for a discussion). The main reason is that the process of border drawing is historically related to the process of state formation and is thus associated invariably with both voluntary and forced movements of people. Our study accounts for some of the shortcomings of cross-country studies. First, it establishes that African borders are to a great degree artificial by showing that there are no systematic differences in geographic, economic, institutional, and cultural characteristics between partitioned ethnicities and groups that have not been separated by the national border. 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 itself. Third, we can control for local factors that affect civil conflict at a fine level. Fourth, our regional approach allows us to condition on country and ethnic-family characteristics.

Structure

In the next section we discuss how we identify partitioned ethnic groups combining pre-colonial ethnic maps with contemporary political ones. We then detail how we use the regional data on civil war to construct estimates of civil conflict incidence, duration, and casualties at the ethnicity level. In Section 3 we examine whether there are systematic differences between partitioned and non-partitioned ethnic groups with respect to an array of geographic,

colonial and pre-colonial features. Section 4 reports our estimates on the local effect of ethnic partitioning on various aspects of civil conflict (number of war incidents, total casualties, and duration). In Section 5 we model spillovers and examine whether civil conflict spreads from the historical homelands of partitioned ethnic groups to adjacent regions. We then explore how the presence of natural resources in partitioned ethnic homelands increases the likelihood of conflict. We also report several sensitivity checks. In section 6 we conclude, discussing some possible avenues for future research.

2 Data

2.1 Identifying Partitioned Ethnic Groups

We identify partitioned groups projecting national borders in 2000, as portrayed in the Digital Chart of the World (Figure 1*b*) on George Peter Murdock's Ethnolinguistic Map (1959) that portrays the spatial distribution of ethnicities at the time of European colonization in the mid/late 19th century (Figure 1*a*).

Figure 1*a*

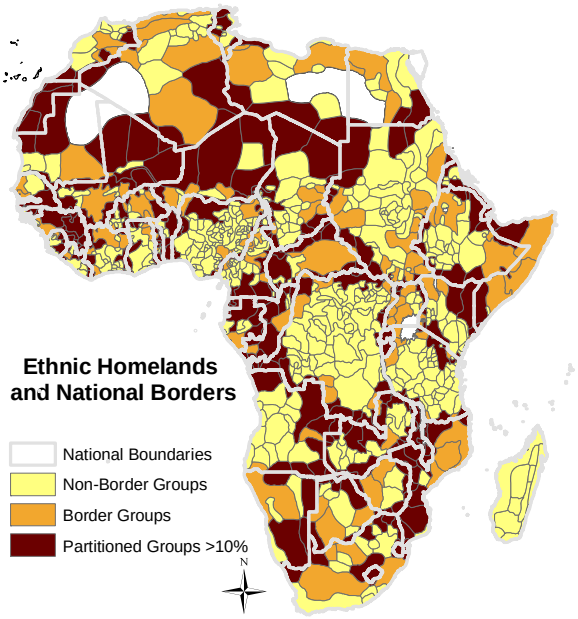


Figure 1*b*

Murdock's map scrambles Africa into 843 ethnic polygons. The mapped ethnicities correspond roughly to levels $7-8$ of the Ethnologue's language family tree. 8 areas are classified as "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 Ethnolinguistic Map, the homeland of 358 groups falls

into more than one contemporary countries. Yet for several of these groups the overwhelming majority of their homeland (usually more than 99%) falls into a single country. For example, 99.5% of the total surface area of the Ahaggaren falls into Niger and only 0.5% falls into Algeria. Since Murdock's map is bound to have some drawing error, we identify as partitioned groups those ethnicities with at least 10% of their total surface area belonging to more than one country ($SPLIT$). 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. Appendix Table A gives the list of all partitioned ethnic groups. When we use a more restrictive threshold of 20% there are 164 ethnicities partitioned across the national border.

Our procedure identifies the major ethnic groups that have been split by African borders. For example, the Maasai have been 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 of large partitioned ethnic groups include the Hausa (split between Nigeria and Niger), the Ababda (split between Egypt and Sudan), and the Bararetta (split between Kenya and Somalia). We also checked whether our codification of partitioned ethnicities is in line with Asiwaju (1985), who provides the only (to our knowledge) codification of partitioned ethnicities in Africa. Our strategy identifies almost all ethnic groups that Asiwaju (1985) lists as partitioned.

We also construct a continuous index of partitioning in the spirit of the cultural fragmentation indicators (e.g. Alesina *et al.* (2003)). The continuous index of partitioning reflects the probability that a randomly chosen pixel of the historical homeland of an ethnic group falls into a different country. Thus the index takes on positive values for partitioned ethnicities and zero otherwise. 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.[15]

[15] Our preferred measure of partitioning is the binary index for four reasons: First, all studies in African historiography suggest that what matters for civil conflict and under-development is whether an ethnicity has been partitioned or not rather than the degree of the split. Second, there is no clear reason on why the propensity to conflict should monotonically increase with the degree of partitioning. Third, as Murdock's tribal map certainly contains drawing error, this will be reflected much more in the continuous measure (as compared to the binary index). Fourth, it is much easier to interpret the estimates on the binary index. Nevertheless, to show that our results are not sensitive to the index of partitioning in Section 3 and Section 5.5 we report specifications with the continuous index ($FRAC$).

2.2 Civil War Data

The main data source for the occurrence and duration of civil wars comes from the Uppsala Conflict Data Program (UCDP)/International Peace Research Institute, Oslo (PRIO) Armed Conflict Dataset, Version 4 – 2006, initially assembled by Petter, Wallensteen, Eriksson, Sollenberg, and Strand (2002) and extended to cover the period 1946 – 2005.[16] We focus on armed conflicts that started after 1970 when the majority of African states had gained independence.[17] Following the literature on civil war we focus on both internal armed conflicts and internationalized internal armed conflicts. The former is defined as conflict between the government of a state and one or more internal opposition group(s) without intervention from other states, whereas the latter occurs between the government of a state and one or more internal opposition group(s) with the intervention from other states (secondary parties) on one or both sides. According to the UCDP/PRIO dataset since 1970 in Africa there have been in total 49 civil wars. 7 are classified as internationalized internal (e.g. the civil war of 1990 in Rwanda when Congo intervened and the civil conflict in the late 1990s in Guinea-Bissau where Guinea and Senegal intervened) whereas the majority, 42 wars, are classified as internal armed conflicts.

In this section we discuss how we construct ethnic-specific measures of the incidence, location, casualties and duration of civil conflict in Africa. Table 1 reports summary statistics for all civil conflict measures. Panel A gives descriptive statistics for the full sample of the 834 ethnicities, while Panels B and C report summary statistics for the partitioned and non-partitioned groups, respectively. Panel D reports the tests of means/medians between the split and non-split groups.

2.2.1 Civil Conflict Incidence and Location

To obtain ethnic-specific measures of civil conflict, we use the dataset of Raleigh, Cunningham, Wilhelmsen, and Gleditsch (2006) that assigns to each conflict reported in the PRIO coding a centroid (in latitude and longitude coordinates) with a corresponding radius in kilometers. The coordinates represent general estimates of where battles have occurred with the radius indicating the largest geographic extent of the conflict zone.[18] The location of several conflicts

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

[17] In particular, we are considering conflicts that are classified with a start date as early as 1970 where the date represents when the conflict for the first time reached 25-battle-related deaths in a calendar year. The results are similar if we include the 1960′s.

[18] There are limitations regarding the construction of geographically referenced variables. For example, the authors note that at a given point in time, the actual conflict zone might be more constrained than the maximum

did change over the duration of the conflict. For example, the long-lasting Liberian civil war took place in 3 different (partially overlapping) conflict zones.[19] Overall, the 49 African civil wars between 1970 and 2005 played out in 77 conflict zones.

Based on Raleigh, Cunningham, Wilhelmsen, and Gleditsch (2006) dataset we generate a map depicting each conflict zone associated with a civil war in Africa from 1970 to 2005. Then we project the constructed map on top of Murdock's ethnolinguistic map (Figure 1*a*). This allows us to identify in a systematic way which ethnicities have been affected by civil conflict during this period. We construct two alternative indexes of civil war incidence at the ethnicity level. The first measure captures the number of civil wars that have affected the historical homeland of each tribe. If the civil war has changed location over time we combine all conflict zones in one. Hence, the index capturing the number of civil wars does not take into account the fact that the zone of a conflict may migrate over time. Thus, we also calculate the number of conflict zones that have affected the historical homeland of each ethnicity. Figures 2*a*-2*b* portray the spatial distribution of civil war incidents and civil conflict zones at the ethnic-homeland level.

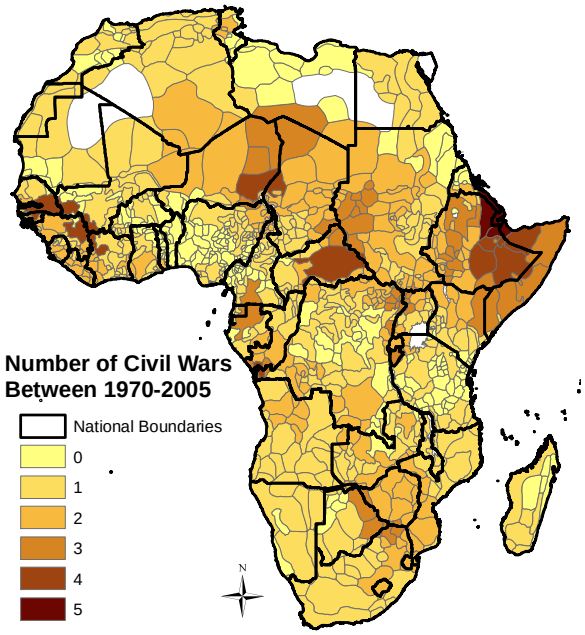


Figure 2*a*

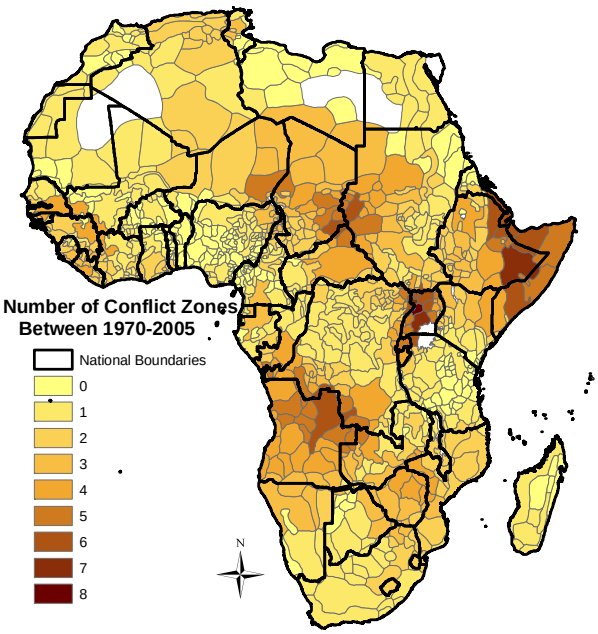


Figure 2*b*

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

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

Table 1 - Panel A illustrates the large variation in the incidence of civil war across ethnic areas. Out of the 834 ethnicities 343 experienced a civil war in their historical homeland; 199 ethnicities experienced two distinct incidents of armed conflict; 54 ethnicities experienced 3 civil wars while 12 tribes were affected by four or even five conflicts. Civil conflict is significantly more pervasive in the historical homeland of partitioned ethnic groups (as compared to non-split ethnicities). On average non-partitioned ethnic groups have experienced one civil conflict (Panel C) while partitioned ethnic groups have experienced 1.41 (Panel B). The groups with the highest incidence of civil war are the Afar and the Esa, which during the period $1970-2005$ have experienced 5 civil wars. Both groups have been greatly impacted by the artificial border design with the Afar being partitioned between Ethiopia, Eritrea and Djibouti, and the Esa being split between Ethiopia and Somalia. The number of conflict zones also varies considerably from 0 to 8 (Table 1 - Panel A). The test of means/medians in Table 1 - Panel D shows that conflict zones are much more likely to affect the historical homeland of partitioned ethnicities.

2.2.2 Civil Conflict Casualties

To construct ethnic-specific measures of combat deaths we use data from Lacina and Gleditsch (2005). This dataset presents for each conflict-year a low, a high, and a best estimate of the number of civilians and combatants killed in the course of combat.[20] After combining casualties with the PRIO data, we construct ethnicity-specific estimates of civil war casualties with the following procedure. First, we calculate for each conflict zone the total number of casualties by summing up the battle-related deaths across all conflict years. For example, in the case of the civil war in Sierra Leone (event ID 187) the total battle fatalities between 1991 and 2000 sum up to $12,997$ individuals. Second, we calculate the overall area that a conflict zone extends to. The Sierra Leone civil war affected a total area of $54,287$ square kilometers. Third, we estimate for each ethnic group how much of its homeland has been affected. In the above example, the Sierra Leone civil war involved $18,770$ out of $22,946$ square kilometers of the Mende's historical homeland. Thus, 35% of the civil war in Sierra Leone has taken place within the Mende's historical territory. Fourth, assuming that total casualties are distributed uniformly across the conflict zone, we derive estimates of ethnic-specific casualties by multiplying the total number of casualties of the conflict zone with the fraction of each ethnicity's land that falls within the conflict zone. This implies that 35% of the $12,997$ battle deaths of the Sierra Leone civil war, i.e. $4,497$, casualties are assigned to the Mende homeland. Fifth, we repeat this calculation for each conflict zone. For instance Mende's homeland takes up 29% of zone 1 of

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

civil war in Liberia -event ID 146- that has caused a total of $4,058$ battle deaths; this translates into $1,175$ casualties for Mende's historical homeland. Similarly, conflict zone 2 of the Liberian civil war -event ID 146- adds an additional 1011 casualties bringing the total number of civil war fatalities of the Mende to $6,680$ individuals.[21] Since larger ethnic areas are more likely to be affected by a conflict, we normalize the tribe-specific battle deaths by the surface area of the historical homeland of each group. Hence, the Mende have a casualty rate of 291 casualties per thousand of square kilometers. Figure $3a$ plots civil conflict casualties for the 834 ethnic areas in our sample.

The number of casualties per thousand of square kilometers is an indicator of how destructive civil war has been for each ethnic homeland. The correlation between casualties and civil war occurrences reported in Appendix Table 1 is positive $(0.36 - 0.58)$ but not overwhelming. Ethnic groups like the Wanga and the Sabei, which have been partitioned by the Kenya-Uganda border, have been involved in only one civil war between 1970 and 2005; however, they have lost thousands of people during this war. The average casualty rate (reported in Table $1A$) is 40 fatalities per thousand of square kilometers. Partitioned ethnic groups have experienced more deadly wars resulting on an average of 50 casualties per thousand of square kilometers (Panel C) as compared to an average of 35 casualties for non-split ethnicities (Panel B). The average rate, however, masks a great deal of heterogeneity. In particular, 50% of tribal areas experienced less than 3.2 casualties per thousand kilometers showing the highly skewed distribution of civil war casualties. There are also sizable (and statistically significant) differences in the median casualty rate in the historical homeland of partitioned and non-partitioned ethnicities amounting to 11 and 2.6 (see Panel D).

2.2.3 Conflict Duration

We obtain ethnic-specific estimates on civil war duration using a procedure analogous to the incidence/location index construction. First, we calculate for each ethnic area the fraction of land that has been affected by a civil war. In the case of the Mende 81% of the historical homeland has been affected by the Sierra Leone civil war -event ID 187- whereas 78% and 85% of its territory has been affected by zone 1 and zone 2 of the Liberian civil war -event ID 146-, respectively. Second, we weight the duration of each conflict with the fraction of the tribal area involved. Since the civil war in Sierra Leone lasted 10 years (1991-2000), the effective duration for Mende's homeland is $10 * 0.81 = 8.1$ years. Third, we sum the effective duration across all conflicts for each ethnicity. Thus, the Mende have been were under civil war for 17.04 years. Figure $3b$ graphs the duration of civil conflict across African ethnic groups.

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

For ethnic homelands that experienced a civil war, conflict duration varies from 1 year (which is the case for 27 conflict zones) to 22 years (corresponding to the long-lasting civil war in Sudan (1983-2004)). There are ethnic areas that have experienced incessant warfare as it is the case of the Alur (partitioned between Zaire and Uganda) registering the highest duration of warfare across African tribes.[22] On average ethnic areas have been under civil conflict for 7.75 years (Table 1*A*). The difference in civil war duration between partitioned and non-partitioned groups is a bit less than a year (8.29 and 7.54 years); the difference in the median duration is, however, a striking 3 years.

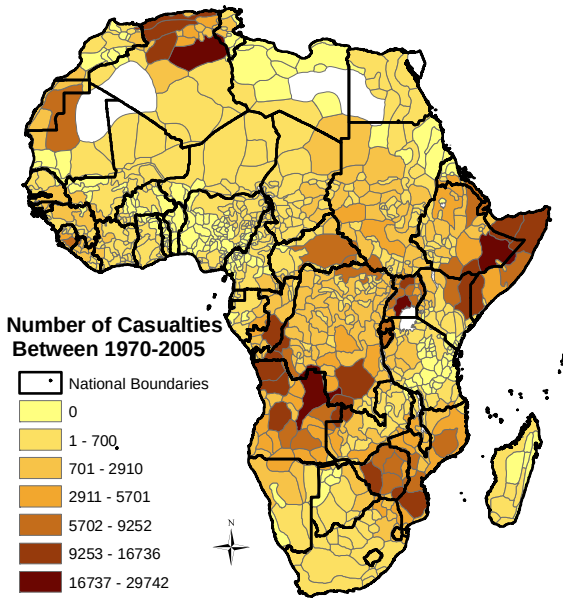


Figure 3*a*

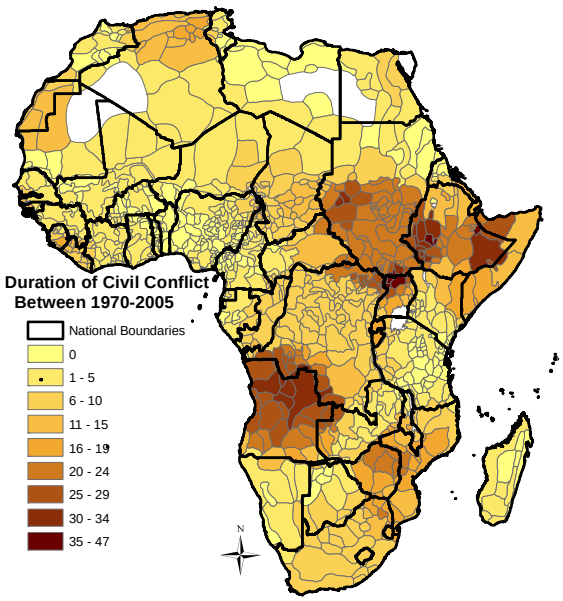


Figure 3*b*

3 Are African Borders Artificial?

3.1 Empirical Specification

While the anecdotal and historical evidence points out that Europeans did not take into account local conditions and ethnic characteristics during the scramble for Africa in the late 19th century, it is important from an econometric standpoint to investigate whether there are systematic differences between partitioned and non-partitioned ethnicities.[23] We test whether there are

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

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

geographic, economic, institutional, and other tribe-specific factors that predict whether an ethnic group is partitioned by the colonial borders. We do so by estimating probabilistic and OLS models of the following form:

$$SPLIT_i(FRAC_i) = a + X_i'\Psi + Z_i'\Theta + a_j + e_t. \quad (1)$$

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 states.[24] We also explore the robustness of our results using the continuous measure of partitioning, $FRAC_i$, as the dependent variable. X_i is a vector of geographic, ecological, natural resource and early development indicators at the ethnicity level; Z_i is a vector of ethnic-specific pre-colonial institutional, cultural, and economic traits, extracted from Murdock's (1967) Ethnographic Atlas available for a subset of African ethnicities. Appendix Table 2 presents the summary statistics for all variables. In all specifications we include region constants (a_j); this allows us to account for the somewhat different timing of colonization across Africa and also for unobserved differences across regions (the regional classification to North, South, East, West and Central Africa follows Nunn (2008)).

3.2 Results

Table 2 reports the results. Odd-numbered specifications report probit (maximum-likelihood) marginal effects, where the dependent variable is the benchmark binary partitioning index ($SPLIT$). Even-numbered columns report LS estimates, where the dependent variable is the continuous measure of partitioning ($FRAC_i$).[25]

Geographical, Ecological, and Natural Resource Measures In Panel A we explore the role of geographic, ecological, and natural resources. In columns (1)-(2) we regress the binary and continuous measures of partitioning on log surface area and log area under water. Ethnic groups spanning large territories in the pre-colonial period are more likely to be partitioned. This finding is in line with the historical evidence that colonizers drew borders in an arbitrary manner, drawing lines across imprecise maps. The estimates further show that ethnicities residing in areas with larger water bodies (lakes and rivers) were more likely to find themselves split by the national boundaries. This result is in accord with the historical

Moreover, we had the secession of Eritrea from Ethiopia (in 1993) and the unification of Tanganyika and Zanzibar (in 1964).

[24] Results are similar if use larger cut-off thresholds, such as a 20% or a 30%, to identify partitioned ethnicities.

[25] The results are similar if we estimate Tobit models that account for truncation of continuous partitioning index.

narrative that Europeans tried to use some natural barriers while delineating the spheres of their influence.

In columns (3)-(4) we augment the specification with an index reflecting land's suitability for agriculture and elevation. Mean elevation enters with an insignificant estimate in both models. Land suitability for agriculture enters with (weakly) significant estimate when we use $SPLIT$ as the dependent variable; yet the estimate turns insignificant when we use the broad partitioning index (in (4)) or when we drop the regional constants (results not shown).

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

In columns (7)-(8) we include in the empirical specification two indicators identifying ethnic areas with diamond mines and petroleum fields. While in the initial phase of colonization Europeans were mostly interested in agricultural goods, adding these two indicators allows us to investigate whether partitioned ethnicities differ from non-partitioned ethnic groups in terms of natural resources. This is important as one of the most robust correlates of civil warfare across countries is oil and other underlying hydrocarbon deposits, as well as diamonds (e.g. Ross (2006); Herge and Sambanis (2006)). Again there are no systematic differences in natural resources among split and non-split groups.

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

In columns (5)-(6) we explore whether colonial powers took into account pre-existing economic development when designing colonial borders. We proxy the pre-slave trade level of economic development using an indicator variable that equals one when a city with population exceeding $20,000$ people in 1400 AD was present in the historical homeland of an ethnicity and zero otherwise (using data from Chandler (1987)). There is no evidence that ethnicities with

urban centers before colonial times managed to escape ethnic partitioning.

In columns (7)-(8) we examine whether there are differences in development between partitioned and non-split ethnic groups at the time of independence. As there is no data on economic performance at the ethnic level for the 1960s, we use log population density in 1960 (using data from UNESCO). In a Malthusian regime where richer areas are more densely populated, the insignificant estimate on log population density in 1960 implies that there were no systematic differences in economic performance and urbanization rates between partitioned and non-partitioned ethnicities. Hence, any contemporary divergence in civil war intensity may not be attributed to differences in economic conditions in the eve of African independence.

Ethnic-Specific Pre-colonial Traits In Panel *C* of Table 2 we examine whether ethnic-specific pre-colonial institutional, cultural, and economic traits correlate with partitioning, using information from Murdock's (1967) Ethnolinguistic Atlas. The sample size drops because Murdock (1967) does not provide information for all the ethnicities in his Ethnographic Map (1959). In all specifications we include region fixed effects and the two geographical measures, size and land area under water, that have been found to be significant predictors of partitioning.

In columns (1)-(2) we investigate whether Europeans took into account the degree of political centralization of the African ethnicities when designing the borders. Following Gennaioli and Rainer (2006, 2007), we proxy political centralization with an indicator variable that equals zero when Murdock assigns an ethnicity either as "*stateless*" or "*a petty chiefdom*" (e.g. Xam or the Ibo); and becomes 1 when the ethnicity is part of either a "*large paramount chiefdom*" or a "*large state*" (e.g. Thonga and Zulu). The political centralization enters with a negative coefficient but the estimate is statistically indistinguishable from zero.

In columns (3)-(4) we examine whether the societal structure correlates with partitioning using Murdock's class stratification index. The index ranges from zero, indicating societies without any class distinctions, to four for ethnicities with significant class and wealth distinctions. There is no evidence that partitioned ethnic groups differ in regards to the complexity of societal structure.

African scholars (e.g. Hopkins (1973); Austin (2008)) argue that pre-colonial economic and institutional development was higher in areas with intensive use of agriculture (Fenske (2010) provides empirical evidence supportive of this conjecture); thus in columns (5)-(6) we augment the specification with a $0-10$ index measuring the importance of agriculture for subsistence at the ethnicity level, failing again to detect a significant correlation with partitioning.

In columns (7)-(8) we examine whether ethnic partitioning is systematically related to

pastoralism using a $0-10$ range index of ethnicity's dependence on animal husbandry. The index enters with an insignificant estimate, suggesting that there are no major differences in pastoral activities between split and non-partitioned ethnic groups.

In Appendix Table 3 we further explore the association between partitioning and numerous other ethnic-specific variables from Murdock (1967) reflecting among others: the size of precolonial settlements, the type of family organization (monogamous/polygyny), the presence of rules for inheritance, the dependence of the economy on various activities and the role of clans, failing again to detect significant differences between split and non-partitioned ethnic groups.

Summary The results reported in Tables 2 and Appendix Table 3 support the anecdotal and historical evidence on the arbitrary design of African borders. Europeans largely ignored local conditions when designing the colonial borders. The analysis further shows that early contact with Europeans or pre-colonial development does not correlate with ethnic partitioning; and partitioned ethnicities do not differ systematically across numerous ethnic-specific institutional, cultural, and economic features. Out of dozens of potentially relevant variables, only surface area and the presence of water streams correlate robustly with partitioning. Perhaps more importantly, the overall explanatory power of the models is poor. Mc Fadden's pseudo-R^2 (that compares the log likelihood value of the constant-only model with that of the full specification) is low across all permutations, at most 0.10. Likewise, the R^2 of the OLS models is always below 0.11. 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 Panel A (not reported) predicts correctly $(G(X_i'\Psi + Z_i'\Theta + a_j) > 0.5)$ only 29 out of the 231 partitions with the benchmark index $(SPLIT)$ whereas in the specification with all the ethnic-specific pre-colonial features used in Panel C the model predicts accurately only 28 out of the 149 partitions. So, although we cannot rule out the possibility that unobservable characteristics may correlate with partitioning, the results suggest that ethnic partitioning is not correlated with observable factors that may independently affect civil conflict.

4 Partitioning and Civil Conflict

4.1 Econometric Specification

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

$$y_i^{f,c} = a_0 + \gamma SPLIT_i + X_i'\Phi + f(LOCUS_i) + a_c + a_f + \varepsilon_i^{f,c}. \quad (2)$$

The dependent variable, $y_i^{f,c}$, reflects civil conflict in the historical homeland of ethnic group i. Ethnicity i belongs to ethnic cluster f and country c.[26] $SPLIT$ is an indicator that takes the value 1 when a tribal area is partitioned by the political boundaries. Thus, the coefficient γ captures the direct (local) effect of ethnic partitioning on civil conflict.

Vector X_i' includes geographical controls, like surface and water area; ecological features, such as a malaria stability index and land's suitability for agriculture; natural resources reflecting the presence of diamond mines and petroleum fields; early development proxies such as having a major city in 1400. In the Data Appendix we provide detailed variable definitions and sources. To capture unobservable location characteristics that vary smoothly in space in many specifications we introduce a cubic polynomial ($f(LOCUS_{i,c})$) in latitude and longitude of the centroid of each ethnic group.[27] Moreover, to further account for the tribal location, we also control for the distance to the coast, the distance to the national border and the distance to the capital city.

We also present specifications including country fixed effects (a_c) and ethnic family fixed effects (a_f). Country constants capture nationwide factors that a vast literature on civil warfare has identified as significant correlates of armed conflict. The ethnic-family fixed effects capture broad cultural, institutional, and other hard-to-observe ethnic-specific factors that may influence the propensity of ethnicities to engage in warfare and development (Murdock assigns the 834 groups into 96 ethnolinguistic clusters/families). Doing so we are able to uncover the effect of partitioning across groups belonging to relatively homogeneous ethnic clusters. The latter include the Bedouin Arabs, the Tuareg, the Southwestern Bantu to name a few.

A notable fraction of the observations on civil warfare (number of civil wars, number of war zones, casualties, duration) takes on the value of zero. Moreover, the civil war measures are skewed since there are many observations close to zero and a few extreme observations in the right tail of the distribution (See Table 1 − Panel A). Hence, we estimate non-linear ML specifications with a Poisson and/or a Negative Binomial process (Wooldridge (2002)). Using non-linear estimators is appealing because they do not require log-linearizing the dependent variable and thus preserve the higher moments of the distribution (see Silva and Tenreyro (2006) and Silva, Tenreyro, and Windmeijer (2010)). To illustrate the robustness of our estimates, we

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

[27] Letting x denote latitude and y denote longitude the polynomial reads: $x+y+x^2+y^2+xy+x^3+y^3+x^2y+xy^2$ (see also Dell (2010)). As there might be some concerns of over-fitting, we report also results without the polynomial. We also estimated specifications adding a cubic polynomial in distance to sea coast finding similar results (not reported).

also report log-linear LS specifications taking the log of one plus the respective civil conflict measures.

In all specifications we account for spatially correlated residuals ($\varepsilon_i^{f,c}$) clustering standard errors at the country level and at the ethnic-family level using the multi-way method of Miller, A.Cameron, and Gelbach (2011). This correction also accounts for arbitrary residual correlation within each country and each ethnic family. Moreover, the inclusion of ethnic family and country fixed effects further accounts for spatial correlation. We also estimated standard errors using Conley's method to account for spatial dependence of an unknown form, finding similar (and if anything less conservative) standard errors.

4.2 Civil Conflict Incidence

In Table 3 we examine the effect of partitioning on the incidence of civil conflict. Both panels report Poisson ML estimates where the dependent variable equals the number of civil conflicts (Panel A) and the number of conflict zones (Panel B) that affect each ethnic homeland.

Specification (1) shows that, conditional on region fixed effects, log population density at independence, surface area and area under water (the only geographic variables found to correlate with partitioning in Table 2), partitioned ethnicities are significantly more likely to experience civil conflict. Adding the cubic polynomial in latitude and longitude (column (2)) has little effect on the estimate. In column (3) we further condition on location characteristics augmenting the model with the distance from the centroid of each ethnic homeland to (i) the national border (ii) the nearest seacoast and (iii) the capital city. Overall, distance to the sea enters with a positive and significant estimate illustrating that there is less conflict in areas close to the sea. Distance to the capital enters also 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 national border enters with a negative sign; yet the coefficient is not significant at standard confidence levels. The partitioning indicator drops but continues to enter with a positive and significant estimate.

Column (4) includes a rich set of controls, reflecting ecology (land suitability for agriculture, elevation, malaria prevalence), early development (dummy variable that equals one if a major city was in the historical homeland in 1400), and natural resources (indicators that equal one if a diamond mine or an oil deposit is present). Accounting for these factors seems a priori important, because the cross-country literature documents significant correlations between many of these variables and various aspects of civil warfare. For example, Fearon and Laitin (2003) find that there is a higher likelihood of civil conflict in mountainous countries. Likewise, both cross-country works (e.g. Ross (2006, 2009)) and regional studies (e.g. Buhaug

and Rod (2006); Bellows and Miguel (2009)) show that conflict is higher in areas with diamond mines and petroleum fields. Moreover, in a recent study Cervellati, Sunde, and Valmori (2011) document a strong positive correlation between the disease environment and civil conflict across countries. The coefficient on the partitioning index remains unaffected. This is consistent with the findings in Table 2 showing that partitioning is largely uncorrelated with these characteristics.[28]

In column (5) we add ethnic-family fixed effects. Doing so, increases the precision of our estimate and if anything the coefficient on the partitioning index increases. In column (6) we add country fixed effects to control for differences in national institutions and policies, the identity and type of colonization, and other common-to-all ethnic regions factors.[29] The estimate on the partitioning indicator continues to be positive and significant at the 99%. The coefficient implies that partitioned ethnicities experience an increase of approximately 0.15 log points in the number of conflicts. This translates into a 16% increase in civil conflict activity ($\exp(0.15) - 1 = 0.16$) in areas where partitioned ethnicities reside.[30]

4.3 Civil Conflict Casualties

The descriptive statistics, reported in Table 1, show that the average death toll from armed conflict has been much higher in areas of partitioned ethnic groups. In Table 4 we formally examine the effect of partitioning on civil war casualties. Panel *A* reports negative binomial (NB) ML estimates where the dependent variable is the number of casualties standardized by land area (due to over-dispersion in the dependent variable, specification tests suggest that the negative binomial model is preferable to the Poisson model). Panel *B* reports analogous LS estimates where the dependent variable is the log of one plus the number of casualties per thousand of square kilometers.

Column (1) shows that casualties from armed conflict have been significantly higher in areas where partitioned ethnic groups reside. The coefficient on ethnic partitioning retains its economic and statistical significance when we include the third order polynomial in latitude and longitude in column (2). The estimate of the negative binomial specification implies that casualties are approximately 45% higher ($\exp(0.3805) - 1 = 0.46$) in the historical homeland of partitioned ethnicities. The coefficient drops moderately in column (3) when we control for the distance measures to the coast, the capital city, and the national border. Yet the coefficient is

[28] In all specifications the natural resource measures enter with a positive sign and are significant in most specifications. There is also some evidence that civil conflict is higher in more mountainous regions.

[29] When we add country fixed effects we lose variation from Rwanda, Swaziland, Burundi, and the Comoros, since in these countries Murdock maps a single ethnic group.

[30] Ordered probit estimation yields similar results. The coefficient on the $SPLIT$ index is highly significant (t-stat higher than 4 in all permutations).

more than two standard errors larger than zero both in the NB-ML and the LS specifications.

In column (4) we condition on the rich set of geographic controls. As these variables do not correlate with partitioning (see Table 2), their inclusion has limited impact on the estimate of the $SPLIT$ index. In column (5) we include a vector of ethnic-family constants, while in column (6) we condition on both ethnic-family and country fixed effects. The coefficient on $SPLIT$ remains intact. The NB-ML estimate in the full specification implies that civil casualties are 50% higher in the homelands of partitioned ethnicities ($\exp(0.4113) - 1 = 0.50$). As a robustness check, in columns (7) and (8) we report double fixed effects estimates using the low estimate and the high estimate of conflict casualties, respectively. The coefficient on $SPLIT$ is positive and highly significant.

4.4 Civil Conflict Duration

In Table 5 we examine the effect of partitioning on the duration of civil conflict. Panel A reports negative binomial (NB) ML estimates while Panel B reports LS estimates where the dependent variable is the log of one plus war duration. The specifications in column (1) suggest that the duration of civil conflict is significantly higher in regions where partitioned ethnicities reside. The estimate on $SPLIT$ is unaffected when we include in column (2) the third order polynomial in latitude and longitude. The coefficient remains unaffected also when we condition on the distance to the national border, the distance to the capital city, and the distance to the coast (in (3)). In column (4) we control for natural resources, ecological features and early development. The coefficient on the partitioning index remains stable. In column (5) we add a vector of ethnic-family fixed effects. The coefficient on the binary partitioning index retains its statistical and economic significance. Column (6) reports the most restrictive specification, where on the top of the controls of the previous specifications we add country fixed factors. The results suggest that civil wars last 25% longer in areas of partitioned ethnic groups, as compared to regions where non-partitioned, non-border groups reside ($\exp(0.23) - 1 = 0.256$). Our estimates are in line with the cross-country results of Fearon (2004) and Fearon and Laitin (2010), showing that wars involving land conflict between a peripheral ethnic minority and state-supported migrants of a dominant ethnic group are on average quite long lived (see also Wimmer, *et al.* (2009)).

5 Further Evidence

5.1 Spillovers

5.1.1 Issues and Approach

The estimates in Tables 3−5 do not take into account the possibility that there may be spillovers (externalities) from the historical homelands of partitioned ethnic groups (the "treatment" group) to neighboring areas where non-split ethnicities reside (the "control" group). Spillovers may emerge for numerous reasons. First, the battleground between a partitioned ethnic group and the central government (or another ethnicity supported by the state) might extend beyond the historical homeland of the partitioned ethnicity. For example, the battles between the Somali and Ethiopian tribes have not taken place solely across the border (where partitioned ethnicities reside). Second, in many cases the conflict in the homeland of the partitioned ethnicity leads to displacement and refugee flows to nearby areas, which in turn may spur conflict (e.g. Salehyan and Gleditsch (2006) and Blattman and Miguel (2010)). For example, the discriminatory policies of Mobutu Sese Seko against the Alur that lead to their displacement from Zaire, affecting adjacent regions both in Zaire and in Uganda. Third, ethnic groups nearby partitioned tribes are usually in the same ethnolinguistic family as the partitioned ethnicity, so they might get involved in the conflict to support their peers. For example, many Ugandan tribes assisted the ethnically similar Southern Sudanese troops in their long fight against the government troops of the ethnically and religiously heterogeneous Northern Sudan. Fourth, we have introduced some spatial correlation with the way the dependent variable is constructed (see Section 2).

To the extent that spillovers exist, the magnitudes of the estimates in Tables $3-5$ are lower bounds of the true effects, because of the non-accounted externalities. In other words, if ethnic conflict due to partitioning spreads to areas where non-split ethnic groups reside, then civil conflict will be higher in the control group; as such the difference between the two groups, reflected in the estimate of the binary partitioning index, will be an under-estimate of the true effect (see Miguel and Kremer (2004) for an analogous discussion of underestimating the impact of deworming due to spatial externalities). In an attempt to uncover and measure these potential externalities we modify the general empirical model as follows:

$$y_i^{f,c} = a_0 + \gamma SPLIT_i + \delta ADJ_i + \lambda SPLIT\_ADJ_i + X_i'\Phi + f(LOCUS_i) + a_c + a_f + \varepsilon_i^{f,c}. \quad (3)$$

Compared to the general econometric model in (2), we now amend the main specification adding: (i) the number of partitioned ethnic groups adjacent to ethnicity i, $SPLIT\_ADJ_i$;

(*ii*) a variable that records the number of ethnic groups adjacent to ethnicity i, ADJ_i. The latter is also a measure of local ethnic diversity [31]

5.1.2 Results

Table 6 reports results with all four civil conflict measures. Odd-numbered columns present the cross-sectional estimates with region constants. Even-numbered columns report specifications including a vector of ethnic-family and country fixed effects.

Civil Conflict Incidence Columns (1)-(2) report Poisson ML estimates using as the dependent variable the number of civil conflict incidents. Columns (3)-(4) report Poisson ML estimates with the number of civil conflict locations in the LHS. Across all permutations, the coefficient on $SPLIT$ that identifies the direct effect of partitioning is positive and highly significant. In line with our previous results on the local effect of partitioning on civil conflict incidence (in Table 3), the estimates suggest that there is a 20% higher likelihood that civil conflict takes place in the historical homeland of partitioned ethnic groups. The variable reflecting the number of adjacent ethnic groups (ADJ_i), that captures the effect of local ethnic heterogeneity, enters with an unstable and imprecise estimate. There is no evidence that the likelihood, duration and casualties of civil conflict increase with the number of nearby ethnic groups.[32] Nevertheless, the variable denoting the number of adjacent partitioned ethnic groups ($SPLIT\_ADJ_i$) that captures externalities also enters with a positive and highly significant coefficient. This implies that civil conflict is more likely to occur in regions that are neighboring the historical homeland of partitioned ethnic groups. The coefficient implies that an ethnic group residing adjacent to a partitioned ethnic homeland is on average 5% more likely to experience civil conflict, ceteris paribus. Interestingly, this estimate is quantitatively similar to the recent cross-country study of Bosker and de Ree (2010) who show that ethnic links to a neighbor country increase the probability of an outbreak of ethnic civil war at home by 6 percentage points.

Civil Conflict Casualties In columns (5)-(6) we study the direct and the indirect effect of partitioning on civil war casualties. The NB-ML specification in (5) shows that war casualties are much higher not only in the area where partitioned ethnicities reside but also in ethnic regions adjacent to the homelands of groups split by the national border. When we add ethnic-family fixed effects and country fixed effects in (6) the coefficient on the $SPLIT\_ADJ_i$

[31] This modeling of spillovers is similar to Kremer and Miguel (2004). In the previous version of the paper we also reported estimates from a general spatial lag model finding similar results.

[32] Using administrative regional data, Cunningham and Weidmann (2010) provide evidence that local ethnic heterogeneity is increasing conflict.

variable that captures externalities drops and turns statistically insignificant, while $SPLIT$ continues to enter with a highly significant estimate.

Civil Conflict Duration Columns (7)-(8) report NB-ML estimates with the number of years of civil conflict as the dependent variable. In line with our evidence in Table 5, the coefficient on the $SPLIT$ index is positive and more than two standard errors larger than zero. The estimate implies that the duration of civil wars is 30% higher in the homeland of partitioned ethnic groups, as compared to regions where non-partitioned groups reside ($\exp(0.28) - 1 = 0.33$). Moreover, $SPLIT\_ADJ_i$ also enters with a positive and highly significant coefficient, implying that civil conflict lasts longer in areas that are nearby the ethnic homelands of partitioned ethnic groups.

Summary Overall the results in Table 6 point out that if anything the estimates in Tables $2-5$ understate the causal effect of ethnic partitioning on civil conflict, as there are significant regional spillovers. Table 6 shows that the deleterious consequences of partitioning are not limited to the homeland of split ethnic groups. Tribal areas adjacent to partitioned groups also experience more civil wars which tend to last longer.[33]

5.2 Results in Border Areas

While in all specifications in Table 6 we account for factors that vary smoothly spatially by including the flexible third order polynomial in latitude and longitude, as well as the distance of the centroid of each homeland to the national border, the coastline, and the capital city, one may still worry that our estimates reflect an overall border effect. Thus, we repeated estimation focusing solely on ethnic groups close to the national border.

Table 7 reports specifications otherwise identical to Table 6 except for limiting the sample to groups close to the national border. We classify ethnicities as being close to the national border when the distance to the border is below the median value of 103 km (see Table 1 - Panel A). The estimate on the binary partitioning index that captures the direct local effect

[33] In the previous version of the paper we included in the specification the total number of casualties per thousand of square kilometers in each country and in each ethnic-family. This allows us to account for spillovers of civil conflict within countries and ethnic families. We found that a 10% increase in civil war casualties in the country where an ethnicity resides (excluding the casualties of the respective ethnicity) is associated with a 2.5% increase in casualties in the homeland of this ethnic group. Likewise a 10% increase in the number of casualties in areas of the same ethnic cluster (again excluding the casualties of the respective ethnicity) is associated with 1.8% increase in casualties for each ethnic group. Similarly, we augmented the specification with the average duration of civil conflict in all other groups within each country and within each ethnic family (netting out the duration in each ethnicity's own location). The estimates suggest that a 10% increase in the duration of armed civil conflict in the same country and ethnic-family is associated with approximately 5% increase in civil war duration locally.

of ethnic partitioning is positive and significant with all four measures of civil conflict. The coefficient's magnitude is quite similar to the (more efficient) estimates in the full sample (Table 6). Moreover, the $SPLIT\_ADJ_i$ index that captures spillovers also enters with a positive coefficient, that is in most models statistically significant. This suggests that across border areas civil conflict is not only concentrated in the historical homeland of partitioned ethnic groups but also spreads across those adjacent to split ethnicities.

5.3 Ethnic Partitioning and Natural Resources

It has been argued that partitioning may spur conflict by interacting with natural resources. This is because ethnicities with significant natural resources may decide to secede or demand autonomy since they depend less on the central government (or/and the other ethnic groups) for transfers. Moreover, the central government has a higher incentive to suffocate ethnicities with significant natural resources so as to get control of the diamond mines or the petroleum fields. In Table 8 we explore in detail this channel augmenting the specification (in (3)) with: (i) a natural resource dummy variable that takes on the value of one if a diamond mine or an petroleum field is present in the historical homeland of ethnic group i and zero otherwise; and (ii) an indicator identifying partitioned ethnic groups that have in their historical homeland diamond mines or petroleum fields. In all specifications we include ethnic-cluster fixed effects, country fixed effects, and the other set of control variables (e.g. the third order polynomial in latitude and longitude, distance to capital city and the seacoast).

In line with our results so far, the $SPLIT$ index that captures the local effects of partitioning enters with a positive and significant coefficient in all specifications. The $SPLIT\_ADJ_i$ index that captures spillovers from partitioned ethnic homelands to adjacent regions, also enters with a positive and usually statistically significant coefficient. On the contrary, the natural resource indicator enters with a positive though insignificant estimate.[34]

The variable that identifies partitioned ethnic homelands with diamond mines or petroleum fields enters with a positive and highly significant in columns (1) and (2). The estimate implies that there is approximately a 25% ($\exp(0.20) - 1 = 0.23$) higher likelihood that a civil conflict will take place in areas of partitioned ethnicities rich in natural resources. The estimate on the interaction between natural resources and partitioning is positive also in the specifications with casualties and duration; yet the coefficient's standard errors are large and thus the estimate is less precisely estimated. Overall, the results in Table 8, though not definite, point

[34] In the cross-sectional models (not reported) the natural resource index enters with a positive and significant, estimate suggesting a higher propensity for civil conflict in natural resource rich areas. Yet when we add country fixed effects, the estimate drops considerably and becomes statistically indistinguishable from zero. This implies that natural resources abundance does not correlate with civil conflict within national territories.

to the direction that natural resource abundance is particularly harmful when it interacts with ethnic partitioning.

5.4 Sensitivity Analysis

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

Continuous Measure of Partitioning First, we repeated estimation using the continuous measure of partitioning. Appendix Table 4 reports results associating the four measures of civil conflict with the continuous index of partitioning ($FRAC$) that reflects the likelihood that a historical ethnic homeland falls into more than one contemporary state. In line with our results so far, the ethnic partitioning measure always 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, more casualties, and prolonged warfare. Moreover, in all permutations the variable that records the number of adjacent partitioned groups enters with a positive estimate, suggesting the presence of sizable externalities.

Excluding Ethnic Areas where Capital Cities Fall Second, since some civil conflicts have taken place in the capital cities, we repeated estimation excluding from the analysis ethnic homelands where capitals fall. Appendix Table 5 reports the results. Across all model permutations the coefficient on the partitioning indicator is positive and significant; moreover, in all specifications the estimate on the $SPLIT\_ADJ_i$ measure is positive and statistically significant at standard confidence levels, illustrating the presence of sizable spillovers.

Excluding Regions Third, we estimated models dropping each time ethnic areas from one of the five African regions (using Nunn's (2008) classification). Appendix Table 6 reports the results. In Panel A we report estimates excluding North Africa; this is useful because Europeans had contacts with the northern part of the continent since the ancient times. In Panel B we drop Southern African countries to account for the fact that during the apartheid period, South Africa intervened in many civil conflicts in nearby countries (Namibia, Angola, and Zimbabwe). In Panel C 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 Panel D and Panel E we exclude from the estimation 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 and the Democratic Republic of Congo. The results in Appendix Table 6 show that the strong positive effect of ethnic partitioning on civil

conflict is not driven by a particular region. In all specifications the partitioning index enters with a highly significant estimate, similar in magnitude to the (more efficient) estimates in Table 6. Moreover, across all thirty specifications reported in Appendix Table 6, the coefficient on the number of adjacent partitioned groups enters with a positive (and in most models statistically significant) coefficient, suggesting that civil conflict spreads from the homelands of partitioned ethnicities to nearby regions.

6 Conclusion

We have examined 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 caused the partitioning of several ethnicities across the newly created African states.

In the first part of our paper we formally establish the artificial nature of African political boundaries. Using regional information from Murdock (1959) on the spatial distribution of ethnicities at the time of colonization, we estimate probabilistic models associating ethnic partitioning with various geographic, ecological, and ethnic-specific pre-colonial characteristics. With the sole exceptions of the size of the historical homeland and water bodies, there are no other significant differences between partitioned and non-partitioned ethnicities. These results support the consensus of the African historiography on the accidental and artificial drawing of colonial and consequently national borders.

Second, we examine in a quasi-natural experimental setting the effect of ethnic partitioning on various aspects of civil conflict (incidence, location, casualties, duration), as this has been hypothesized to be the major effect of the scramble for Africa on development. In contrast to most works on the long-run effects of colonization on civil conflict and development, our analysis is based on regional data spanning 834 ethnic areas across Africa. Our ethnic-homeland level dataset allows us to control at a very fine level for natural resources, geography, the disease environment, early development, and other factors that cross-country works have identified as significant correlates of civil warfare. We also include country fixed effects and ethnic-family fixed effects to account for differences in both national policies and broad cultural/economic/social characteristics across ethnic families. We find that partitioned ethnicities have suffered systematically more from civil conflict compared to groups that have not been directly affected by the improper border design.

Third, we search for spillovers/externalities. Our estimates show that civil conflict is not only concentrated in the historical homeland of partitioned ethnic groups, but groups adjacent to split ethnicities are also more likely to experience conflicts which last longer and result in

more casualties.

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

7 Data Appendix

Partitioning Index (SPLIT): Indicator variable that equals 1 if at least 10% of the historical homeland of an ethnic group is partitioned into different countries. *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 A 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.*

Number of Civil Conflicts: See text. Section 2.2.1. *Source: Uppsala Conflict Data Program (UCDP)/International Peace Research Institute, Oslo (PRIO) Armed Conflict Dataset, Version 4-2006; and Raleigh, Cunningham, Wilhelmsen, and Gleditsch (2006).*

Number of Civil War Conflict Zones: See text. Section 2.2.1. *Source: Uppsala Conflict Data Program (UCDP)/International Peace Research Institute, Oslo (PRIO) Armed Conflict Dataset, Version 4-2006; and Raleigh, Cunningham, Wilhelmsen, and Gleditsch (2006).*

Civil Conflict Casualties (Best/High/Low Estimate): See text. Section 2.2.2. *Source: Uppsala Conflict Data Program (UCDP)/International Peace Research Institute, Oslo (PRIO) Armed Conflict Dataset, Version 4-2006; and Lacina and Gleditsch (2005).*

Civil Conflict Duration: See text. Section 2.2.3. *Source: Uppsala Conflict Data Program (UCDP)/International Peace Research Institute, Oslo (PRIO) Armed Conflict Dataset, Version 4-2006; and Raleigh, Cunningham, Wilhelmsen, and Gleditsch (2006).*

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

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

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

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

Land Suitability for Agriculture: Average land quality for cultivation within the area of each ethnic-country observation. The index is the product of two components reflecting

the climatic and soil suitability for cultivation. *Source: Michalopoulos (2011); Original Source: Atlas of the Biosphere. Available at http://www.sage.wisc.edu/iamdata/grid_ data_ sel.php.*

Malaria Stability Index: The index takes into account the prevalence and type of mosquitoes indigenous to a region, their human biting rate, their daily survival rate, and their incubation period. The index has been constructed for 0.5 degree by 0.5 degree grid-cells. We use the average value for each ethnic homeland. *Source: Kiszewski, Mellinger, Spielman, Malaney, Sachs, and Sachs (2004)*

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

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

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

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

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

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

Longitude: Longitude of the centroid of each ethnic group. *Source: Constructed using ArchGIS Software.*

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

Slavery: Number of persons of each ethnic group that were shipped during the trans-Atlantic and Indian Ocean slave trades. Following Nunn (2008) in the regressions we use the

log of one plus the number of slaves per 1000 of square kilometers. *Source: Nunn (2008) and Nunn and Watchkon (2011).*

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

Distance to Explorer's Routes: The geodesic distance of the centroid of each group to the nearest route of the principal European explorers. *Source: The "Century Atlas, Africa" digitized by Nunn (2010).*

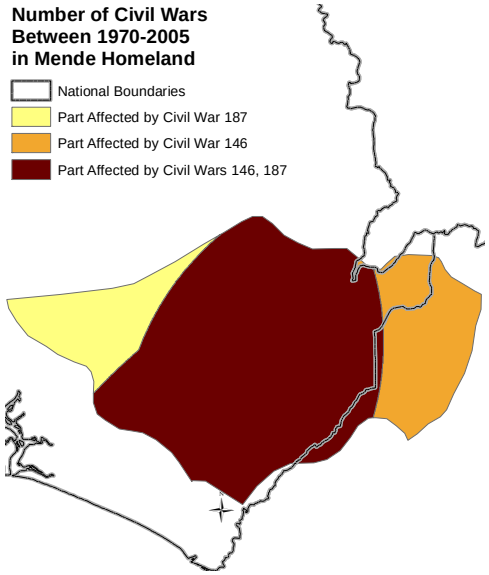
Political Centralization: The binary index is constructed using Murdock's Jurisdictional Hierarchy beyond Local Community $0-4$ index that indicates the number of jurisdictional levels (political complexity) in each society above the local level**.** The political centralization index takes the value 0 if the Jurisdictional Hierarchy beyond Local Community variable equals 0 or 1 (when the society is classified as either stateless or forming a small chiefdom). The index takes on the value 1 if the Jurisdictional Hierarchy beyond Local Community variable equals 2, 3, and 4 (when the society is classified as being part of large paramount chiefdom or a large state). This aggregation follows Gennaioli and Rainer (2006, 2007). *Source: Murdock (1967).*

Class Stratification: Ordered variable ranging from 0 to 4 quantifying "*the degree of class differentiation, excluding purely political and religious statuses*". A zero score indicates "*absence of significant class distinctions among freemen, ignoring variations in individual repute achieved through skill, valor, piety, or wisdom.*" A score of 1 indicates "*the presence of wealth distinctions, based on possession or distribution of property, which however have not crystallized into distinct and hereditary social classes.*" A score of 2 indicates "*elite stratification in which an elite class derives its superior status from control over scarce resources, particularly land, and is thereby differentiated from a propertyless proletariat or serf class*". A score of 3 indicates a "*dual stratification into a hereditary aristocracy and a lower class of ordinary commoners or freemen, where traditionally ascribed noble status is at least as decisive as control over scarce resources.* A score of 4 indicates "*complex stratification into social classes correlated in large measure with extensive differentiation of occupational statuses.*" *Source: Murdock (1967); variable code in the Entholinguistic Atlas v67.*

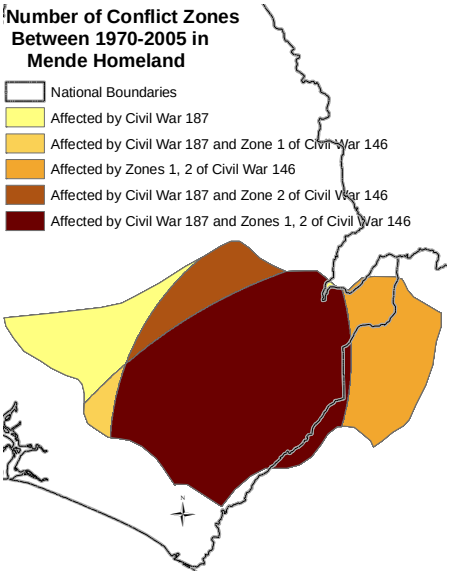
Dependence on Agriculture: $0-10$ scale index reflecting the dependence of each ethnicity in agriculture at the time of colonization. *Source: Murdock (1967).*

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

7.1 Appendix Figures



Appendix Figure 1a



Appendix Figure 1b

References

Acemoglu, D., M. A. Bautista, P. Querubin, and J. A. Robinson (2008): "Economic and Political Inequality in Development: The Case of Cundinamarca, Colombia," in *Institutions and Economic Performance*, ed. by E. Helpman. Harvard University Press.

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

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

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

Aghion, P., A. Alesina, and F. Trebbi (2004): "Endogenous Political Institutions," *Quarterly Journal of Economics*, 119(2), 565–612.

Alesina, A., R. Baqir, and W. Easterly (1999): "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 114(4), 1243–1284.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003): "Fractionalization," *Journal of Economic Growth*, 8, 155–194.

Alesina, A., W. Easterly, and J. Matuszeski (2011): "Artificial States," Journal of the European Economic Association, forthcoming.

Alesina, A., and E. L. Ferrara (2005): "Fractionalization," *Journal of Economic Literature*, 43, 762–800.

Alesina, A., and E. Spolaore (2003): *The Size of Nations*. MIT Press.

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

Arbesu, H. A. L. (2011): "Colonialism and Economic Development: Evidence from a Natural Experiment in Colonial Nigeria," mimeo MIT.

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

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

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

Bates, R. H. (1981): in *States and Markets in Africa*University of California Press, Berkeley, CA.

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

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

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 WP 8055*.

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

Cervellati, M., U. Sunde, and S. Valmori (2011): "Disease Environment and Civil Conflicts," IZA Discussion Paper No. 5614.

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

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

——— (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., A. Hoeffler, and M. Soderbom (2004): "On the Duration of Civil War," *Journal of Peace Research*, 41(3), 253–273.

Collier, P., and N. Sambanis (2005): *Understanding Civil War. Volume 1: Africa*. World Bank, Washington, DC.

Collier, P., and A. Venables (2008): "Trade and Economic Performance: Does Africa's Fragmentation Matter?," paper prepared for the ABCDE Conference.

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

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

Cunningham, K. G., and N. B. Weidmann (2010): "Shared Space: Ethnic Groups, State Accommodation, and Localized Conflict," *International Studies Quarterly*, 54, 1035–1054.

Dell, M. (2010): "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 78.

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

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

——— (2009): "The European Origins of Economic Development," Working Paper, mimeo Brown University.

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

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

Esteban, J., L. Mayoral, and D. Ray (2010): "Ethnicity and Conflict: An Empirical Investigation," mimeo NYU Department of Economics.

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

Fearon, J. (2004): "Why Do Some Civil Wars Last So Much Longer Than Others?," *Journal of Peace Research*, 41(3), 275–302.

Fearon, J., and D. Laitin (2003): "Ethnicity, Insurgency and Civil War," *American Political Science Review*, 97(1), 75–90.

——— (2010): "Sons of the Soil, Migrants, and Civil War," World Development, forthcoming.

Fenske, J. (2010): "Does Land Abundance Explain African Institutions," mimeo Oxford University.

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

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

Glaeser, E. L., R. L. Porta, F. L. de Silanes, and A. Shleifer (2004): "Do Institutions Cause Growth?," *Journal of Economic Growth*, 9(3), 271–303.

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

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

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

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.

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

Iyer, L. (2010): "Direct versus Indirect Colonial Rule in India: Long-term Consequences," *Review of Economics and Statistics*, 92(4), 693–713.

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.

La-Porta, R., F. L. de Silanes, A. Shleifer, and R. Vishny (1999): "The Quality of Government," *Journal of Law Economics and Organization*, 15(1), 222–279.

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

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

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

Michalopoulos, S., and E. Papaioannou (2010): "Divide and Rule or the Rule of the Divided? Evidence from Africa," CEPR Discussion Paper, CEPR DP No. 6414.

Miller, D., C. A.Cameron, and J. Gelbach (2011): "Robust Inference with Multi-Way Clustering," .

Montalvo, J. G., and M. Reynal-Querol (2005a): "Ethnic Diversity and Economic Development," *Journal of Development Economics*, 76, 293–323.

——— (2005b): "Ethnic Polarization, Potential Conflict and Civil War," *American Economic Review*, 95, 796–816.

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

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

Naritomi, J., R. R. Soares, and J. J. Assunção (2009): "Institutional Development and Colonial Heritage within Brazil," Working Paper, IZA DP No. 4276.

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

Nunn, N., and L. Wantchekon (2011): "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review, forthcoming*.

Petter, G. N., P. Wallensteen, M. Eriksson, M. Sollenberg, and H. Strand (2002): "Armed Conflict 1946-2001: A New Datase," *Journal of Peace Research*, 39(5), 615–637.

Raleigh, C., D. Cunningham, L. Wilhelmsen, and N. P. Gleditsch (2006): "Conflict Sites: 1946-2005. v2.0," Centre for the Study of Civil War, PRIO.

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

Salehyan, I., and K. S. Gleditsch (2006): "Refugees and the Spread of Civil War," *International Organization*, 60(02), 335–366.

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

Silva, J. S., S. Tenreyro, and F. Windmeijer (2010): "Is it Different for Zeros? Discriminating Between Models for Non-Negative Data with Many Zeros," CeMMAP working papers CWP20/10.

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

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

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

Table 1: Summary Statistics for Civil Conflict Measures

	Obs.	mean	st. dev.	min	p25	median	p75	max
Panel A: All Ethnicities								
Number of Civil Conflicts	834	1.14	0.95	0.00	0.00	1.00	2.00	5.00
Number of Civil War Conflict Zones	834	1.58	1.52	0.00	0.00	1.00	2.00	8.00
Civil Conflict Casualties (Best Estimate)	834	39.62	82.25	0.00	0.00	3.21	40.12	589.53
Civil Conflict Casualties (Low Estimate)	834	33.57	77.59	0.00	0.00	1.90	40.12	561.34
Civil Conflict Casualties (High Estimate)	834	45.23	91.47	0.00	0.00	4.32	40.68	639.37
Civil Conflict Duration	834	7.75	9.39	0.00	0.00	3.86	11.63	46.79
Panel B: Partitioned (*SPLIT=1*) Ethnicities								
Number of Civil Conflicts	231	1.41	1.03	0.00	1.00	1.00	2.00	5.00
Number of Civil War Conflict Zones	231	1.98	1.62	0.00	1.00	2.00	3.00	7.00
Civil Conflict Casualties (Best Estimate)	231	50.28	94.00	0.00	0.30	11.02	59.35	589.53
Civil Conflict Casualties (Low Estimate)	231	40.46	88.34	0.00	0.07	5.75	37.00	561.34
Civil Conflict Casualties (High Estimate)	231	61.01	109.77	0.00	0.49	13.35	69.58	639.37
Civil Conflict Duration	231	8.29	9.24	0.00	0.48	6.00	11.85	46.79
Panel C: Non-Partitioned (*SPLIT=0*) Ethnicities								
Number of Civil Conflicts	603	1.04	0.89	0.00	0.00	1.00	2.00	4.00
Number of Civil War Conflict Zones	603	1.42	1.45	0.00	0.00	1.00	2.00	8.00
Civil Conflict Casualties (Best Estimate)	603	35.53	76.97	0.00	0.00	2.59	40.12	551.87
Civil Conflict Casualties (Low Estimate)	603	30.93	72.96	0.00	0.00	1.47	40.12	535.48
Civil Conflict Casualties (High Estimate)	603	39.18	82.72	0.00	0.00	4.06	40.12	580.52
Civil Conflict Duration	603	7.54	9.44	0.00	0.00	3.00	11.53	42.74

Panel D: Test of Means & Medians

	Test of Means		Test of Medians	
	diff.	p-value	Pearson's chi^2	p-value
Number of Civil Conflicts	0.37	*0.00*	24.20	*0.00*
Number of Civil War Conflict Zones	0.56	*0.00*	27.34	*0.00*
Civil Conflict Casualties (Best Estimate)	14.74	*0.03*	6.52	*0.01*
Civil Conflict Casualties (Low Estimate)	9.53	*0.15*	3.74	*0.05*
Civil Conflict Casualties (High Estimate)	21.82	*0.01*	5.75	*0.02*
Civil Conflict Duration	0.75	*0.30*	3.74	*0.05*

The table reports descriptive statistics for all outcome variables employed in the empirical analysis. Panel A reports summary statistics in the full sample of ethnicities. Panel B reports summary statistics in the sample of partitioned ethnicities. Panel C reports summary statistics in the sample of non-partitioned ethnicities. Panel D reports a test of mean and median equality between the group of partitioned ethnicities and the group of non-partitioned ethnicities. The Data Appendix gives detailed variable definitions and data sources.

Table 2 - Are African Borders Artificial?

Panel A: Geographical, Ecological and Natural Resources Features

	SPLIT	FRAC	SPLIT	FRAC	SPLIT	FRAC	SPLIT	FRAC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Land Area under Water	0.3209***	0.0612***	0.3743***	0.0668***	0.3417***	0.0665***	0.3353***	0.0624***
	(0.0759)	(0.0125)	(0.0822)	(0.0122)	(0.0716)	(0.0122)	(0.0810)	(0.0134)
Land Area	0.0912	0.0154*	0.1184**	0.0187**	0.0979*	0.0172**	0.0745	0.0137*
	(0.0524)	(0.0080)	(0.0540)	(0.0082)	(0.0559)	(0.0084)	(0.0479)	(0.0071)
Elevation			-0.0628	-0.0181				
			(0.2532)	(0.0304)				
Suitability for Agriculture			0.5788**	0.0732				
			(0.3362)	(0.0464)				
Malaria Stability Index					0.1215	0.0284		
					(0.3204)	(0.0405)		
Distance to the Coast					-0.0001	0.0001		
					(0.0002)	(0.0000)		
Diamond Mine Indicator							0.1598	0.0176
							(0.1796)	(0.0289)
Oil Indicator							-0.004	0.0009
							(0.1709)	(0.0352)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.051	__	0.057	__	0.052	__	0.052	__
Adjusted R-squared	__	0.084	__	0.091	__	0.090	__	0.085
Observations	834	834	834	834	834	834	834	834

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 - Are African Borders Artificial?

Panel B: Historical Features

	SPLIT	FRAC	SPLIT	FRAC	SPLIT	FRAC	SPLIT	FRAC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Land Area under Water	0.3287***	0.0624***	0.3202***	0.0613***	0.3157***	0.0605***	0.3223***	0.0612***
	(0.0747)	(0.0123)	(0.0766)	(0.0124)	(0.0819)	(0.0130)	(0.0755)	(0.0125)
Land Area	0.0879	0.0148*	0.0903	0.0155*	0.0867	0.0148*	0.0981*	0.0152*
	(0.0511)	(0.0079)	(0.0519)	(0.0079)	(0.0522)	(0.0080)	(0.0566)	(0.0084)
Slave Exports	0.0227	0.0034						
	(0.0238)	(0.0036)						
Major City in 1400AD			0.0437	-0.0057				
			(0.1981)	(0.0351)				
Distance to Explorer's Routes					-0.0004	0.0001		
					(0.0003)	(0.0000)		
Population Density 1960							0.0213	-0.0006
							(0.0638)	(0.0079)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.053	__	0.051	__	0.054	__	0.051	__
Adjusted R-squared	__	0.086	__	0.084	__	0.087	__	0.084
Observations	834	834	834	834	834	834	834	834

Table 2- Panel B reports probit marginal effects (in odd-numbered columns) and OLS estimates (in even-numbered columns) associating ethnic partitioning with historical variables. In odd-numbered specifications, the dependent variable is an indicator that equals one when at least 10% of the historical ethnic homeland (as portrayed in Murdock's (1959) Ethnolinguistic map) falls to more than one contemporary country. In even-numbered columns, the dependent variable is a continuous index of ethnic partitioning that reflects the probability that a randomly chosen pixel of the historical homeland of an ethnic group falls into a different country. All specifications include a set of region fixed-effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 2 - Are African Borders Artificial?

Panel C: Pre-colonial Ethnic Features

	SPLIT	FRAC	SPLIT	FRAC	SPLIT	FRAC	SPLIT	FRAC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Land Area under Water	0.2452**	0.0465***	0.3029***	0.0575***	0.2803**	0.0542***	0.2546**	0.0501***
	(0.0976)	(0.0167)	(0.0924)	(0.0164)	(0.0952)	(0.0166)	(0.0943)	(0.0160)
Land Area	0.1672**	0.0262***	0.1659**	0.0235**	0.1721**	0.0259***	0.1662**	0.0251***
	(0.0637)	(0.0094)	(0.0677)	(0.0102)	(0.0598)	(0.0083)	(0.0586)	(0.0083)
Political Centralization	-0.1865	-0.0307						
	(0.1668)	(0.0210)						
Class Stratification			-0.0237	-0.0037				
			(0.0474)	(0.0072)				
Share of Agriculture					0.028	0.0041		
					(0.0328)	(0.0054)		
Animal Husbandry							-0.0014	0.0003
							(0.0350)	(0.0055)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
pseudo R-squared	0.069	__	0.079	__	0.068	__	0.067	__
adjusted R-squared	__	0.106	__	0.123	__	0.107	__	0.106
Observations	440	440	396	396	490	490	490	490

Table 2- Panel C reports probit marginal effects (in odd-numbered columns) and OLS estimates (in even-numbered columns) associating ethnic partitioning with pre-colonial ethnic variables (using data from Murdock (1967)). In odd-numbered specifications, the dependent variable is an indicator that equals one when at least 10% of the historical ethnic homeland (as portrayed in Murdock's (1959) Ethnolinguistic map) falls to more than one contemporary country. In even-numbered columns, the dependent variable is a continuous index of ethnic partitioning that reflects the probability that a randomly chosen pixel of the historical homeland of an ethnic group falls into a different country. All specifications include a set of region fixed-effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 3: Partitioning and the Incidence of Civil Conflict (Poisson ML Estimates)

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A - Dependent Variable: Number of Civil Conflicts					
SPLIT - Partitioning	0.2680***	0.2660***	0.1687*	0.1656*	0.2039***	0.1472***
	(0.0593)	(0.0596)	(0.0914)	(0.0916)	(0.0525)	(0.0498)
Distance to the Border			-0.0007	-0.0008	-0.0002	-0.0001
			(0.0006)	(0.0006)	(0.0003)	(0.0003)
Log Likelihood	-1036.88	-997.67	-970.54	-965.08	-855.59	-819.32
	Panel B - Dependent Variable: Number of Conflict Zones					
SPLIT - Partitioning	0.3063***	0.3061***	0.1691*	0.1864**	0.2266***	0.1512***
	(0.0730)	(0.0689)	(0.0961)	(0.0881)	(0.0580)	(0.0516)
Distance to the Border			-0.0011	-0.0011	-0.0002	-0.0001
			(0.0007)	(0.0007)	(0.0004)	(0.0004)
Log Likelihood	-1281.79	-1229.68	-1176.51	-1165.10	-969.96	-918.956
Controls & Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial Latitude & Longitude	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes
Additional Controls	No	No	No	Yes	Yes	Yes
Ethnic Family FE	No	No	No	No	Yes	Yes
Country FE	No	No	No	No	No	Yes
Observations	834	834	834	834	834	830

The table reports Poisson ML estimates associating civil war incidence with ethnic partitioning. In Panel A the dependent variable is the number of civil wars that have taken place in the historical homeland of an ethnic group between 1970 and 2005. In Panel B the dependent variable is the number of conflict zones associated with 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. All specifications include a set of region fixed effects (constants not reported), log land area, log land area under water (lakes, rivers, and other streams), and log population density around independence (in 1960).

Columns (2)-(6) include a cubic polynomial in latitude and longitude of the centroid of each ethnic group. Columns (3)-(6) include the distance of each ethnic homeland to the national border, the distance to the capital city, and the distance to the closest sea coast (Location Controls). Columns (4)-(6) include a rich set of control variables, namely 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. Columns (5) and (6) include a set of ethnic-family fixed effects (constants not reported). Column (6) includes a set of country fixed effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 4: Partitioning and Civil Conflict Casualties

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Negative Binomial ML				
SPLIT - Partitioning	0.4429***	0.3805***	0.2532**	0.3141**	0.4381***	0.4113**	0.3968**	0.4233**
	(0.1915)	(0.1093)	(0.1178)	(0.1399)	(0.1516)	(0.1969)	(0.1747)	(0.1706)
Distance to the Border			-0.0012	-0.0013	0.0000	-0.0006	-0.0005	-0.0006
			(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.0008)	(0.0009)
Log Likelihood	-3196.85	-3004.05	-2992.58	-2960.79	-2699.69	-2535.66	-2325.29	-2645.20
				Panel B: Log Linear OLS				
SPLIT - Partitioning	0.4854**	0.4336***	0.3723**	0.4229***	0.4213***	0.3149*	0.2850*	0.3372*
	(0.2120)	(0.1574)	(0.1511)	(0.1497)	(0.1485)	(0.1701)	(0.1526)	(0.1735)
Distance to the Border			-0.0004	-0.0004	0.0003	-0.0003	-0.0002	-0.0003
			(0.0010)	(0.0012)	(0.0009)	(0.0008)	(0.0007)	(0.0008)
Adjusted R-squared	0.214	0.413	0.486	0.513	0.757	0.839	0.850	0.831
Controls & Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial Lat. & Long.	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	No	No	No	Yes	Yes	Yes	Yes	Yes
Ethnic Family FE	No	No	No	No	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	No	Yes	Yes	Yes
Observations	834	834	834	834	834	830	830	830

The table reports Negative Binomial Maximum Likelihood (ML) estimates (in Panel A) and OLS estimates (in Panel B), associating civil war casualties with ethnic partitioning. The dependent variable in Panel A is the number of casualties per thousand of square kilometers from civil conflicts that have taken place in the historical homeland of an ethnic group between 1970 and 2005. The dependent variable in Panel B is the log of one plus the number of casualties per thousand of square kilometers. In columns (1)-(6) we use the best estimate of civil conflict casualties, while in column (7) and (8) we use the high and the low war casualties estimates, respectively. 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. All specifications include a set of region fixed effects (constants not reported), log land area, log land area under water (lakes, rivers, and other streams), and log population density around independence (in 1960).

Columns (2)-(6) include a cubic polynomial in latitude and longitude of the centroid of each ethnic group. Columns (3)-(6) include the distance of each ethnic homeland to the national border, the distance to the capital city, and the distance to the closest sea coast (Location Controls). Columns (4)-(6) include a rich set of control variables, namely 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 mine indicator. Columns (5) and (6) include a set of ethnic-family fixed effects (constants not reported). Column (6) includes a set of country fixed effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 5: Partitioning and Civil Conflict Duration

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Negative Binomial ML			
SPLIT - Partitioning	0.2907**	0.3474**	0.3236**	0.3202**	0.2750**	0.2333*
	(0.1449)	(0.1302)	(0.1289)	(0.1320)	(0.1115)	(0.1321)
Distance to the Border			-0.0003	-0.0003	0.0002	-0.0003
			(0.0010)	(0.0010)	(0.0007)	(0.0006)
Log Likelihood	-2377.90	-2281.71	-2252.00	-2236.48	-1912.89	-1752.21
			Panel B: Log Linear OLS			
SPLIT - Partitioning	0.2484**	0.2828***	0.2638**	0.2542**	0.2516**	0.2317**
	(0.1148)	(0.1041)	(0.1110)	(0.1124)	(0.0988)	(0.1134)
Distance to the Border			-0.0001	-0.0002	0.0002	-0.0001
			(0.0007)	(0.0008)	(0.0006)	(0.0006)
adjusted R-squared	0.231	0.391	0.483	0.513	0.774	0.846
Controls & Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial Latitude & Longitude	No	Yes	Yes	Yes	Yes	Yes
Location	No	No	Yes	Yes	Yes	Yes
Additional Controls	No	No	No	Yes	Yes	Yes
Ethnic Family FE	No	No	No	No	Yes	Yes
Country FE	No	No	No	No	No	Yes
Observations	834	834	834	834	834	830

The table reports Negative Binomial Maximum Likelihood (ML) estimates (in Panel A) and OLS estimates (in Panel B), associating civil war duration with ethnic partitioning. The dependent variable in Panel A is the total number of years that an ethnic homeland has been under civil conflict over the period 1970-2005. The dependent variable in Panel B is the log of one plus the number of number of years under conflict. 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. All specifications include a set of region fixed effects (constants not reported), log land area, log land area under water (lakes, rivers, and other streams), and log population density around independence (in 1960).

Columns (2)-(6) include a cubic polynomial in latitude and longitude of the centroid of each ethnic group. Columns (3)-(6) include the distance of each ethnic homeland to the national border, the distance to the capital city, and the distance to the closest sea coast (Location Controls). Columns (4)-(6) include a rich set of control variables, namely 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 mine indicator. Columns (5) and (6) include a set of ethnic-family fixed effects (constants not reported). Column (6) includes a set of country fixed effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 6: Partitioning and Civil Conflict: Accounting for Spillovers

	Incidents		Locations		Casualties		Duration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT - Partitioning	0.1788*	0.1832***	0.2095**	0.2017***	0.3385**	0.4510***	0.3317**	0.2817*
	(0.0934)	(0.0552)	(0.0895)	(0.0593)	(0.1385)	(0.2133)	(0.1415)	(0.1643)
ADJ - Neighbors	-0.0141	0.0101	-0.0321**	0.0222	-0.0749	-0.0447*	-0.0468	-0.0051
	(0.0153)	(0.0116)	(0.0147)	(0.0131)	(0.0514)	(0.0288)	(0.0315)	(0.0148)
SPLIT-ADJ - Partitioned Neighbors	0.0562**	0.0483***	0.0843***	0.0655***	0.1638**	0.0733	0.1088*	0.0777*
	(0.0221)	(0.0140)	(0.0275)	(0.0163)	(0.0627)	(0.0570)	(0.0614)	(0.0432)
Distance to the Border	-0.0005	0.0001	-0.0006	0.0002	-0.0006	-0.0004	0.0002	0.0001
	(0.0007)	(0.0004)	(0.0007)	(0.0004)	(0.0009)	(0.0009)	(0.0010)	(0.0010)
Log Likelihood	-962.96	-817.79	-1157.94	-914.44	-2955.76	-2533.71	-2232.08	-1746.04
Region FE	Yes	No	Yes	No	Yes	No	Yes	No
Polynomial Latitude & Longitude	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rich Conditioning Set	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Family FE	No	Yes	No	Yes	No	Yes	No	Yes
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	834	830	834	830	834	830	834	830

The table report Poisson (in columns (1)-(4)) and Negative Binomial (in columns (5)-(8)) 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. The dependent variable in columns (5)-(6) is the number of casualties per thousand of square kilometers from civil conflicts that have taken place in the historical homeland of an ethnic group between 1970 and 2005. The dependent variable in columns (7)-(8) is the total number of years that an ethnic homeland has been under civil conflict over 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. ADJ (Neighbors) is the number of ethnic groups adjacent to the ethnic homeland. SPLIT-ADJ is the number of adjacent ethnic groups that are partitioned by the national border.

Odd-numbered specifications include a set of region fixed effects (constants not reported). Even-numbered specifications include a set of ethnic-family fixed effects (constants not reported) and a set of country fixed effects (constants not reported). All specifications include a cubic polynomial in latitude and longitude of the centroid of each ethnic group. Moreover, 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 density around independence (in 1960), the distance of each ethnic homeland to the national border, the distance to the capital city, the distance to the closest sea coast, land suitability for agriculture, mean elevation, a malaria stability index, an indicator of early development that equals one when a major city was in the ethnicity's historical homeland in 1400, an oil indicator and a diamond indicator. The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 7: Partitioning and Civil Conflict in Border Areas

	Incidents		Locations		Casualties		Duration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT - Partitioning	0.1833***	0.1454***	0.2404***	0.1845***	0.4743***	0.2959*	0.3622***	0.1744
	(0.0475)	(0.0469)	(0.0671)	(0.0473)	(0.1436)	(0.1807)	(0.1338)	(0.1144)
ADJ - Neighbors	-0.0347	0.0021	-0.0652***	0.0105	-0.0687	-0.0810*	-0.0806**	-0.0115
	(0.0251)	(0.0182)	(0.0224)	(0.0167)	(0.0619)	(0.0455)	(0.0387)	(0.0166)
SPLIT-ADJ - Partitioned Neighbors	0.0860**	0.0541***	0.1278***	0.0751***	0.2318***	0.0807	0.1811***	0.1004**
	(0.0407)	(0.0173)	(0.0425)	(0.0260)	(0.0768)	(0.0997)	(0.0786)	(0.0467)
Log Likelihood	-485.06	-407.00	-577.83	-451.80	-1462.82	-1250.86	-1068.82	-791.84
Region FE	Yes	No	Yes	No	Yes	No	Yes	No
Polynomial Latitude & Longitude	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rich Conditioning Set	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Family FE	No	Yes	No	Yes	No	Yes	No	Yes
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	417	417	417	417	417	417	417	417

Table 7 - Notes

The table reports Poisson (in columns (1)-(4)) and Negative Binomial (in columns (5)-(8)) Maximum Likelihood (ML) estimates associating various measures of civil war with ethnic partitioning in areas close to the national border (using as a cutoff the median distance to the national border). 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. The dependent variable in columns (5)-(6) is the number of casualties per thousand of square kilometers from civil conflicts that have taken place in the historical homeland of an ethnic group between 1970 and 2005. The dependent variable in columns (7)-(8) is the total number of years that an ethnic homeland has been under civil conflict over 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. ADJ (Neighbors) is the number of ethnic groups adjacent to the ethnic homeland. SPLIT-ADJ is the number of adjacent ethnic groups that are partitioned by the national border.

Odd-numbered specifications include a set of region fixed effects (constants not reported). Even-numbered specifications include a set of ethnic-family fixed effects (constants not reported) and a set of country fixed effects (constants not reported). All specifications include a cubic polynomial in latitude and longitude of the centroid of each ethnic group. Moreover, 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 density around independence (in 1960), the distance of each ethnic homeland to the capital city, the distance to the closest sea coast, land suitability for agriculture, mean elevation, a malaria stability index, an indicator of early development that equals one when a major city was in the ethnicity's historical homeland in 1400, an oil indicator and a diamond indicator. The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 8: Partitioning, Natural Resources, and Civil Conflict

	Incidents		Locations		Casualties		Duration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT - Partitioning	0.1633***	0.1624***	0.2142***	0.1677***	0.2498*	0.4371**	0.2732**	0.2563**
	(0.0510)	(0.0481)	(0.0617)	(0.0520)	(0.1359)	(0.2204)	(0.1344)	(0.1414)
NR - Natural Resources Indicator	0.0390	-0.0446	0.1066	-0.0323	0.4740*	-0.0435	0.2587**	0.0007
	(0.0668)	(0.0557)	(0.0770)	(0.0527)	(0.2757)	(0.1613)	(0.0980)	(0.0786)
SPLIT-NR - Partitioned & Natural Resources	0.2583**	0.1902**	0.1537	0.1590	0.1820	0.3009	0.0592	0.0968
	(0.1147)	(0.0958)	(0.1070)	(0.1047)	(0.3050)	(0.2079)	(0.1797)	(0.1314)
ADJ - Neighbors	-0.0255*	0.0011	-0.0408***	0.0112	-0.0782	-0.0471	-0.0415	-0.0120
	(0.0171)	(0.0121)	(0.0173)	(0.0127)	(0.0690)	(0.0324)	(0.0398)	(0.0196)
SPLIT-ADJ - Partitioned Neighbors	0.0691***	0.0621***	0.0965***	0.0784***	0.1364**	0.0878**	0.0923	0.0874**
	(0.0220)	(0.0139)	(0.0299)	(0.0186)	(0.0662)	(0.0503)	(0.0666)	(0.0461)
Log Likelihood	-991.23	-819.79	-1215.02	-920.49	-2994.32	-2552.21	-2275.56	-1761.42
Region FE	Yes	No	Yes	No	Yes	No	Yes	No
Polynomial Latitude & Longitude	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Family FE	No	Yes	No	Yes	No	Yes	No	Yes
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	834	830	834	830	834	830	834	830

The table reports Poisson (in columns (1)-(4)) and Negative Binomial (in columns (5)-(8)) Maximum Likelihood (ML) estimates associating various measures of civil war with ethnic partitioning and natural resources. 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. The dependent variable in columns (5)-(6) is the number of casualties per thousand of square kilometers from civil conflicts that have taken place in the historical homeland of an ethnic group between 1970 and 2005. The dependent variable in columns (7)-(8) is the total number of years that an ethnic homeland has been under civil conflict over 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. ADJ (Neighbors) is the number of ethnic groups adjacent to the ethnic homeland. SPLIT-ADJ is the number of adjacent ethnic groups that are partitioned by the national border. NR is an indicator that identifies ethnic homelands with oil/petroleum fields or diamond mines. NR-SPLIT is an indicator that takes on the value of one for partitioned ethnicities that also have oil/petroleum fields or diamond mines. Odd-numbered specifications include a set of region fixed effects (constants not reported). Even-numbered specifications include a set of ethnic-family fixed effects (constants not reported) and a set of country fixed effects (constants not reported). All specifications include a cubic polynomial in latitude and longitude of the centroid of each ethnic group. The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Appendix Table 1: Correlation Structure of the Dependent Variables

Number of Civil Conflicts	1					
Number of Civil War Conflict Zones	0.8484*	1				
Civil Conflict Casualties (Best Estimate)	0.3669*	0.5851*	1			
Civil Conflict Casualties (Low Estimate)	0.3383*	0.5560*	0.9780*	1		
Civil Conflict Casualties (High Estimate)	0.3926*	0.5900*	0.9787*	0.9344*	1	
Civil Conflict Duration	0.6411*	0.7101*	0.5445*	0.5081*	0.5392*	1

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

Appendix Table 2: Summary Statistics for the Conditioning Variables

	Obs.	mean	st. dev.	min	p25	median	p75	max
				Panel A: Control Variables				
Land Area under Water	834	34.06	58.97	0.24	6.13	14.44	35.94	604.90
Land Area	834	0.86	2.25	0.00	0.01	0.17	0.66	27.66
Latitude	834	4.46	13.11	-33.09	-2.92	6.59	11.08	36.58
Longitude	834	17.85	15.47	-16.41	7.32	18.40	31.22	49.25
Population Density at Independence	834	17.47	25.71	0.00	3.17	9.26	20.47	321.53
Distance to the Capital City	834	504.58	375.03	11.31	256.43	392.82	633.13	1846.93
Distance to the Sea Coast	834	603.20	436.60	0.22	210.74	556.89	924.79	1721.30
Distance to the National Border	834	142.05	126.32	0.00	45.08	102.85	206.59	636.87
Mean Elevation	834	0.62	0.43	0.00	0.30	0.49	0.93	2.17
Land Suitability for Agriculture	834	0.41	0.24	0.00	0.25	0.42	0.57	0.98
Malaria Stability Index	834	0.75	0.36	0.00	0.57	0.98	1.00	1.00
Diamond Mine Indicator	834	0.12	0.33	0.00	0.00	0.00	0.00	1.00
Oil Deposit Indicator	834	0.12	0.40	0.00	0.00	0.00	0.00	4.00
Major City in 1400 AD Inidcator	834	0.04	0.20	0.00	0.00	0.00	0.00	2.00
Slave Exports	834	551.70	3289.21	0.00	0.00	0.00	17.67	41045.08
Distance to Explorer's Routes	834	205.10	223.39	0.23	58.48	126.77	264.64	1280.46
				Panel B: Pre-colonial Variables (from Murdock, 1967)				
Political Centralization	440	0.34	0.48	0.00	0.00	0.00	1.00	1.00
Class Stratification	396	1.3056	1.40	0.00	0.00	1.00	3.00	4.00
Share of Agriculture	490	5.63	1.79	0.00	5.00	6.00	7.00	9.00
Animal Husbandry	490	2.37	1.86	0.00	1.00	2.00	3.00	9.00

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

Appendix Table 3: Pre-colonial Ethnic Features and Partitioning

	Specification A - Unconditional Relationship		Specification B - Conditional Relationship	
	Additional Variable	Obs.	Additional Variable	Obs.
	(1)	(2)	(3)	(5)
Gathering	-0.0941	490	-0.0604	490
	(0.1721)		(0.0504)	
Hunting	-0.0311	490	0.0428	490
	(0.1229)		(0.0448)	
Fishing	0.2442*	490	-0.0039	490
	(0.1312)		(0.0513)	
Milking	0.0545	455	-0.0122	455
	(0.0410)		(0.0386)	
Polygyny	0.1782	490	-0.0041	490
	(0.1389)		(0.0491)	
Clan Communities	-0.1066**	490	0.0614	490
	(0.0438)		(0.0599)	
Settlement Pattern	0.0382	454	0.0046	454
	(0.1111)		(0.0122)	
Complex Settlements	-0.2337	454	0.0137	454
	(0.2851)		(0.0530)	
Jurisdictional Hierarchy of Local Community	-0.0779	443	0.0013	443
	(0.1381)		(0.0333)	
Elections	-0.0171	346	-0.0533	346
	(0.0353)		(0.0903)	
Property Rights	0.2341	375	0.021	375
	(0.1618)		(0.0898)	

The table reports OLS (linear probability model) estimates associating ethnic partitioning with pre-colonial ethnic variables (using data from Murdock (1967)). In all specifications the dependent variable is an indicator that equals one when at least 10% of the historical ethnic homeland (as portrayed in Murdock's (1959) Ethnolinguistic map) falls to more than one contemporary country. Specifications A are simple unconditional models. Specifications B include a set of region fixed-effects (constants not reported), log land area under water, and log land area. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Variable Definitions:

Gathering: Binary index that indicates the reliance of the economy on "the collection of wild plans and small land fauna." The index equals zero when the dependence is between 0% and 5%; the index equals one when dependence is greater than 5% dependence. Source: Murdock (1967); variable code in the Ethnographic Atlas v1.

Hunting: Binary index that indicates the intensity in hunting (including trapping and fowling). The index equals zero when the dependence is between 0% and 5%; the index equals one when dependence is greater than 5%. Source: Murdock (1967); variable code in the Ethnographic Atlas v2.

Fishing: Binary index that indicates the intensity in fishing (including shell fishing and the pursuit of large aquatic animals). The index equals zero when the dependence is between 0% and 5%; the index equals one when dependence is greater than 5%. Source: Murdock (1967); variable code in the Ethnographic Atlas v3.

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

Polygyny: Indicator that equals one when polygyny is practised and zero otherwise. The indicator equals one when the original variable indicates that polygyny is common and when large extended families are present. Source: Murdock (1967); variable code in the Ethnographic Atlas v8.

Clan Communities: Indicator that equals one when Murdock's community marriage organization variable (v15) equals 6 ("clan communities or clan barrios") and zero otherwise. Source: Murdock (1967); variable code in the Ethnographic Atlas v15.

Settlement Pattern: Ordered variable ranging from 1 to 8 quantifying "settlement pattern of each group". 1 indicates fully nomadic (migratory) groups, 2 indicates semi-nomadic, 3 indicates semi-sedentary, 4 identifies groups that live in compact and impermanent settlements, 5 indicates societies those in neighborhoods of dispersed family homes, 6 indicates for groups in separated hamlets forming a single community, 7 indicates societies living in compact and relatively permanent settlements, and 8 denotes the groups residing in complex settlements. Source: Murdock (1967); variable code in the Ethnographic Atlas v30.

Complex Settlements: Indicator that equals one for ethnicities living in compact and relatively permanent settlements (v30=7) or in complex settlements (v30=8), and zero otherwise. Source: Murdock (1967); variable code in the Ethnographic Atlas v30.

Jurisdictional Hierarchy of Local Community: Ordered variable ranging from 0 to 2 reflecting the hierarchy of local community organization. A zero score indicates the theoretical minimum of two (e.g., family and band), while a score of 2 indicates the theoretical maximum of four levels (e.g., nuclear family, extended family, clan barrio, village levels). Source: Murdock (1967); variable code in the Ethnographic Atlas v32.

Elections: Indicator that equals 1 when succession to the office of the local headman is conducted via "election or other formal consensus, nonhereditary" and zero otherwise. Source: Murdock (1967); variable code in the Ethnographic Atlas v72.

Property Rights: Indicator that equals one when some form of inheritance rule of real property (land) is present; the binary indicator equals zero when there is "absence of individual property rights". Source: Murdock (1967); variable code in the Ethnographic Atlas v74; the construction of the index follows Fenske (2009).

Appendix Table 4: Continuous Measure of Ethnic Partitioning and Civil Conflict

	Incidents		Locations		Casualties		Duration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
FRAC - Continuous Partitioning	0.5179**	0.5268***	0.4882**	0.4698***	0.7787**	0.8500*	0.7390**	0.6852**
	(0.2524)	(0.1490)	(0.2288)	(0.1537)	(0.4445)	(0.4575)	(0.3661)	(0.2798)
ADJ - Neighbors	-0.0165	0.0077	-0.0338**	0.0218	-0.0774	-0.0437	-0.0469	-0.0058
	(0.0149)	(0.0114)	(0.0143)	(0.0132)	(0.0515)	(0.0299)	(0.0310)	(0.0156)
SPLIT-ADJ - Partitioned Neighbors	0.0540***	0.0473***	0.0814***	0.0609***	0.1592***	0.0582	0.1041*	0.0719*
	(0.0212)	(0.0136)	(0.0267)	(0.0163)	(0.0619)	(0.0546)	(0.0597)	(0.0371)
Distance to the Border	-0.0004	0.0002	-0.0006	0.0002	-0.0006	-0.0006	0.0002	-0.0001
	(0.0007)	(0.0004)	(0.0007)	(0.0004)	(0.0010)	(0.0008)	(0.0010)	(0.0006)
Log Likelihood	-962.264	-817.307	-1158.484	-914.767	-2956.208	-2537.114	-2233.127	-1746.839
Region FE	Yes	No	Yes	No	Yes	No	Yes	No
Polynomial Latitude & Longitude	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rich Conditioning Set	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Family FE	No	Yes	No	Yes	No	Yes	No	Yes
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	834	830	834	830	834	830	834	830

The table reports Poisson (in columns (1)-(4)) and Negative Binomial (in columns (5)-(8)) Maximum Likelihood (ML) estimates associating various measures of civil war with a continuous index of ethnic partitioning (FRAC). 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. The dependent variable in columns (5)-(6) is the number of casualties per thousand of square kilometers from civil conflicts that have taken place in the historical homeland of an ethnic group between 1970 and 2005. The dependent variable in columns (7)-(8) is the total number of years that an ethnic homeland has been under civil conflict over the period 1970-2005. FRAC 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. ADJ (Neighbors) is the number of ethnic groups adjacent to the ethnic homeland. SPLIT-ADJ is the number of adjacent ethnic groups that are partitioned by the national border.

Odd-numbered specifications include a set of region fixed effects (constants not reported). Even-numbered specifications include a set of ethnic-family fixed effects (constants not reported) and a set of country fixed effects (constants not reported). All specifications include a cubic polynomial in latitude and longitude of the centroid of each ethnic group. Moreover, 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 density around independence (in 1960), 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 5: Ethnic Partitioning and Civil Conflict. Excluding Capital Cities

	Incidents		Locations		Casualties		Duration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT - Partitioning	0.1489	0.1765***	0.2218***	0.2025***	0.3105**	0.4404***	0.3003***	0.2586***
	(0.0999)	(0.0604)	(0.0927)	(0.0638)	(0.1311)	(0.1634)	(0.1125)	(0.1109)
ADJ - Neighbors	-0.0151	0.0027	-0.0279*	0.0174	-0.0686*	-0.0544**	-0.0334	-0.0107
	(0.0157)	(0.0116)	(0.0142)	(0.0130)	(0.0453)	(0.0279)	(0.0308)	(0.0181)
SPLIT-ADJ - Partitioned Neighbors	0.0599***	0.0500***	0.0879***	0.0694***	0.1682***	0.0812	0.0960*	0.0846**
	(0.0256)	(0.0162)	(0.0305)	(0.0178)	(0.0657)	(0.0557)	(0.0515)	(0.0419)
Distance to the Border	-0.0005	0.0002	-0.0006	0.0003	-0.0005	-0.0005	0.0001	-0.0002
	(0.0007)	(0.0004)	(0.0007)	(0.0004)	(0.0007)	(0.0007)	(0.0006)	(0.0004)
Log Likelihood	-902.723	-767.185	-1085.483	-858.857	-2754.551	-2388.294	-2100.799	-1655.854
Region FE	Yes	No	Yes	No	Yes	No	Yes	No
Polynomial Latitude & Longitude	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rich Conditioning Set	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Family FE	No	Yes	No	Yes	No	Yes	No	Yes
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	787	787	787	787	787	787	787	787

The table reports Poisson (in columns (1)-(4)) and Negative Binomial (in columns (5)-(8)) Maximum Likelihood (ML) estimates associating various measures of civil war with ethnic partitioning excluding ethnic homelands where capital cities fall. 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. The dependent variable in columns (5)-(6) is the number of casualties per thousand of square kilometers from civil conflicts that have taken place in the historical homeland of an ethnic group between 1970 and 2005. The dependent variable in columns (7)-(8) is the total number of years that an ethnic homeland has been under civil conflict over 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. ADJ (Neighbors) is the number of ethnic groups adjacent to the ethnic homeland. SPLIT-ADJ is the number of adjacent ethnic groups that are partitioned by the national border.

Odd-numbered specifications include a set of region fixed effects (constants not reported). Even-numbered specifications include a set of ethnic-family fixed effects (constants not reported) and a set of country fixed effects (constants not reported). All specifications include a cubic polynomial in latitude and longitude of the centroid of each ethnic group. Moreover, 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 density around independence (in 1960), 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 6: Partitioning and Civil Conflict

Sensitivity Analysis - Excluding African Regions
Country and Ethnic-Family Fixed Effects Specifications

	Incidents	Locations	Casualties		Duration	
	Poisson ML	Poisson ML	NB - ML	OLS	NB-ML	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Excluding North Africa			
SPLIT - Partitioning	0.1538***	0.1623***	0.2883**	0.2339*	0.1530*	0.1724**
	(0.0484)	(0.0523)	(0.1460)	(0.1273)	(0.0875)	(0.0849)
ADJ - Neighbors	0.0018	0.0132	-0.0618**	-0.0324	-0.0071	-0.0077
	(0.0100)	(0.0099)	(0.0266)	(0.0223)	(0.0135)	(0.0145)
SPLIT-ADJ	0.0526***	0.0704***	0.0855	0.0572	0.0631*	0.0519
- Partitioned Neighbors	-0.0093	-0.0134	-0.0532	-0.0423	-0.0372	-0.0366
Distance to the Border	0.0002	0.0002	-0.0003	-0.0001	-0.0002	0.0000
	(0.0004)	(0.0004)	(0.0010)	(0.0009)	(0.0007)	(0.0007)
Log Likelihood	-757.99	-847.60	-2306.28	__	-1572.71	__
adjsuted R-squared	__	__	__	0.866	__	0.876
Observations	766	766	766	766	766	766
			Panel B: Excluding South Africa			
SPLIT - Partitioning	0.1799***	0.2156***	0.5197**	0.3720*	0.3119**	0.2977**
	(0.0593)	(0.0640)	(0.2197)	(0.2131)	(0.1472)	(0.1456)
ADJ - Neighbors	0.0161	0.0248*	-0.0319	-0.0013	-0.0033	0.0047
	(0.0117)	(0.0138)	(0.0229)	(0.0239)	(0.0156)	(0.0173)
SPLIT-ADJ	0.0511***	0.0698***	0.0832*	0.0390	0.0891*	0.0640
- Partitioned Neighbors	(0.0140)	(0.0188)	(0.0450)	(0.0467)	(0.0473)	(0.0439)
Distance to the Border	0.0001	0.0002	-0.0005	-0.0001	-0.0002	0.0003
	(0.0004)	(0.0004)	(0.0008)	(0.0008)	(0.0007)	(0.0007)
Log Likelihood	-714.715	-790.842	-2137.61	__	-1457.389	__
adjsuted R-squared	__	__	__	0.864	__	0.866
Observations	740	740	740	740	740	740

Panel C: Excluding West Africa

SPLIT - Partitioning	0.2122***	0.2215***	0.4595*	0.4311*	0.2316*	0.2634
	(0.0699)	(0.0740)	(0.2347)	(0.2560)	(0.1305)	(0.1687)
ADJ - Neighbors	0.0072	0.0202	-0.0234	-0.0115	0.0048	0.0138
	(0.0115)	(0.0123)	(0.0258)	(0.0363)	(0.0119)	(0.0203)
SPLIT-ADJ	0.0423***	0.0533***	0.0371	0.054	0.0334	0.0414
- Partitioned Neighbors	(0.0159)	(0.0187)	(0.0523)	(0.0590)	(0.0332)	(0.0432)
Distance to the Border	0.0001	0.0002	-0.0005	-0.0003	-0.0002	-0.0002
	(0.0004)	(0.0004)	(0.0010)	(0.0010)	(0.0007)	(0.0007)
Log Likelihood	-612.545	-702.12	-2161.104	__	-1443.076	__
adjsuted R-squared	__	__	__	0.795	__	0.826
Observations	577	577	577	577	577	577

Panel D: Excluding East Africa

SPLIT - Partitioning	0.1845***	0.1902***	0.4686***	0.3548*	0.2924*	0.2837**
	(0.0512)	(0.0548)	(0.1811)	(0.1820)	(0.1604)	(0.1392)
ADJ - Neighbors	-0.0055	0.0065	-0.0576*	-0.0186	-0.0132	-0.0084
	(0.0118)	(0.0133)	(0.0305)	(0.0233)	(0.0200)	(0.0165)
SPLIT-ADJ	0.0617***	0.0797***	0.1131**	0.0825**	0.1111**	0.0906**
- Partitioned Neighbors	(0.0094)	(0.0165)	(0.0553)	(0.0411)	(0.0500)	(0.0402)
Distance to the Border	0.0000	0.0003	0.0001	0.0003	0.0000	0.0000
	(0.0004)	(0.0005)	(0.0011)	(0.0009)	(0.0008)	(0.0007)
Log Likelihood	-577.344	-643.636	-1692.446	__	-1161.97	__
adjsuted R-squared	__	__	__	0.864	__	0.851
Observations	609	609	609	609	609	609

Panel E: Excluding Central Africa

SPLIT - Partitioning	0.2023***	0.2369***	0.6627***	0.4696**	0.4326***	0.3382**
	(0.0692)	(0.0660)	(0.2049)	(0.1996)	(0.1391)	(0.1352)
ADJ - Neighbors	0.0166	0.0209*	-0.0789**	-0.0241	-0.0138	0.0000
	(0.0118)	(0.0124)	(0.0329)	(0.0298)	(0.0225)	(0.0221)
SPLIT-ADJ	0.0292	0.0531***	0.1277**	0.0851	0.0951**	0.0705
- Partitioned Neighbors	(0.0208)	(0.0203)	(0.0604)	(0.0529)	(0.0460)	(0.0430)
Distance to the Border	0.0005*	0.0005	-0.0001	0.0003	0.0004	0.0006
	(0.0003)	(0.0003)	(0.0010)	(0.0009)	(0.0004)	(0.0005)
Log Likelihood	-579.129	-634.134	-1695.282	__	-1202.611	__
adjsuted R-squared	__	__	__	0.822	__	0.841
Observations	617	617	617	617	617	617
Polynomial Latitude & Longitude	Yes	Yes	Yes	Yes	Yes	Yes
Rich Conditioning Set	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Family FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes

The table reports Poisson ML (in columns (1) and (2)), Negative Binomial ML (in columns (3) and (5)) and OLS (in columns (4) and (6)) estimates associating various measures of civil war with ethnic partitioning. In Panel A we exclude ethnic homelands in Northern Africa. In Panel B we exclude ethnic homelands in Southern Africa. In Panel C we exclude ethnic homelands in West Africa. In Panel D we exclude ethnic homelands in Eastern Africa. In Panel E we exclude ethnic homelands in Central Africa. The regional classification follows Nunn (2008).

The dependent variable in column (1) 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 column (2) 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. The dependent variable in column (3) is the number of casualties per thousand of square kilometers from civil conflicts that have taken place in the historical homeland of an ethnic group between 1970 and 2005. The dependent variable in column (4) is the log of one plus the number of casualties per thousand of square kilometers from civil conflicts that have taken place in the historical homeland of an ethnic group between 1970 and 2005.The dependent variable in column (5) is the total number of years that an ethnic homeland has been under civil conflict over the period 1970-2005. The dependent variable in column (6) is the log of one plus the total number of years that an ethnic homeland has been under civil conflict over 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. ADJ (Neighbors) is the number of ethnic groups adjacent to the ethnic homeland. SPLIT-ADJ is the number of adjacent ethnic groups that are partitioned by the national border. All specifications include a set of ethnic-family fixed effects (constants not reported) and a set of country fixed effects (constants not reported). All specifications include a cubic polynomial in latitude and longitude of the centroid of each ethnic group. Moreover, 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 density around independence (in 1960), 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.