

NBER WORKING PAPER SERIES

THE LONG-RUN EFFECTS OF THE SCRAMBLE FOR AFRICA

Stelios Michalopoulos
Elias Papaioannou

Working Paper 17620
<http://www.nber.org/papers/w17620>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2011

We would like to thank the Editor and 4 referees for insightful comments and useful suggestions. We also thank Alberto Alesina, Maarten Bosker, Chris Blattman, Francesco Caselli, Jeremiah Dittmar, Joan Esteban, James Fenske, Fabrizio Dell'Acqua, Federico Finan, David Laitin, Horacio Larreguy Arbesú, Stathis Kalyvas, Sebastian Hohmann, Janina Matuszeski, Ted Miguel, Nathan Nunn, Gregorios Siourounis, Neils Weidman, Andreas Wimmer, and participants at the AEA meetings in Denver, Yale, UC Berkeley, Princeton, ALBA, UC Irvine, UC Merced, American University, Brown, the Institute for Economic Analysis, Autonoma University, George Washington University, CERGE-EI, Surrey, City University, the NBER Political Economy meetings, the CEPR meeting on the Political Economy of Conflict, the 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2011 by Stelios Michalopoulos and Elias Papaioannou. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Long-Run Effects of the Scramble for Africa
Stelios Michalopoulos and Elias Papaioannou
NBER Working Paper No. 17620
November 2011, Revised March 2015
JEL No. N17,N47,O10,Z10

ABSTRACT

We examine the long-run consequences of ethnic partitioning, a neglected aspect of the Scramble for Africa caused by the colonial border drawing, and uncover the following regularities. First, apart from the land mass and presence of water bodies, historical homelands of split and non-split groups are similar across a wealth of observable characteristics. Second, using geo-referenced data on conflict and exploiting within-country variation, we show that the incidence, severity and duration of violence are higher in the historical homelands of partitioned groups. Third, we shed some light on the mechanisms showing that military interventions from neighboring countries and conflict between government forces and rebels that aim at countering state authority are much more likely in the homelands of split groups. Fourth, our exploration of the status of ethnic groups in the political arena reveals that partitioned ethnicities are systematically discriminated from the national government and are more likely to participate in ethnic civil wars. Finally, using micro-level data we find that individuals identifying with split groups have lower access to public goods and lower education. The uncovered evidence brings in the foreground the detrimental repercussions of ethnic partitioning.

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

Elias Papaioannou
London Business School
Regent's Park
Sussex Place
London NW1 4SA
United Kingdom
and NBER
papaioannou.elias@gmail.com

1 Introduction

The predominant explanations on the deep roots of contemporary African underdevelopment are centered around the influence of Europeans during the colonial period (Acemoglu *et al.* (2001, 2002, 2005)), but also in the centuries before colonization when close to 20 million slaves were exported from Africa (Nunn (2008)). Yet in the period between the ending of the slave trades and the colonial rule, another major event took place in European capitals that according to the African historiography had malicious long-lasting consequences. The "Scramble for Africa" starts with the Berlin Conference of 1884–1885 and is completed by the turn of the 20th century. During this period, Europeans partitioned Africa into spheres of influence, protectorates, and colonies. The borders were designed in European capitals at a time when Europeans had barely settled in Africa and had little knowledge of the local conditions. Despite their arbitrariness these boundaries endured after African independence in the 1960s. As a result in many African countries a significant fraction of the population belongs to ethnic groups that are partitioned among different states.¹

A considerable body of work in African historiography (e.g., Asiwaju (1985), Dowden (2008), Wes-seling (1996), Thomson (2010)) argues that the main channel of Europeans' influence on African development was not colonization per se, but the improper border design. Herbst (2000) summarizes "*for the first time in Africa's history [at independence], territorial boundaries acquired salience...The boundaries were, in many ways, the most consequential part of the colonial state.*" The artificial borders fostered ethnic struggles, patronage politics, and conflict primarily by splitting groups across the newly-minted African states. Ethnic partitioning led to irredentism and helped create an ideology of secession and nationalism (Horowitz (1985)). Moreover, split groups have often been used instrumentally by governments to destabilize neighboring countries, setting the stage for discrimination of split ethnicities in the political sphere and a rationale for governments to push them to the other side of the border.

Despite the wealth of anecdotal evidence, there is little work formally examining the ramifications of ethnic partitioning in the context of the Scramble for Africa. Some cross-country studies have touched upon this issue, showing, that the likelihood of conflict increases when there is an ethnic war in adjacent nations (Bosker and de Ree (2014)) and that countries with straight borders, where a large share of the population belongs to ethnicities that are present in nearby nations, perform economically worse (Alesina, Easterly, and Matuszeski (2011)).² Nevertheless, to the best of our knowledge there is no empirical work directly exploring the consequences of ethnic partitioning for African groups where the arbitrary border design and the large number of split groups offer the opportunity to cleanly identify the importance of partitioning and explore the mechanisms at work. This study is a step in this direction.

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

²We discuss below the key differences of our work with Alesina, Easterly, and Matuszeski (2011).

Results To formally assess the claim that African borders were drawn with little respect to the local political geography, we explore whether partitioned ethnicities differ systematically across several geographic-ecological traits vis a vis non-split groups. With the exceptions of land mass of the historical ethnic homeland and presence of lakes, there are no significant differences between split and non-split homelands along a comprehensive set of covariates. We further show that there are no systematic differences across several pre-colonial, ethnic-specific, institutional, cultural, and economic features, such as the size of settlements, the type of the subsistence economy, and proxies of pre-colonial conflict. Our results, hence, offer empirical support to a long-standing assertion within African historiography regarding the largely arbitrary nature of the African border design, at least with respect to ethnic partitioning.

We then employ the Scramble for Africa as a "quasi-natural" experiment to assess the impact of ethnic partitioning on civil conflict. Using a newly-assembled dataset (Armed Conflict Location & Event Data Project (ACLED)) that reports geo-referenced information for 1997–2013 on the location of incidents of political violence, including battles between government forces, rebels and militias and violence against civilians, we show that civil conflict is higher in the homelands of partitioned ethnicities. This applies to conflict intensity, duration, and casualties as well as the likelihood of conflict. We obtain similar results when we restrict estimation to ethnic homelands close to the national borders. Our estimates suggest that civil conflict intensity is approximately 40% higher in areas where partitioned ethnicities reside as compared to the homelands of ethnic groups that have not been separated by the national borders. We estimate that the likelihood of any conflict is approximately 8% higher in the homelands of split ethnicities.

We then exploit the richness of ACLED to shed light on the mechanisms. First, we examine the thesis of the African historiography that split groups are used by neighboring countries to stage proxy wars and destabilize the government on the other side of the border. We find clear evidence supporting this conjecture. Military interventions from neighboring countries are more common in the homelands of partitioned groups, rather than in nearby areas where non-split groups reside. Second, we show that ethnic partitioning matters crucially for conflict between government troops and rebel groups "*whose goal is to counter an established national governing regime by violent acts*". This specific pattern is corroborated using an alternative geo-referenced conflict database (Uppsala Conflict Data Program Geo-referenced Event Dataset, UCDP - GED) that records deadly events associated with major civil wars. In contrast, there is no link between ethnic partitioning and riots and protests, which are mostly present in the capitals; and there is no association between partitioning and conflict between non-state actors. This result is in accord with African historiography that points out that partitioned groups face discrimination from the national government and often engage in rebellions to counter state presence.

In an attempt to dig deeper on the partitioning - discrimination - civil war nexus we use the Ethnic Power Relations (EPR) dataset (Wimmer, Cederman, and Min (2009)) that provides an assessment of formal and informal degrees of political participation of ethnic groups in the political arena over the full post-independence period. The within-country analysis shows that partitioned ethnicities are significantly

more likely to engage in major civil wars that have an explicit ethnic dimension; moreover, the likelihood that split ethnicities are subject to political discrimination from the national government is approximately 9 percentage points higher compared to non-split groups.

Finally, we complement the conflict-based analysis with individual-level evidence from the Demographic and Health Surveys (DHS) spanning more than 85,000 households across 20 African countries. Members of partitioned groups, irrespective of their current residence, have fewer household assets, poorer access to utilities, and worse educational outcomes, as compared to individuals from non-split ethnicities in the same country (and even in the same enumeration area).

Related Literature Our paper belongs to the genre of studies that investigate the historical origins of comparative development (see Nunn (2013) for a review). The literature has mainly focused on the impact of colonization mainly via the early institutions (e.g., Acemoglu, Johnson, and Robinson (2005)), such as the system of indirect rule (e.g., Acemoglu, Reed, and Robinson (2014)), infrastructure (e.g., Huillery (2009), Jedwab and Moradi (2015)), and human capital (e.g., Easterly and Levine (2012), and Wantchekon, Klasnja, and Novta (2015)). We emphasize, an aspect of colonization, the drawing of political boundaries in the end of the 20th century that resulted in a large number of partitioned ethnicities, that has been neglected by economics research. As such our work is mostly related to Alesina, Easterly, and Matuszeski (2011) who show with cross-country regressions that "artificial states" with straight borders and where a significant part of the population resides in more than one country, under-perform economically compared to more "organic" countries.

A related body of research searches for the origins of African countries' weak state capacity in the pre-colonial period. Nunn (2008) and Nunn and Wantchekon (2011) show that the slave trades (1400 – 1900) have crucially shaped African development mostly by spurring ethnic conflict and by lowering trust. Gennaioli and Rainer (2006, 2007) and Michalopoulos and Papaioannou (2013) show that deeply-rooted pre-colonial ethnic institutions correlate significantly with contemporary development. Our paper relates to these contributions, as we also study the long-run implications of historical legacies focusing on the ethnic dimension. Yet rather than focusing on the pre-colonial period, we examine the impact of ethnic partitioning during the early stages of colonization, a topic that has not received much attention. Assessing the impact of deeply-rooted features seems important, as Michalopoulos and Papaioannou (2014) show that the impact of contemporary national institutions on regional African development is small, mostly because state's power to broadcast power diminishes rapidly outside the capitals (see also Herbst (2000)).³ So, while these studies are related to our paper, they examine intrinsically different themes.

³In particular in Michalopoulos and Papaioannou (2014) we employ a spatial regression discontinuity design that examines the impact of national institutions on regional development (as reflected on satellite images on light density at night) at the border, exploiting within-ethnicity across-country variation. The analysis reveals two key results. First, we find that on average differences in national institutions (rule of law, control for corruption) do not translate on average on differences in development. Second, we find that the average non-effect masks considerable heterogeneity, which is linked to the limited penetration of national institutions in areas far from the capital.

Our paper also contributes to the literature on the origins of civil conflict that mainly examines the role of country-level characteristics (see Collier and Hoeffler (2007) and Blattman and Miguel (2010) for reviews and Collier and Sambanis (2005) for case studies in Africa). Of most relevance are works that link a country’s ethnic composition to civil war. While the correlation between ethnic fragmentation and civil war is weak (Fearon and Laitin (2003)), recent studies document interesting cross-country correlations associating various aspects of the societal structure with armed conflict. Montalvo and Reynal-Querol (2005) and Esteban, Mayoral, and Ray (2012) show a strong association between ethnic polarization and conflict, while Esteban and Ray (2008, 2011) emphasize the role of differences between and within ethnic groups. Huber and Mayoral (2014) show a link between within-group inequality 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. We complement this research showing that ethnic partitioning affects both the likelihood of ethnic-based civil wars and political discrimination. We further show that the link between ethnic partitioning and civil conflict partially operates via discrimination.⁴

The correlations found in studies linking cross-country variation in border features and the distribution of ethnicities to development proxies (income or conflict) are informative (e.g., Alesina, Easterly, and Matuszeski (2011), Englebort, Tarango, and Carter (2002), Bosker and de Ree (2014)), but they cannot be casually interpreted (see Blattman and Miguel (2010) and Fuchs-Schundeln and Hassan (2015)). The main endogeneity concern is that the process of border drawing is usually an outcome of state formation that affects directly economic performance and conflict. As the recent literature on state capacity shows, nation building, development, and conflict are inter-linked and jointly driven by hard-to-account-for factors related to societal structure, geography, and historical legacies (e.g., Besley and Persson (2011)). Thus, selection, reverse causation, and omitted variables are non-negligible issues. Likewise, due to measurement error in the main independent variables, multi-colinearity, and the limited degrees of freedom, the cross-country correlations are sensitive to small permutations and data revisions (see Hegre and Sambanis (2006) and Ciccone and Jarocinski (2010)).

By exploiting variation across ethnic homelands, we account for many shortcomings of cross-country works. First, by showing that there are no systematic differences in geographic, economic, institutional, and cultural characteristics between split and non-split ethnicities, our analysis offers large-scale econometric evidence supportive to the African historiography on the accidental nature of African borders, at least with respect to the ethnic partitioning dimension.⁵ Second, the use of information on the spatial distribution of ethnicities in the end of 19th century, well before the current national boundaries came into effect, alleviates concerns related to migratory flows ignited by the border design. Since borders were drawn by

⁴Englebort, Tarango, and Carter (2002) show a positive cross-country correlation between proxy measures of suffocation and dismemberment and violence, secession attempts, border disputes, and warfare.

⁵Admittedly, we cannot entirely rule out that some unobserved factor may have been taken into account in the process of border drawing. Nevertheless, given the exhaustive list of covariates considered in the analysis and the overwhelming evidence of the African history on the arbitrariness of borders, our results suggest that the impact of unobservable factors are unlikely to be of first-order significance.

Europeans with limited knowledge of local conditions and did not change at independence, we focus on cases where country boundaries were not the result of political, economic, or military developments; thus selection and reverse causation issues are secondary -if not absent. Third, focusing on ethnic groups is conceptually appealing in the context of Africa, where ethnic identification is strong, ethnic segregation high and conflict has a strong ethnic component. In their synthesis of the case-study evidence on conflict in Africa and the results of cross-country regressions, Collier and Sambanis (2005) note "*the country-year is not the appropriate unit of observation to study such wars. Instead it would be more appropriate to focus on the ethnic group or we should analyze patterns of violence in a geographical region that does not necessarily correspond to predefined national boundaries.*" Fourth, by looking into different subsets of conflict and exploiting ethnic-level data from the Ethnic Power Database on political discrimination and individual-level data from the DHS we shed some light on the mechanisms at work.

Structure The next section provides a synopsis of the historical background and presents the key arguments on the impact of the Scramble for Africa. In Section 3 we first discuss how we identify partitioned ethnicities and then examine whether there are systematic differences between split and non-split groups with respect to an array of geographic and historical features that may independently affect conflict. Section 4 reports our estimates on the effect of partitioning on various aspects of civil conflict (likelihood, intensity, duration and fatalities). In Section 5 we explore some the mechanisms. In Section 6 we explore the connection between partitioning, ethnic-based discrimination and ethnic wars. Section 7 presents the individual-level analysis linking education and access to utilities to ethnic partitioning. In Section 8 we summarize and discuss avenues for future research.

2 Historical Background

2.1 The Scramble for Africa

The "Scramble for Africa" starts in the 1860s when the French and the British begin the systematic exploration of Western Africa and sign bilateral agreements assigning spheres of influence. In the next 40 years, Europeans signed hundreds of treaties that divided 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. The key consideration was to preserve the "status quo" preventing conflict among Europeans for Africa, as the memories of the European wars of the 18th-19th century were alive. As a result, in the overwhelming majority of cases, European powers drew borders without taking into account local conditions. African

leaders were not invited and had no say.⁶ In many cases, European leaders were in such a rush that they didn't wait for the information arriving from explorers, geographers, and missionaries. As the British prime minister at the time Lord Salisbury put it, "*we have been engaged in drawing lines upon maps where no white man's feet have ever trod; we have been giving away mountains and rivers and lakes to each other, only hindered by the small impediment that we never knew exactly where the mountains and rivers and lakes were.*" Asiwaaju (1985) summarizes that "*the study of European archives supports the accidental rather than a conspiratorial theory of the marking of African boundaries.*" In line with the historical evidence, Alesina, Easterly, and Matuszeski (2011) document that eighty percent of African borders follow latitudinal and longitudinal lines, more than in any other part of the world.

Several factors have been proposed to rationalize the accidental border design. First, at the time Europeans had little knowledge of local geographic conditions, as with the exception of few coastal areas, the continent was unexplored. 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.⁷ Second, Europeans were not drawing borders of prospective states, but of colonies and protectorates; clearly at the time none could foresee independence. Third, demarcation was poor.⁸ Fourth, Europeans were unwilling to sacrifice their commitment not to go to war for any part of Africa and hence were reluctant to change colonial borders.⁹ Fifth, as locals could freely move across colonial borders, African chiefs did not oppose much the colonial design, as little changed on the ground.¹⁰ Wesseling (1996) summarizes: "*The partition of Africa was recorded by the Europeans on their maps, but the matter rested there for the time being....In Europe conquests preceded the drawing of maps; in Africa the map was drawn, and then it was decided what was going to happen. These maps did not therefore reflect reality but helped to create it.*"

African independence occurred at a fast speed that not even the key protagonists expected (Herbst (2000)). The independence of Northern African countries in the 1950s was soon followed by Ghana's and Guinea's independence in 1957 and in 1958, respectively. By the end of 1966, 40 countries had become independent. While at the time, many proposed changing the borders, African leaders and leaving Europeans did not touch this issue. The leaders of the newly-crafted African states believed that nation-building and

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

⁷ An illustrative example is the annexation of Katanga in Congo Free State that turned out to be its richest province. King Leopold got Katanga in exchange for the Niari-Kwilu area that the French insisted on getting themselves. Wesseling (1996) writes "*what impelled him [Leopold] was a general imperialist surge, the desire for compensation for the Niari-Kwilu, and the objective of making the new state as large as possible and filling as much of the Congo basin as possible.*"

⁸ Poor demarcation and imprecise colonial treaties on exact boundaries have contributed to conflict after independence. Examples include the war between Tanzania and Uganda in 1978 over the Kagera salient (a 1800 km² strip of land) and the conflict between Burkina Faso and Mali over the Agacher strip in 1985.

⁹ 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.*"

¹⁰ Asiwaaju (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.*"

industrialization would sideline ethnic divisions. National leaders feared that border realignment would threaten their position, whereas Europeans' main objective was to maintain the special rights and corporate deals with their former colonies, and, as such, they were also reluctant to open the border issue. Almost all African countries accepted the colonial borders when signing the Charter of the Organization of African Union (OAU) in 1964. Only Somalia and Morocco did not accept the borders, while Ghana and Togo raised some objections on their boundary that splits the Ewe, but the border did not change. The freezing of the colonial borders by the OAU compact allows us to explore their consequences in a quasi-experimental setting that is key for identifying causal effects, something that would not have been possible if post-colonial states bargained and redrew the colonial borders.

2.2 Channels and Case Studies

Irredentism, secession, and autonomy The literature has stressed the impact of ethnic partitioning on generating irredentist demands, as split ethnicities may want to unify with their peers across the border.¹¹ In line with this argument, Wimmer, Cederman, and Min (2009) estimate that 20% of all civil wars in Africa have a secessionist demand.¹² While, compared to the number of civil wars in Africa, there have been few cases of secession, irredentism and the associated ideology has played a key role in some major conflicts, mostly in Somalia, Mali, and Senegal. Somali people, for example, were split during colonization between three different European colonies, while Ethiopia also got a slice. The five-pointed star in the flag of Somalia symbolizes the five regions inhabited by Somali clans (Italian Somaliland, North Kenya, Southern Ethiopia, French Djibouti, and British Somaliland); three long-lasting wars have been partly driven by the desire of Somalis in Ethiopia and Kenya to become part of Somalia (Meredith (2005)).¹³ Similarly, in the initial years after independence Kenya experienced substantial conflict in the Northern Frontier District as Somali insurgents (*shifta*) were fighting for annexation to Somalia (Touval (1967)). Another example are the Tuareg who in the early 1990s declared autonomy both in Niger and in Mali, with some of their leaders envisioning a unified Tuareg state. Even in cases that obtaining autonomy

¹¹Horowitz (1985) notes "a quick tour d'horizon reveals the rich range of possibilities (for conflict and irredentism). The Ghana-Togo border divides the Ewe, as the Nigeria-Benin border divides the Yoruba. There are Hausa in Nigeria and Hausa in Niger. There are Fulani across a wide belt of West and Central Africa, Beteke in Gabon and Congo (Brazzaville), and Fang in Cameroon, Gabon, and Equatorial Guinea. The Bakongo are divided among, Zaire, Congo (Brazzaville) and Angola; the Lunda among Zaire, Zambia, and Angola. There are Somalis in Somalia, Ethiopia,, Kenya, and Djibouti. There are Wolof in Mauritania, in Gambia, and in Senegal, Kakwa in Sudan and in Uganda. And various Berber groups are distributed among more than one North African state."

¹²Civil wars with a secession demand are almost absent in Central and South America. Besides Africa, secession-driven conflicts are found in the Middle East, India, and the Caucasus.

¹³UCDP describes the conflict in the Ogaden region in Southern Ethiopia as follows: "The independent Somali Republic was created in 1960 by the merger of only two of these entities (British Somaliland and Italian Somaliland); something that did not satisfy the aspirations of Somali nationalism. Subsequently, in August 1960 the government of the newly independent state published a manifesto that called for the formation of a Greater Somalia, which would include Djibouti, the Northern Frontier District of Kenya, and above all the Ogaden desert and adjacent areas in south-eastern Ethiopia. The Somali population in Ethiopia had received little from the Addis Ababa government since it came under its rule in the late 1880s. When Somalia became independent and began spreading the idea of Somali nationalism, it found fertile soil in the Ogaden region. Irredentist agitation and armed clashes soon commenced, and increased as the Ethiopian government launched its first systematic attempt to collect taxes in the region."

or independence is not the ultimate goal, ethnic partitioning creates an ideology that local ethnic parties exploit in pursuit of their special interest (Horowitz (1985)).

Discrimination Groups in Africa often become targets of the central governments. Ethnic-based discrimination is pervasive and a large body of research provides ample evidence on ethnic-based politics (Posner (2005)). National governments frequently attempt to suffocate ethnicities by seizing property, imposing high taxation and restrictions on the activities of specific groups (Bates (1981)). Examples include the (Hu)Ambo and the Chokwe in Angola, the I(g)bo in Nigeria, Tuareg clans in Mali and Niger, and the Oromo and Somalis in Ethiopia. What is different between partitioned and non-split groups, though, is that split ones can seek shelter in their ancestral homeland on the other side of the border. Split ethnicities can re-organize, obtain arms, and get assistance from their co-ethnics across the border both when they are on the defense and when they attack. Moreover, partitioned groups are regularly used by African governments to destabilize neighboring countries. An example is the case of the Ewe, a large group split between Ghana and Togo. The Ewe in Togo helped Flt.-Lt. Jerry Rawlings (half Ewe) in his coup in 1979 and 1981 to overthrow the government in Ghana, as the Ewe have been discriminated for most of the post-independence years. The recurrent conflict in the Casamance region in Southern Senegal, where the partitioned Diola (Jola) reside, offers another example. As Gambia effectively splits Senegal into a Northern and a Southern part, the Casamance province is disconnected from the central government in Dakar. The independence "Movement of the Democratic Forces of Casamance (MDFC)" was supported by the neighboring Guinea-Bissau, where the Diola exert significant political influence (Humphreys and Mohamed (2005)).

Spillovers Spillovers also occur because of population displacements, as refugee flows may change the ethnic composition in adjacent countries igniting conflict. A pertinent example is that of the Alur, a group split between the Belgian Congo and the British Protectorate of Uganda during the late phase of the Scramble for Africa (1910 – 1914). When Mobutu Sese Seko initiated the subjugation of several minority groups in Zaire, a large portion of the Alur were pushed to Uganda. This in turn generated opposition from the Buganda in Southern Uganda leading to conflict (Asiwaju (1985)). Fearon and Laitin (2011) report that 31% of civil wars (and 57% of ethnic wars) involve "*members of a regional ethnic group that considers itself to be the indigenous sons-of-the-soil and recent migrants from other parts of the country*".¹⁴

Other Aspects of the Scramble for Africa Besides ethnic partitioning, the artificial border design may have contributed to underdevelopment and conflict via other channels (that we do not consider in this paper). Border drawing shaped a host of country-specific geographical and cultural characteristics including a country's ethnic heterogeneity, polarization, land size, and access to the coast that affect

¹⁴Fearon and Laitin (2011) list eight conflicts in Africa (26% of all wars) that involved indigenous versus *within-country* migrants (e.g., Tuareg in Mali in 1989, Senegal in 1989 involving Diolas in Casamance, etc.).

development. Herbst (2000) argues that civil conflict is more pervasive in large African countries that find it hard to broadcast power across their territories. Collier (2007) discusses how the border design resulted in Africa having the largest proportion of landlocked countries limiting their growth potential. While our analysis focuses on a single aspect of the Scramble for Africa, that of ethnic partitioning, we are able to account for these other aspects with the inclusion of country fixed effects that absorb common-to-all-homelands, country-specific characteristics.¹⁵

3 Ethnic Partitioning and Border Artificiality

3.1 Identifying Partitioned Ethnic Groups

We identify partitioned groups projecting contemporary borders, as portrayed in the 2000 Digital Chart of the World, on George Peter Murdock’s Ethnolinguistic Map (1959) that depicts the spatial distribution of African ethnicities at the time of European colonization in the mid/late 19th century and early 20th century (Figure 1a).¹⁶ Murdock’s map divides Africa into 843 regions. The mapped ethnicities correspond roughly to levels 7 – 8 of the Ethnologue’s language family tree. 8 areas in the Sahara are "uninhabited upon colonization" and are therefore not considered. We also drop the Guanche, a small group in the Madeira Islands that is currently part of Portugal and the Comorians, as none of the conflict databases covers Comoros. This leaves us with 833 groups. We also exclude 8 regions where population according to the earliest post-independence census is zero.¹⁷ So our analysis focuses on 825 ethnicities.¹⁸

Out of a total of 833 ethnicities in Murdock’s Map, the homeland of 357 groups falls into more than one country. Yet for several of these groups the overwhelming majority of their homeland’s area (usually more than 99%) falls into a single country. For example, 99.5% of the area of the Ahaggaren falls into Algeria and only 0.5% falls in Niger. Since Murdock’s map is bound to be drawn with some error, we identify as partitioned 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 229 ethnicities (27.7% of the sample) with at least 10% of their historical homeland falling into more than one contemporary state (Figure 1b).¹⁹ When we use a broader threshold of 5% we identify 266 partitioned

¹⁵In the Supplementary Appendix we explore how these different nationwide by-products of border drawing interact with ethnic partitioning in mitigating or magnifying conflict.

¹⁶Murdock’s map is based on various primary sources covering approximately the period 1860 – 1940. Most observations correspond to 1890, 1900, and 1910. We drop regions of less than 100 km^2 because these small areas are most likely an outcome of measurement error in the underlying historical mapping of ethnicities.

¹⁷These groups are the Bahariya, the Fertit, the Ifora, the Kimr, the Matumbi, the Midobi, the Mituku, and the Popoi. The results are identical if we were to retain these ethnic areas, assigning to them a very small population number.

¹⁸Since in our empirical analysis we primarily explore within-country variation, in many specifications we lose observations in countries with either a single ethnicity or without variability in ethnic partitioning. These countries are Burundi, Djibouti, Swaziland, Madagascar, and Western Sahara.

¹⁹We apply the same threshold, as in our work assessing the within-ethnicity across-the-border impact of national institutions on contemporary development. In Michalopoulos and Papaioannou (2014) we focus, however, on 220 split groups. The 9-groups difference emerges because: (i) three ethnicities were dropped in Michalopoulos and Papaioannou (2014) as they are split between Western Sahara and Morocco and there are no data on national institutions for Western Sahara; (ii) six groups were dropped because the population estimate is zero in *one* of the two partitions in 2000.

groups. Appendix Table A lists partitioned ethnicities.

Our procedure identifies most major ethnic groups that have been split by African borders. For example, the Maasai are partitioned between Kenya and Tanzania (shares 62% and 38%), the Anyi between Ghana and the Ivory Coast (shares 58% and 42%), and the Chewa between Mozambique (50%), Malawi (34%), and Zimbabwe (16%). Other examples include the Hausa (split between Nigeria and Niger), the Ababda (split between Egypt and Sudan), the Ewe (split by the Togo-Ghana border), and the Barareta clans (split between Kenya and Somalia). We also checked whether our coding 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.²⁰



Figure 1a

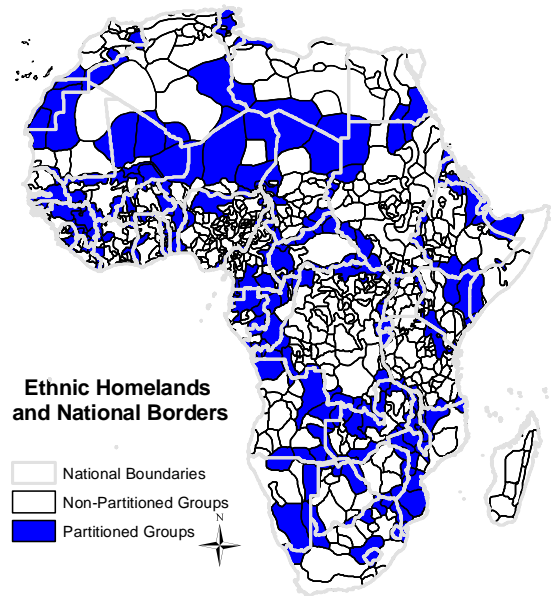


Figure 1b

3.2 Border Artificiality

The African historiography provides ample evidence arguing that, in the majority of cases, Europeans did not consider ethnic features and local geography in the design of colonial borders. In a few instances, nevertheless, Europeans did try taking into account political geography, as, for example, in Swaziland, Burundi, and in some parts of Uganda. And few borders were delineated in the late stage of colonization in the early 20th century, when Europeans conceivably had some knowledge of local conditions. Moreover, some contemporary borders in Western Africa follow the French administrative divisions. And in two cases

²⁰We should stress here that this approach is imperfect, as ethnic groups overlap and because there is certainly noise in Murdock's map. As such the ethnic partitioning index is likely to contain error-in-variables. For example, our procedure identifies as non-split the conflict-prone Ogaden region-clan (it enters as partitioned when we adopt the 5% threshold for the classification of groups) and the Sab clan in Ethiopia, although all our readings suggest that these groups have been impacted by the Ethiopian-Somali border. Since our procedure is simply based on the overlap of the historical map with the contemporary boundaries, this error is unlikely to be systematic (correlated with contemporary conflict or the key controls).

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

We search for potential correlates of ethnic partitioning estimating simple (linear probability) models that associate the binary ethnic partitioning index (*SPLIT*) with various geographic, ecological, natural resource variables and proxies of pre-colonial conflict and development at the ethnicity level.²¹ Table 1 reports the results. In all specifications we include region-specific constants to account for the different timing and patterns of colonization. Below the estimates, we report double-clustered standard errors at the country and at the ethnic-family level using the method of Cameron, Gelbach, and Miller (2011) that accounts for spatial correlation and arbitrary residual correlation within each dimension.²²

Geography, Ecology, Natural Resources and Ethnic Partitioning In Table 1 - Panel A we examine the impact of geography, ecology and natural resources. The positive and highly significant estimate of (log) land area in column (1) suggests that ethnic groups spanning large territories in the pre-colonial period were more likely to be partitioned. This is consistent with the arbitrary nature of border drawing. In column (2) we augment the specification with indicators that identify ethnic homelands with a large lake and a main river. The coefficient on the lake dummy is positive and significant at the 90% level, while the river indicator enters with a small and statistically indistinguishable from zero coefficient. These results are again in accord with the historical evidence that Europeans attempted to use natural barriers while delineating spheres of influence. In column (3) we add an index reflecting land quality for agriculture and elevation. Both variables enter with small and insignificant coefficients. In column (4) we examine the role of ecological conditions augmenting the empirical model with a malaria stability index and distance to the coast. Since Europeans settled mostly in coastal areas and regions where malaria was less pervasive, these specifications shed light on whether early contact with colonizers affected partitioning. Both indicators enter with insignificant estimates. In column (5) we include on the RHS indicators identifying ethnic areas with diamond mines and petroleum fields. While in the initial phase of colonization Europeans were mostly interested in agricultural goods and minerals, adding these indicators allows us to investigate whether partitioned groups differ from non-partitioned ones in terms of natural resources whose presence has been linked to conflict (see Ross (2012)). There are no systematic differences between the two sets of ethnic homelands. In column (6) we augment the specification with the share of

²¹Appendix Table 1 reports summary statistics for all variables at the ethnic homeland level. The Data Appendix gives variable definitions and sources. The results are similar with probit and logit ML estimation.

²²Cameron, Gelbach, and Miller (2011) explicitly cite spatial correlation as an application of the multi-way clustering method. Murdock (1959) assigns the 833 ethnic groups into 96 ethnolinguistic clusters/families. We also estimated standard errors using the method of Conley (1999) to account for spatial dependence of an unknown form, finding similar results.

adjacent ethnicities that are of the same ethnolinguistic family, to examine whether Europeans took into account local ethnic-linguistic differences when delineating borders. This does not seem to be the case. Column (7) includes all the geographic, ecological, and natural resource measures. No factor other than the size of the ethnic area (and the presence of lakes) correlates with ethnic partitioning.

Pre-colonial Features and Ethnic Partitioning While at the time of colonial border design Europeans had limited understanding of local political geography, it is necessary to examine the association between ethnic partitioning and pre-colonial conflict, as recent cross-country works (Fearon and Laitin (2012)) and cross-regional studies in Africa show a significant legacy of conflict from the pre-colonial to contemporary times (Besley and Reynal-Querol (2014), Nunn and Wantchekon (2011)). Table 1 - Panel *B* examines the association between ethnic partitioning and two proxies of pre-colonial conflict, one based on wars between large pre-colonial African kingdoms and the other reflecting the intensity of slave trades.

Besley and Reynal-Querol (2014) show that contemporary conflict is higher in regions that suffered from pre-colonial wars (such as the Songhai-Gourma conflict in contemporary Mali in the end of the 15th century or the conflict between the Banyoro and Buganda kingdoms around 1600 in contemporary Uganda). Specification (1) shows the lack of a systematic association between ethnic partitioning and pre-colonial violence, as reflected by an indicator that takes the value one for ethnic homelands that experienced conflict over the period 1400 – 1700. Similarly, column (2) shows that ethnic partitioning and proximity to the nearest pre-colonial conflict are not related (the results are similar with log distance). This pattern suggests that ethnic partitioning captures a potential source of contemporary conflict distinct to that emphasized by Besley and Reynal-Querol (2014).

Africa experienced conflict during the slave trades, as the most common method of enslavement was “*through raids and kidnapping by members of one ethnicity of another or even between members of the same ethnicity*” (Nunn and Puga (2012)). Nunn and Wantchekon (2011) document a negative correlation between historical enslavement and proxies of social capital, arguing that the slave trades spurred animosity between African ethnicities. Similarly, Djankov and Reynal-Querol (2010) present cross-country evidence of a significant association between enslavement and civil war. In column (3) we regress ethnic partitioning on an indicator that equals one for ethnicities that were affected directly by the slave trades, while in column (4) we follow Nunn (2008) and use the log of one plus the number of slaves normalized by the area of each homeland. The coefficient on slave trades is quantitatively small and statistically insignificant, assuaging concerns that the ethnic partitioning index captures pre-colonial violence.

In columns (5) and (6) we associate ethnic partitioning to the homeland falling within the boundaries or being close to a large pre-colonial kingdom, using data from Besley and Reynal-Querol (2014). There is no systematic association between ethnic partitioning and the homeland being part of a large kingdom or the distance to the centroid of the closest pre-colonial kingdom.

In column (7) we associate ethnic partitioning to the pre-slave trade level of economic development

using an indicator that equals one if a city with population exceeding 20,000 people in 1400 *AD* was present in the historical homeland and zero otherwise (using data from Chandler (1987)). There is no evidence that ethnicities with historical urban centers were differentially treated during the early stage of colonization when borders were drawn in European capitals.

Further Checks In Appendix Table 8 we provide further evidence on the lack of systemic association between ethnic partitioning and various other measures of pre-colonial societal economic, political and cultural traits, such as the family organization, the presence of rules for inheritance, local elections, settlement patterns, using data from Murdock (1967) that are available for 450 – 490 groups.

Perhaps more importantly, in Appendix Table 9 we examine whether there are differences in various geographic, ecological, and natural resource characteristics between partitioned and non-split ethnic homelands in the same country. We report "balancedness tests" both for the full sample of country-ethnic homelands (the unit of analysis in our regressions linking ethnic partitioning to conflict) and for country-ethnic homelands close to the national border. The "similarity regressions" show that with the exception of (log) land area, there are no systematic differences in numerous observable characteristics between split and non-split ethnic homelands in the same country.

Summary Our results are consistent with the historical narrative on the largely arbitrary design of African borders. Yet, they are not a proof that *all* African borders were randomly designed; this is clearly not the case. What our large-scale econometric evidence suggests is that -on average- there are no systematic differences between partitioned and non-split ethnic homelands across observable characteristics that may independently affect conflict.

4 Ethnic Partitioning and Civil Conflict

This section reports the baseline estimates associating various aspects of civil conflict to ethnic partitioning. First, we present the conflict data. Second, we lay down the econometric specification and discuss estimation. Third, we report the benchmark estimates along with additional results.

4.1 Main Conflict Data

Our baseline data come from the Armed Conflict Location and Event Dataset (ACLED 4, Raleigh, Linke, and Dowd (2014)) that provides information on the location and some other characteristics of violent events across all African countries from 1997 till 2013. Political violence is understood as the use of force by a group with a political purpose or motivation. ACLED is by far the most complete geo-referenced conflict dataset; and while the data are noisy they have several desirable features.²³

²³Parallel works studying various driving forces of civil conflict using ACLED data, include Besley and Reynal-Querol (2014), Harari and La Ferrara (2014), and Berman, Couttenier, Rohner, and Thoenig (2014).

First, ACLED does not only record conflicts that take place within the context of a civil war, but also "*violent activity that occurs outside of civil wars, particularly violence against civilians, militia interactions, communal conflict and rioting*". The reporting of violence against civilians is particularly desirable, as Africa is plagued by civil strife that the standard data sources of civil war miss. Not only violence against civilians, such as child soldiering raids, rapes, and abductions is recurrent, but these incidents are often deadly, economically harmful, and devastating for the victims and the local community.

Second, ACLED categorizes conflict into four main categories, allowing for a closer examination. The main categories (percentage of total events) are: (1) Battles, either without change of control (32%) or where rebels or the government gain control (4%); (2) Violence against civilians, where armed groups (rebels, militias or government forces) attack unarmed civilians (31.5%); (3) Riots and protests (25%); and (4) Non-violent activities by violent actors, such as recruiting rallies (7.5%).

Third, ACLED reports an estimate of casualties and as such we can study the impact of partitioning on conflict intensity. Battles and violence against civilians are by far the most deadly types, as on average 45% of these incidents result in at least one fatality; in contrast, only 6.5% of riots and protests result in casualties and non-violent acts of conflict actors almost never result in casualties (less than 1%).

Fourth, the data classifies events by main conflict actor (government, rebels, militias, foreign interventions) and thus allows for a closer examination. This is important, as the available case studies suggest that ethnic partitioning is associated with discrimination from the government and thus more likely to ignite state-driven conflict. Moreover, one key mechanism linking partitioning to conflict is that nearby countries use split groups to launch proxy wars intervening in neighboring countries.

Original Sources. The data are based on a diverse set of sources. For almost all countries data come from more than ten different sources, while for the war-prone nations data come from around twenty sources. This diversity assuages concerns of systematic biases in reporting from government controlled media. The data are mostly based on international sources, such as the BBC (more than 10,000 incidents), Reuters (more than 5,000 incidents), the Associated Press (around 2,500 incidents), and the Agence France Press (around 5,000 incidents). A considerable fraction (around 10%) comes from media outlets from the United Kingdom, Portugal, Canada, the United States, and Australia. ACLED also relies on reports from NGOs, such as *Human Rights Watch* and *Amnesty International*, and the United Nations.²⁴ Even in cases of data coming from local sources (around 25% of the sample), most incidents come from pan-African news agencies, such as the *All Africa* network and independent newspapers.

²⁴Going over the documentation it seems that the data are based on verified information and not simply the reproduction of government statements and state press releases. For example, in Zimbabwe, besides information coming from BBC and Reuters many incidents are reported by the Zimbabwe Human Rights NGO Forum, a coalition of nineteen human rights NGOs that get data from their representatives on the ground. Likewise, for Somalia the international NGO CARE via its Security and Preparedness Project that "*aims to reduce the risks posed to programme personnel and assets of NGOs operating in Somalia*" is the source for several events. In Kenya many incidents are based on reports from the Peace and Development Network Trust, a local NGO, co-founded by Oxfam.

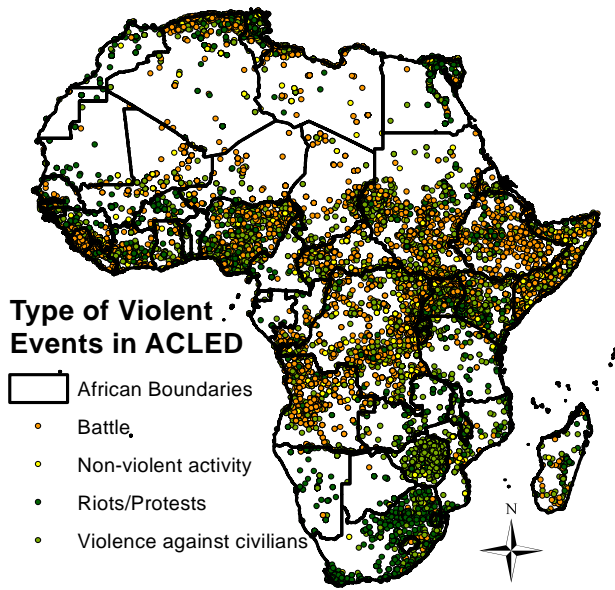


Figure 2a

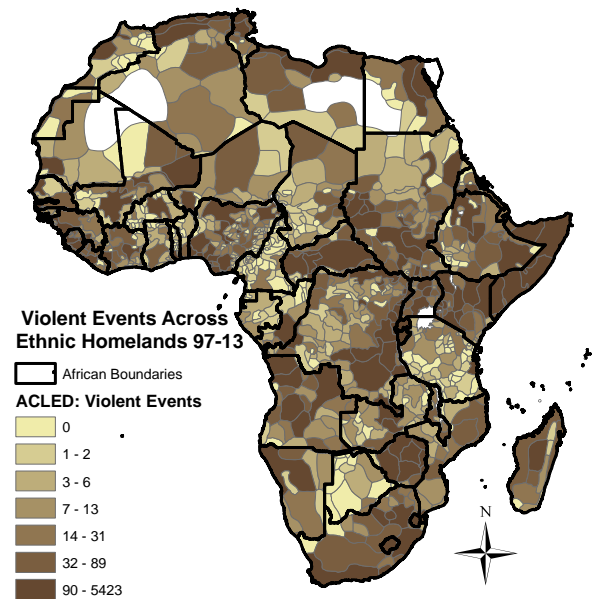


Figure 2b

Figure 2a maps the spatial distribution of conflict events. The map plots 64,650 high quality geo-precision events over the period 1997 – 2013. (In total there are 79,765 recorded events, but given the nature of our study, we drop low quality geo-precision events). There is significant heterogeneity in the incidence of political violence across countries (see Appendix Table 6). There are numerous events in Central Africa, mostly in Eastern Congo, Rwanda, Burundi, and Uganda. In Western Africa, conflict and political violence are mostly present in Nigeria and in Sierra Leone. Violence is also pervasive in Somalia, Ethiopia, and Zimbabwe. In contrast, there are few events in Botswana, Zambia, Tanzania, Namibia, and Gabon. There is considerable variation within countries. For example, while conflict incidence in Tanzania is low, there are several violent events along the border with Kenya and Rwanda. Likewise, most of the conflict in Angola is close to the northern border with Congo and in the Cabinda enclave. We examine separately the effect of partitioning on the number of battles and violence against civilians, as these are the most frequent and harmful types of conflict. The correlation between battles and violence against civilians is high, but far from perfect (0.64; see Appendix Table 7). For example, in Zimbabwe we observe lots of violence against civilians (3,701 incidents) and few battles (59). Conversely in Ethiopia and Angola we predominantly observe conflict between government and rebels rather than unilateral violence against civilians. In Uganda and Sudan battles go hand in hand with civilian violence (660 and 657 events in Uganda and 1411 and 1431 in Sudan, respectively). To construct conflict intensity at the country-ethnic homeland level, we project ACLED’s mapping of conflict (Figure 2a) on the intersection of Murdock’s ethnolinguistic map with contemporary borders (Figure 1b). Figure 2b portrays the spatial distribution of all civil conflict incidents at the country-ethnic homeland level.

4.2 Econometric Specification and Estimation

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

$$y_{i,c} = \exp(a_c + \gamma SPLIT_{i,c} + \phi SPIL_{i,c} + X'_{i,c}\Phi + \varepsilon_{i,c}). \quad (1)$$

The dependent variable, $y_{i,c}$, reflects civil conflict in the historical homeland of ethnic group i in country c . $SPLIT_{i,c}$ is a binary (dummy) variable that identifies partitioned ethnic areas in each country. Each partition of group i is assigned to the corresponding country c . For example, conflict in the part of the Lobi in Ivory Coast is assigned to Ivory Coast, while conflict in Lobi's homeland in Burkina Faso is assigned to Burkina Faso. At the country-ethnic homeland level, we have 518 partitioned areas and 694 non-split homelands. Given the lack of systemic correlation between the ethnic partitioning index and various historical, ecological, and geographical variables that correlate with conflict (Table 1 and the "balancedness tests" in Appendix Table 9), the γ coefficient captures the local average treatment effect of ethnic partitioning. To capture potential spatial externalities of ethnic partitioning, we augment the specification with a spillover index ($SPIL$), reflecting the fraction of adjacent groups that are partitioned. In the sample of 1212 country-ethnic areas, we have 274 areas without a partitioned neighbor, 146 areas are fully surrounded by split ethnicities, while most ethnic regions have at least one neighbor that has been split. [The mean (standard deviation) of $SPIL$ is 0.41 (0.32).]

Vector $X'_{i,c}$ includes log land area, log population according to the first post-independent census, indicators for the presence of rivers and lakes as well as several geographic, ecological, and natural resource measures. a_c denotes country-specific constants that account for all common for all ethnic homelands in a given country factors that may affect conflict, related for example to the type of colonial rule, colonial institutions, national policies, etc.

As the dependent variable is a count, we estimate negative binomial (NB) models with maximum likelihood (ML) (Wooldridge (2002), Cameron and Trivedi (2013)).²⁵ The negative binomial model accounts for the many zeros and for some extreme observations in the right tail of the distribution of the dependent variable. Following Cameron and Trivedi (2013), we use the unconditional negative binomial (NB2) model with country constants that allows for arbitrary over-dispersion.²⁶ To further account for outliers, we report specifications excluding homelands hosting the capital city or homelands where the dependent variable is in the top 1%. In the Appendix we also report fixed-effects Poisson ML estimates

²⁵Due to overdispersion in the dependent variable, specification tests reject the Poisson, favoring the negative binomial model. Across all specifications in Tables 2 – 5 the χ^2 value of the likelihood ratio test for the null hypothesis of a Poisson model (where the mean equals standard deviation) exceeds 100 [p – value : 0.00], and as such the negative binomial model is adopted. This LR test is asymptotically equivalent to a t -test on whether the alpha overdispersion parameter is zero.

²⁶This model reduces to the Poisson when the overdispersion parameter converges to zero. While the estimation of the fixed-effects suffers from the "incidental parameters" problem, the estimator has good properties (Greene (2005), Guimaraes (2008), Allison and Waterman (2002)). The NB2 model with fixed-effects has been used recently by Fisman and Miguel (2007), Aghion, Reenen, and Zingales (2013), and Bloom, Schankerman, and Reenen (2013).

dropping the top 5% of the dependent variable. This is useful both because it reassures that our estimates are not driven by outliers and because the fixed-effects Poisson model has good small sample properties (Wooldridge (2002)). To isolate the impact of ethnic partitioning on the likelihood of conflict, we always report linear probability models (MPM) where the dependent variable is a indicator that takes on the value of one if a country-ethnic area has been affected by any conflict over the sample period and zero otherwise. And we also estimate non-linear models focusing on conflict duration and fatalities.

4.3 Ethnic Partitioning and Civil Conflict

Table 2 reports the baseline country-fixed-effects specifications. Panel *A* gives (unconditional) negative binomial ML estimates with country constants focusing on conflict events, while Panel *B* gives country-fixed-effects OLS (linear probability model) estimates focusing on the likelihood of conflict.

Let us start with the NB specifications. The coefficient on the ethnic partitioning index in the parsimonious specifications in (1) and (2) is positive and more than two standard errors larger than zero. In column (3) we control for distance to the national border, the sea coast, the capital, and also include a capital city dummy and an indicator for coastal homelands. The coefficient on the ethnic partitioning index slightly increases and becomes more precisely estimated.²⁷ Column (4) includes controls reflecting geography-ecology (land quality for agriculture, elevation, malaria, an island dummy) and natural resources (indicators for diamond mines and oil deposits). We also include in the set of controls an indicator for the presence of a city in 1400. The coefficient on the partitioning index remains unaffected. This is consistent with our findings that partitioning is uncorrelated with these characteristics. In column (5) we drop outliers (top 1% of the dependent variable), while in column (6) we exclude regions where capitals fall. The estimates imply that partitioned ethnicities experience an increase of approximately 145 log points in the number of civil conflict incidents. This translates into an 57% increase in civil conflict activity ($\exp(0.45) - 1 = 0.568$) in areas where partitioned ethnicities reside (as compared to the homelands of non-split ethnicities). The effect of ethnic partitioning on civil conflict is quantitatively as strong as the effect of the petroleum indicator that enters with a significant coefficient (0.44 in specification (4)). The share of adjacent partitioned ethnicities (to the total number of neighboring ethnic areas) also enters with a positive estimate that is in some specifications statistically significant at the 90% level. This implies that the negative repercussions of ethnic partitioning are not present solely in split homelands, but also affect nearby regions. The coefficient on *SPIL* (0.43–0.49) suggests that conflict intensity is approximately 30% higher in the homelands of groups that are surrounded by 50% of split groups ($(\exp(0.47) - 1) * 0.5 = 0.30$).

In columns (7)-(12) we restrict estimation to areas close to the border, using the median distance from the centroid of each country-ethnic homeland to the border (61.3 *km*). This allows us to compare

²⁷Distance to the sea enters with a positive and significant estimate suggesting that there is less conflict in areas closer to the coast. Distance to the capital enters with a positive estimate suggesting that there is more conflict in regions further from the capitals, though the coefficient is not always significant. Distance to the border enters with a negative though insignificant coefficient. As violence against civilians, riots, and protests often take place in the capitals, the capital city indicator enters with a positive and highly significant coefficient.

conflict between partitioned and other at-the-border groups. Across all permutations the coefficient on the partitioning index is positive (around 0.60) and highly significant, reassuring that our estimates in the full sample are not capturing an overall border effect (which itself could reflect the impact of partitioning). The coefficient in the border sample is somewhat larger compared to the estimate in the full sample; yet a Hausman-Chow test shows that these differences are not statistically significant. The coefficient on *SPIL* is also stable (around 0.45), though standard errors increase and the estimate loses significance.²⁸

Table 2-Panel *B* reports linear probability model estimates. While by solely looking at the "extensive" margin, we do not exploit the richness of the data, we account for the non-linear nature of the dependent variable. Moreover, these specifications shed light on the margin along which ethnic partitioning operates. The coefficient of interest implies that the likelihood of conflict is approximately 7% – 8% higher for partitioned groups. The magnitude is similar (0.08 – 0.09) when we restrict estimation to groups close to the national border. The LPM estimates reveal sizable spillovers. In all specifications *SPIL* enters with a highly significant estimate (at the 99% level). The specification in (4) implies that compared to an ethnic homeland where none of the nearby groups is split ($SPIL = 0$), in ethnic homelands where half of adjacent groups are partitioned ($SPIL = 0.5$) the likelihood of conflict increases by 7%.

4.4 Ethnic Partitioning and Conflict Intensity

ACLED reports events that have resulted in some casualties as well non-deadly incidents (that, nevertheless, involve major conflict actors). We thus explored whether conflict severity is also higher in split homelands, focusing on the number of deadly incidents, the likelihood of deadly conflict, the total number of casualties and conflict duration. Doing so allows us to have a more comprehensive understanding of the impact of ethnic partitioning. Moreover, we account for the concern that due to the comprehensive nature of ACLED it puts together events of political violence that differ significantly in the intensity/casualties (Eck (2012)).

Table 3 reports the results. Columns (1) and (6) give negative binomial ML estimates looking solely on deadly events in the full and the border sample, respectively. The coefficient on *SPLIT* is 0.335 and 0.465, implying that deadly conflict is 40% – 60% higher in the homelands of partitioned ethnicities. This effect is quite similar to that of the petroleum dummy (coefficient 0.41). Columns (2) and (7) report linear probability estimates, where the dependent variable is a binary index identifying country-ethnic homelands that have experienced at least one deadly event over the sample period and zero otherwise. The estimates suggest that there is a 6% to 8% increased probability of a deadly incident in the homelands of split groups. Again the LPM estimates reveal sizable spillovers. Columns (3) and (8) report NB ML estimates associating total fatalities (aggregated across all events in all years for each country-ethnic area) to ethnic partitioning. Given the extreme skewness of casualties, the estimate is somewhat unstable;²⁹ yet *SPLIT*

²⁸The estimates in columns (10) and (11) are identical because all outliers (observations where all conflict incidents exceed the 99th percentile) are not in the border sample. The border sample is somewhat smaller than 606 observations, because there is no variability on ethnic partitioning for some countries when we zoom in the border.

²⁹The mean (median) of fatalities is 317 (3) with a standard deviation of 3,307. This is because of few extreme outliers.

enters with a significantly positive coefficient both in the full and the border sample. In columns (4) and (9) we focus on conflict duration, the number of years that there has been some conflict in each homeland, while in columns (5) and (10) we focus on the duration (in years) of deadly conflict. Since outliers is not an issue when we examine conflict duration and the mean - variance equality holds, we report country-fixed-effects Poisson ML estimates. There is a strong link between ethnic partitioning and conflict duration across all permutations. The estimate in column (10) implies that conflict duration is on average 55% higher in the homelands of partitioned ethnicities ($\exp(0.435) - 1 = 0.55$). The highly significant estimate on *SPIL* further suggests that if a given homeland is surrounded exclusively by split groups then conflict duration increases further by 60%, as compared to homelands where none of the adjacent groups is partitioned.

Example³⁰ Senegal offers an illustration of our results. ACLED records 565 events across its 12 constituent ethnic homelands. In the isolated Casamance region in the South, where the Diola and the Banyun ancestral homelands are split by the colonial border, we observe 154 and 85 events, respectively.³¹ This is 42.5% of all events (63% if we exclude the capital) though these two regions capture 11.2% of Senegal’s area and only 6% of the country’s population. Not only conflict is concentrated in Casamance but intensity is quite high. In these two split homelands, we observe 77 and 35 deadly events, out of a total of 182 deadly events (61.5% of all deadly events). The overwhelming majority of these events involve government troops (69 and 60 events) and/or rebels (114 in the homeland of the Diola and 63 in Banyun territory). And in both ethnic areas we observe conflict for 16 out of the 17 years between 1997 – 2013, much longer than in all other ethnic regions (with the exception of the capital, the mean is 5). The duration of deadly events is 15 and 12 years for Diola and Banyun, while the mean for the other ten ethnic areas is 3.2 years. Conflict is also quite deadly. Out of a total of 1,210 fatalities in Senegal, 540 (45%) and 354 (29%) have occurred in the two ethnic regions in Casamance.

4.5 Ethnic Partitioning and Type of Conflict

In Table 4 we take advantage of the fact that ACLED distinguishes between battles, violence against civilians, and riots and protests to examine the effect of ethnic partitioning on each of type of conflict separately. Panel *A* gives unconditional NB ML estimates with country constants and Panel *B* reports country-fixed-effects LS (linear probability model) estimates.

Battles Examples of battles between armed actors include the fights of the Lord’s Resistance Army (LRA), the Sudanese People’s Liberation Army, and Uganda’s People Defence Force (UPDF); the constant

For example, the threshold for the top 1% percentile is 435 and the maximum value is 107,554. See Appendix Table 2.

³⁰In the Supplementary Appendix we discuss as a case study conflict in Eastern Congo (the most conflict intense region in our sample).

³¹The contemporary border follows the 1886 convention between Portuguese Guinea and (French) Senegal. The seeds of the current conflict may be traced in early 1900s, when the Diolas opposed the French, who fought the local resistance and imprisoned King Sihalebe and other chiefs. Even during the colonial era, the Diolas were organizing their resistance from the Portuguese side of the border (Tomas (2006)).

fighting between the Rwandan forces against Hutu rebels in Rwanda and in Eastern Congo; and the battles between Kikuyu rebel groups against Maasai militias. Battles result usually (on average 47%) in fatalities; for example, ACLED description indicates that in a single event in September 1999 the Ugandan army killed 42 Pian warriors, coming from the Karamojong ethnic group that is split between Uganda, Sudan, and Kenya. The specifications in columns (1) and (4) show that battles between government forces, militias, and rebel groups are more pervasive in the homelands of partitioned groups; the NB estimate suggests that partitioned groups experience on average 55% – 60% ($exp(0.45) - 1 = 0.57$) more battles. The LPM estimate on *SPLIT* is also positive and significant implying that battles are 9% more likely to take place on the historical homeland of partitioned ethnicities. *SPIL* also enters with a positive (though noisy) estimate, suggesting that in some cases there are also spillovers.

Civilian Violence A useful feature of the ACLED is the reporting of violence against the civilian population, a socially and economically devastating aspect of conflict that the commonly employed civil war datasets leave unaccounted. Approximately 20% of violence against civilians is perpetrated by government troops, 20% from rebel groups with the remaining events coming from militias. Examples include the raids of the Janjaweed militias against civilians in Darfur; the assaults of the Central Intelligence Organization in Zimbabwe; the killings of civilians in Northern and Western Rwanda by the Interahamwe Hutu militias. Violent events include the burning of churches, hostage-taking and child-soldiering raids by rebels in Nigeria and in Sierra Leone. Going over the narratives of each event reveals that they may be also quite devastating (43% of these events result in at least in one fatality). For example, in a single day in Eastern Congo in May 1997 "*ADLF rebels moved in and took control of Mbandaka slaughtering 200 Rwandan Hutu refugees*". The NB estimate in the full sample (in column (2)) implies that there are 55% ($exp(0.43) - 1 = 0.54$) more violent events against civilians in the homelands of partitioned ethnicities. Restricting estimation to ethnic regions close to the national border (in (5)) yields somewhat higher estimates (although the difference is not statistically significant). The LPM estimate on *SPLIT* is 0.052 and 0.065 in the full and the border sample, respectively. While the estimate is insignificant (t -stat around 1.5), it implies that the likelihood of violence against the civilian population is approximately 5% – 6% higher in the homeland of split ethnicities. The linear probability models reveal sizable spillovers. The coefficient on *SPIL* is 0.18 and highly significant (at the 99% level) in both specifications. The estimate implies that a one standard deviation (0.34) increase in the share of adjacent groups that are split increases the likelihood of violence against the civilians by 6.5%.

Riots and Protests In columns (3) and (6) we examine the link between ethnic partitioning and riots and protests. Protests and riots are (relatively) non-violent events taking place usually in major urban centers. Examples include the protests in South Africa during and after the Marikana miners' strike (in 2012) and the protests in Zimbabwe during the periods of hyperinflation and food shortages

(2005 – 2009) and the Arab Spring events. Given the nature and usual location of these events perhaps it is not surprisingly to find that there is no significant association between partitioning and riots and protests.

4.6 Sensitivity Checks

We have performed numerous sensitivity checks -that for brevity we report and discuss in the on-line Supplementary Appendix. Specifically: (1) As the number of conflict events recorded in ACLED increased considerably in 2011, 2012 and 2013, we repeated estimation focusing on the period 1997 – 2010. (2) We estimated the specifications with the conditional negative binomial model of Hausman, Hall, and Griliches (1984) that parameterizes the over-dispersion parameter rather than the mean. (3) To further account for outliers we dropped the top 5% of the dependent variable and estimated country fixed-effects Poisson ML models as in this case the mean variance equality approximately holds. (4) We did not account for spillovers. (5) We reclassify groups into split and non-split using a 5% land area threshold. (6) We augment the specification with a 3rd (or a 4th) order polynomial in distance to the border to further account for unobserved effects that vary smoothly close to the border. (7) We include ethnic-family fixed effects (on top of country fixed-effects) to account both for local conditions and broad cultural, institutional, and other hard-to-observe ethnic-family factors. (8) To account for different colonial and post-independence policies we drop iteratively ethnic homelands from each of the five main African regions. (9) We estimate formal spatial models that account for spillovers and other features of the nearby ethnic regions. (10) We account for conflict spillovers from regions in the same country or/and the same ethnolinguistic cluster. (11) We control for the historical legacy of violence in the pre-colonial period and the slave trade period. (12) We condition on regional proxies of income finding that the negative effect of ethnic partitioning works on top of any impact on average regional well-being (overall there is a small and usually insignificant effect of partitioning on luminosity and other proxies of regional income). Across all these permutations the coefficient on the ethnic partitioning index retains its economic and statistical significance. And most models reveal sizable spillovers of ethnic partitioning.

4.7 Heterogeneous Effects

We also searched for potential heterogeneous effects of ethnic partitioning. We discuss and report these results in the Supplementary Appendix. In particular, we explored whether the coefficient on partitioning varies by ethnic features related to: (1) the population share of the group in the country; (2) the population of a group’s co-ethnics on the other side of the border; (3) the share of adjacent groups that belong to the same ethnic family; (4) the share of groups in the country that belong to the same ethnic family; (5) the share of partitioned groups among neighboring groups; (6) whether the bilateral border that intersects a given group is straight or wiggly; (7) whether a group is split within the same colonizer or between different colonizers, and (8) the number of countries a split group falls to. The analysis does not reveal

much heterogeneity. We also examined heterogeneity of ethnic partitioning with respect to nationwide features, related to ethnic/linguistic/religious diversity and polarization, country size, and whether the country is landlocked or not. The effect of ethnic partitioning on conflict is quite homogeneous across these country features. (There is some weak evidence that the impact of ethnic partitioning on civil conflict is particularly deleterious for ethnicities in landlocked countries.)

5 Mechanisms and Channels

In this Section we utilize ACLED's classification of events by the conflict actors involved to shed some light on the mechanisms at work. We then complement the analysis using geo-referenced data on major civil wars using an alternative conflict database (UCDP GED).

5.1 Hypotheses

H1 : African historiography suggests that ethnic partitioning is primarily associated with state conflict; if governments discriminate against partitioned ethnicities and/or attempt to push them on the other side of the border and split groups respond by organizing rebel forces to counter such policies, then one should observe in the homelands of split groups more conflict between government troops and rebels.

H2 : If neighboring countries intervene to assist their co-ethnics across the border either because the latter are targets of state violence and discrimination or they are simply used instrumentally as a pretext to destabilize the neighboring governments, then one would expect a link between ethnic partitioning and military interventions from adjacent countries.

5.2 Ethnic Partitioning and Conflict by Key Actors (ACLED)

ACLED categorizes events by the main conflict actors. There are 8 categories: (1) government forces; (2) rebels groups, "*defined as political organizations whose goal is to counter an established national governing regime by violent acts. Rebel groups have a stated political agenda for national power, are acknowledged beyond the ranks of immediate members, and use violence as their primary means to pursue political goals*"; (3) political and (4) ethnic militias, groups that "*are not subsumed within the category of government or opposition, but are noted as an armed associated wing*"³²; (5) riots and (6) protests, defined "*as violent and non-violent spontaneous groupings (respectively)*"; (7) one-sided violence against civilians (that are perpetrated by militias, rebels and government forces); and (8) outside/external forces, which are either international peace-keeping armed forces (United Nations or African Union) or government troops of neighboring countries.³³

³²As ACLED notes, "*militias are more difficult to assess since they can be created for a specific purpose or during a specific time period (i.e., Janjaweed) and may be associated with an ethnic group, but not entirely represent it (i.e., Kenyan Luo ethnic militias).*"

³³Appendix Table 5 gives the distribution of all and deadly conflict events by the interaction of conflict actors.

Table 5 reports NB-ML (in Panel *A*) and linear probability model (in Panel *B*) estimates linking conflict by each actor separately to ethnic partitioning.³⁴ We merge rebels and militias (since there is some degree of arbitrariness distinguishing between them) and distinguish between foreign interventions by neighboring countries and by international forces.

Government Armed Forces. The specifications in columns (1) and (5) reveal a strong link between ethnic partitioning and conflict where government troops are involved. The NB estimate in the full sample implies that there are 70% more conflicts with government troops participation in the homelands of partitioned ethnicities. The linear probability estimates imply that the likelihood of conflict involving state forces is 11%–12.5% higher in the ancestral homelands of partitioned groups. The LPM specifications also indicate sizable spillovers of ethnic partitioning to nearby ethnic areas. The estimate in (5) implies that a one standard deviation (0.34) increase in the share of adjacent groups that are split increases the likelihood of violence against the civilians by 7%.

Rebels and Militias. There is a significant association between ethnic partitioning and conflict where rebels and militias participate (columns (2) and (6)). The coefficient on *SPLIT* in the LPM suggests that the probability of conflict involving rebel groups is approximately 6.5%–8.5% higher in the homelands of partitioned ethnicities. As ACLED requires that rebel groups explicitly challenge national authority via violent means, these results further show that the partitioning - conflict link operates (to some extent at least) via groups challenging the government. In line with this interpretation when we separately focus on rebels and militias, we find a much stronger effect of ethnic partitioning for conflict of rebel groups as compared to militias. For example, in the full sample the NB-ML (linear probability model) estimate with rebels is 0.88 (0.087), while for militias is 0.23 (0.056). Moreover, events featuring rebels are quite deadly, especially when fighting government troops.

External Interventions from Neighboring Countries. In columns (3) and (7) we examine whether interventions from neighboring countries are related to ethnic partitioning. While we do report NB specifications (where the ethnic partitioning enters with a highly significant coefficient), we focus on the linear probability model estimates, as the dependent variable is highly skewed (with many zeros). Overall 269 country-ethnic homelands (22.2%) experienced an invasion from an adjacent country. Examples include the interventions of Ugandan and Rwandan troops in DRC, the fighting of Military Forces of Kenya with rebels in Southern Somalia, and the interventions of the military forces of Chad in Mali and the Central African Republic. The estimates (both in the full and the border sample) imply that there is a 7% increased likelihood of a military intervention from a neighboring country in the homelands of split groups. A simple test of means illustrates the regression estimates. In the border sample (606 observations) that consists of 416 partitioned homelands and 190 non-split ethnic areas, we observe interventions from neighboring countries in 113 country-ethnic regions (19%). Yet 94 of these regions that experience a foreign intervention

³⁴Since we have already reported specifications with riots and protests and violence against civilians (in Table 4) we do not repeat them in Table 5.

(83%) are partitioned, while overall 69% of homelands are partitioned. Not only foreign interventions are more likely to occur in partitioned homelands, but they are much more common. In the border sample, we observe 708 events of interventions from nearby countries in the homelands of split ethnicities, as compared to 100 events in the homelands of non-split groups.

Interventions by International Forces. ACLED also reports conflicts associated with international, usually peace-keeping forces, such as the United Nations/African Union Hybrid Operation in Darfur, the Economic Community of West African States Monitoring Group and United Nations Mission in Sierra Leone, Liberia, and Guinea at the end of the civil war, the UN Mission in DRC (1999 – 2010), as well as the military interventions of NATO and Western countries in Libya. We thus examined whether ethnic partitioning correlates with such type of outside interventions –that we use as a "placebo" as a priori the this type of conflict is not likely to be associated with partitioning. We focus on the LPM estimates as the variable is highly skewed. The coefficient on *SPLIT* is small (0.0065 and 0.035 in the full and border sample) and statistically indistinguishable from zero.

Summary The conflict actor evidence supports two often-cited links between ethnic partitioning and conflict. First, political violence involving state forces and rebel groups opposing the government lies in the heart of the conflicts erupting across split homelands. This pattern is consistent, albeit not a proof, of ethnic-based discrimination from the state. Second, the results that military interventions from neighboring governments' troops are more common in the homelands of split groups shows that partitioned ethnicities are often used to destabilize nearby countries and that partitioned ethnicities, when faced with suffocation, often rearm and regroup within their ancestral homeland on the other side of the border.

5.3 Ethnic Partitioning and Conflict by Key Actors (UCDP GED)

5.3.1 Data

To shed further light on the link between ethnic partitioning and the type of conflict we used data from the Uppsala Conflict Data Program Geo-referenced Events Dataset (UCDP GED) that covers the period 1989 – 2010 (Sundberg, Lindgren, and Padsokocimaite (2010), Sundberg and Melander (2013)). The UCDP focuses on deadly incidents associated with civil wars, as identified by the UCDP-PRIO Armed Conflict Database; this database identifies civil wars when conflict results in at least 25 fatalities in a given year. UCDP conflicts are grouped into three mutually exclusive categories.

(1) State-based 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 in one calendar year"*. Examples of state-based conflict where ethnic partitioning seems to have played a role include the fights between the Ethiopian government and rebels seeking self-determination in the Afar and the Ogaden region³⁵ and the fighting between Tuareg

³⁵UCDP summarizes: *"The conflict over the status of the Ogaden region in south-eastern Ethiopia has seen the active*

rebels with government troops in Mali and Niger seeking autonomy and independence. In total there are 7,512 events taking place mostly in Algeria, Somalia, Angola, Ethiopia, Sierra Leone, and Burundi (see Appendix Table 6).

(2) One-sided violence defined as the "*use of armed force by the government of a state or by a formally organized group against civilians which results in at least 25 deaths in a year*". Examples include the public killings and executions against civilians carried in Angola by both government troops and UNITA (mostly during 1997 – 2003); and ethnic-based violence during the transition to democracy in Congo (1997 – 1999). In total there are 5,219 with DRC, Sierra Leone, South Africa and Sudan being the countries where one-sided civilian violence takes the maximum value.

(3) Non-state conflict takes place "*between two organized armed groups, neither of which is the government of a state, which results in at least 25 battle-related deaths in a year.*" Examples include conflict between the various factions of the African National Congress (ANC) (e.g., "Greens" and "Reds") and between the ANC and other anti-apartheid movements (Azanian People's Organization, United Democratic Movement) during the democratic transition in South Africa; and conflict between various militias in Kenya's Rift Valley. The data also includes many cases of relatively low-intensity conflict between neighboring pastoral groups over land and livestock. In total UCDP includes 3,645 events though 60% of these incidents are from South Africa during the transition. Appendix Figures 4 and 5a – 5c portray the distribution of conflict events across ethnic homelands according to UCDP-GED.

5.3.2 Results

Table 6 reports the results linking various aspects of civil conflict to ethnic partitioning.³⁶

State-driven Conflict. Let us start with the results in Panel A that associate state-driven conflict to partitioning. The NB estimate in column (1) is positive and significant implying that state conflict intensity is 65% higher in the homelands of split, as compared to non-split, groups ($\exp(0.50) - 1 = 0.65$). So, the impact of ethnic partitioning is quite similar to the one obtained with ACLED's estimates on conflict involving government troops. (The correlation between state-driven conflict in UCDP and battles in ACLED is 0.72.) The coefficient on *SPLIT* retains significance when we restrict estimation to border areas (in (4)). Columns (2) and (5) give LPM estimates. The estimate implies that the likelihood of state-driven conflict is 5% – 8% higher in the homelands of partitioned ethnicities. Columns (3) and (6) report NB estimates (specification tests reject the mean - variance equality) associating the duration (in years) of state-driven conflict with ethnic partitioning. *SPLIT* enters with a highly significant coefficient both in the full and the border sample; the estimate implies that conflict duration is higher by approximately 75% in the homelands of partitioned ethnic groups. Across all specifications *SPIL* also enters with a positive

participation of a number of rebel groups with more or less diverging goals. It has also been closely entangled with the inter-state conflict between Ethiopia and Somalia, which concerned the same territory, to which the latter laid claim. The first rebel group to emerge was the Ogaden Liberation Front in 1963. Over a decade later WSLF (Western Somali Liberation Front) appeared and was subsequently followed by ONLF (Ogaden National Liberation Front) and AIAI (al-Itihad al-Islami)."

³⁶Appendix Table 9 reports tests of means and medians across ethnic homelands.

coefficient -that is statistically significant in the full sample. This suggests that there are sizable spillovers. We also estimated NB models linking fatalities from state-driven conflict to ethnic partitioning (results not shown). While the casualty estimates are highly skewed and the results somewhat unstable (especially in the border sample) there is a strong link between ethnic partitioning and fatalities. For example, the coefficients (s.e.) on *SPLIT* and *SPIL* in the specifications in the full sample are 0.78 (0.35) and 1.74 (0.70), respectively, implying economically large effects.

One-sided Violence. Panel *B* reports estimates associating one-sided violence to ethnic partitioning. The coefficient on *SPLIT* is positive both in the NB specifications with the number of incidents ((1) and (4)) and duration ((3) and (6)) and the LPM specifications focusing on the likelihood of conflict (in (2) and (5)). Yet the estimates do not pass the standard significance thresholds. The same applies to the positive coefficient on the spillover measure (*SPIL*). The results are similar with fatalities (results not shown). *SPLIT* and *SPIL* enter with positive though weakly significant coefficients (*t*-stat around 1.5). Overall, the UCDP GED data point out that there is a weak link between ethnic partitioning and one-sided violence. The key difference with ACLED -where ethnic partitioning appears to have a stronger impact on civilian violence- is that UCDP covers way fewer events, as it records events where conflict actors can be succinctly identified and linked to a major civil war. In contrast, ACLED reports many events that are not part of a full-scale civil war and/or incidents where the opposing groups are not clearly identified. This difference in the coverage of events involving violence against civilians also explains the modest correlation (0.43) between the ACLED-based and UCDP GED-based measures.

Non-State Conflict. In Panel *C* we look at on non-state-driven conflict. Across all permutations the ethnic partitioning index enters with a small, unstable, and statistically insignificant coefficient. This should not be surprising as the non-state conflict events predominantly reflect communal violence in South Africa after the fall of the apartheid (that is clearly unrelated to ethnic partitioning) and other low-intensity communal violence mostly between pastoral groups over livestock and land.³⁷

Example. The UCDP GED mapping of conflict in Senegal illustrates the results. 36 and 45 from a total of 91 events of state-driven conflict have taken place in the homelands of the partitioned Banyun and the Diola, respectively. Those events have resulted into 322 and 427 fatalities, out of a total of 827 deaths from state-driven conflict. Looking at one-sided violence yields a similar picture. There have been 41 and 38 violent events against the civilian population in the homelands of the Banyun and the Diola resulting into 125 and 118 fatalities. One-sided violence in Senegal outside these two areas is minimal (there are just 15 events in all other homelands). The UCDP documentation states that all these events involved either state conflict against the MFDC or one-sided violence of the MFDC against the civilians. [Humphreys and Mohamed (2005) provide an eloquent overview of conflict in Casamance and the role of ethnic partitioning, pointing to discrimination from the national government in Dakar against the Diolas,

³⁷Non-state conflict is weakly correlated with all other aspects of conflict both in the UCDP and the ACLED (correlations around 0.15). See Appendix Table 7.

and the resulting insurgency in the region.]

Accounting for Measurement Error in the Conflict Data. Arguably both UCDP and ACLED conflict incidents are measured with noise. To account for error-in-variables in the Supplementary Appendix we combine the two datasets in a variety of ways to obtain a more precise picture on the presence of conflict. For example, we defined an ordered variable ranging from 0 – 2 that takes on the value of two when both databases record conflict in a given ethnic area, a value of one when country-ethnic areas have experienced conflict based on only one of the two databases, and a value of zero when both datasets report no conflict. And we defined binary measures that equal one when an ethnic region experiences conflict with both databases (and zero otherwise). The link between ethnic partitioning and civil conflict is quite strong.

Summary The results with the UCDP GED dataset that focuses on deadly incidents associated with major civil wars further point out that ethnic partitioning matters crucially for state-driven conflict where government troops fight against rebels and militias. The analysis also shows that there is a rather weak link between ethnic partitioning and one-sided violence; and no association with conflict where only non-state actors are involved. Thus the results reveal that ethnic partitioning has for the most part adversely impacted the relations between split-by-the-border groups and the national government.

6 Partitioning and Ethnic Power Relations

Data The strong link between ethnic partitioning and state-driven conflict (both with the ACLED and UCDP GED) suggests that political discrimination against split groups may be one of the channels at work. While this result is by no means a proof, it accords well with anecdotal evidence and case studies, pointing out that split groups face discrimination from the central government, as they are often excluded from power and are targets of abusive policies (Asiwaju (1985)).

In an effort to shed light on the ethnic partitioning - discrimination - conflict nexus, we used data from the Ethnic Power Relations (EPR) dataset (Wimmer, Cederman, and Min (2009)) that provides ethnic-specific information on political representation, discrimination, and ethnic wars for a large number of countries during the postwar period. EPR focuses on politically relevant ethnic groups and relies on expert input to assess formal and informal degrees of political participation and exclusion along ethnic lines. An ethnic group is classified as politically relevant *"if at least one significant political actor claims to represent the interests of that group in the national political arena, or if members of an ethnic entity are systematically and intentionally discriminated against in the domain of public politics."* EPR provides information on 758 politically relevant ethnic groups in 134 states. The coverage for Africa spans 40 countries and 196 groups. Using a multitude of sources, we linked the EPR groups to the Murdock ethnicities matching the 196 groups to 593 Murdock (1959) groups.³⁸ 234 groups (39.5%) are partitioned,

³⁸Such sources include the *Joshua Project*, the *Ethnologue* dataset and the *A-MAR* project. In several instances the matching procedure is straightforward. For example, the "San (Bushmen, Basarwa)" group in Namibia in the EPR is linked to those

while the remaining 359 (60.5%) are non-split. Using the latest vintage of the EPR data and focusing on the post-independence period (1960 – 2010), we examine the link between ethnic partitioning, political discrimination and ethnic-based civil wars.

Discrimination EPR codes ethnic discrimination when "*group members are subject to active, intentional, and targeted discrimination with the intent of excluding them from both regional and national power. Such active discrimination can be either formal or informal. Formal discrimination legally limits access to government positions to citizens who speak a certain mother tongue, display certain phenotypic features, or are members of certain religious groups. Informal discrimination actively and intentionally inhibits individuals with certain ethnic backgrounds from rising within the ranks of government.*"

During the post-independence period, out 110 groups have being discriminated by the national government at some point (18.5%). This average, however, masks considerable differences between partitioned and non-split ethnicities. 58 of the 234 partitioned groups have been subject to political discrimination by the government (25%), while the likelihood of discrimination for non-split groups is ten percentage points lower, 15%, as only 52 of the 359 non-split groups faced discrimination. Examples of partitioned groups that have faced political discrimination include the Bushmen (San/Basarwa) groups in Botswana that have faced restrictions on residence, limited access to the civil service and higher offices and no recognition of their traditional chiefs, something allowed for Tswana groups. Another example is the Karamojong in Uganda, a group split along the Kenyan-Ugandan border that has suffered from land confiscation, abuses, and raids by government forces (MercyCorps (2011)).

Table 7, columns (1)-(4), reports OLS (linear probability model) estimates that explore formally the association between ethnic discrimination and partitioning. Column (1) reports the unconditional specification; so the estimate on the partitioning index gives the simple test of means. In column (2) we include country constants. The coefficient in *SPLIT* is 0.078 and significant, implying that within countries split groups are more likely to be marginalized from the political arena. Controlling for group size in terms of (log) land area and (log) population in 1960 and the rich array of location and geographic traits does not alter the economic or statistical significance of the estimate. The coefficient in column (4) implies that the likelihood of discrimination is approximately 7% for partitioned, as compared to non-split, groups. *SPIL* enters with a small and statistically indistinguishable from zero coefficient. Perhaps not surprisingly there are no spatial spillovers on political discrimination from partitioning.

groups in Murdock (1959) that belong to the "Bushmen and their kin" cluster. In other instances, the matching is less straightforward. For example, in Nigeria EPR lists the "Hausa-Fulani and Muslim Middle Belt" as a single category. In this case we used the *A-MAR* correspondence created by Wilkenfeld, Brancati, Fearon, Gurr, Laitin, Pate, and Saideman (2014). We also took advantage of the fact that the EPR has a geo-referenced version so one may identify the corresponding location of groups on the Murdock map by intersecting the two maps. This method is the least satisfactory and, hence, was only used for roughly 10% of cases. Results are unaffected if such matches are excluded.

Ethnic Wars Another useful feature of the EPR is the matching of ethnic groups to civil wars that have an explicit ethnic dimension. The coding of civil wars is based on the widely-used UCDP/PRIO Armed Conflicts Data Set (Petter, Wallensteen, Eriksson, Sollenberg, and Strand (2002)). Civil war is defined as an armed and organized confrontation between government troops and rebel organizations, or between army factions, that reaches an annual battle-death threshold of 25 people. From this dataset, EPR identifies ethnic wars as those that "*typically involve conflicts over ethno-national self-determination, the ethnic balance of power in the government, ethno-regional autonomy, ethnic or racial discrimination (whether alleged or real), and language and other cultural rights.*" EPR linked ethnic conflicts to the politically relevant ethnic category. We thus examined the link between major ethnic civil wars and ethnic partitioning. Using ethnic wars as the outcome of interest has straightforward benefits. First, instead of relying on the incidence of conflict in a given location, we directly assess whether members of partitioned groups have participated in an ethnic war *irrespective* of the location of actual violence. By doing so, we account for the imprecision in the anthropological maps and the geo-referenced conflict data. Second, we look at major breakouts of violence. Third, the data has a long time horizon covering the entire post-independence period for most countries. Fourth, by focusing on politically relevant groups, we check the robustness of our findings by focusing on ethnic entities with a presence in the political spectrum.

A simple tabulation reveals the stark disparities in ethnic conflict between split and non-split groups. On the one end, 72 out of the 234 split groups (31%) have participated in a civil war with an ethnic dimension. On the other end, 69 out of the 359 non-split groups (19%) have participated in an ethnic war. Examples of split groups that have been involved in major ethnic wars include the Afar groups in North-East Ethiopia, which since the mid-1970s faced large-scale discrimination and marginalization policies by the central government. For example, in 1975 the *Dergue* administration (that ruled Ethiopia from 1974 till 1987) nationalized all land and effectively annulled the de facto autonomy of the Afar leading to a secessionist rebellion under the Afar Liberation Front (Vaughan (2003)). With the assistance of Somalia, the various Somali clans (such as the Esa, the Ogaden, and the Ishaak) residing in the Southern of Ethiopia initiated a major war with the Ethiopian government in 1977 trying to obtain more autonomy or join Somalia. And while the Somali-Ethiopia war ended in 1978 the separatist Ogaden National Liberation Front continued its guerrilla war for the next ten years.

In columns (5)-(8) of Table 7 we formally assess the impact of ethnic partitioning on major ethnic wars. Column (1) tests for cross-sectional mean differences in the likelihood of ethnic wars between split and non-split groups. In column (2) we add country fixed effects, while in (3) and (4) we account for differences across groups in terms of population in 1960, land area, and the presence of water bodies (river or lake), location and geography. The pattern is robust. The estimate in the specification with the rich set of controls suggests that within African countries a partitioned group has roughly an 11% increased likelihood of participating in an ethnic war, as compared to non-split groups. This magnitude is quite similar to the linear probability estimate focusing on conflict where government forces are involved using

the ACLED (0.11, column (1), Panel *B* of Table 5). In line with the baseline coefficients (in Tables 3 – 5) the EPR-based estimates also reveal sizeable externalities. The coefficient on *SPIL* suggests that a one-standard-deviation increase (0.25) in the share of adjacent partitioned groups increases the probability of involvement in an ethnic war for a given group by roughly 7% contributing significantly to the eruption of ethnically tainted civil wars.

Political Discrimination and Ethnic Wars A tenable hypothesis of the literature on the consequences of the Scramble for Africa is that ethnic partitioning has led to political discrimination along ethnic lines spurring conflict. While most likely the link between political discrimination and civil strife reflects a two-way relationship, we examined whether the impact of ethnic partitioning on ethnic wars weakens once we account for discrimination by augmenting the specification with the binary political discrimination index. Column (9) in Table 7 reports the results. The coefficient on the discrimination proxy is positive (0.61) and highly significant, showing, unsurprisingly, that groups facing political repression are also more likely to be involved in ethnic conflicts (and vice versa).³⁹ The coefficient on the ethnic partitioning index retains its statistical significance though it drops from 0.11 to 0.066. This implies that the link between ethnic partitioning and conflict works partly via political discrimination.

A Note It is perhaps instructive to keep in mind that the disproportionate incidence of discrimination among split groups does not imply that the latter are never part of the ruling elite. For example, the Yakoma in Central African Republic, the Oroma and the Tigray in Ethiopia, and the Alur, the Madi and the Lugbara in Uganda although they have been subject to discrimination and have participated in ethnic wars, they also seem to have taken part in various governments over time (either as junior or senior partners). This result is similar to Francois, Rainer, and Trebbi (2015), who show that even groups that face discrimination from the national government at some point in time, do have some representation in the government at some other time (for example, the Ewe in Ghana). Econometrically, this oscillation of several split groups between a state of discrimination at some point and partners in government coalitions at other instances translates into an insignificant relationship between partitioning and the probability of a group having ever been a senior, junior or dominant partner in the government. This pattern suggests that the political status of partitioned groups in African countries is more nuanced than a picture of constant repression and discrimination.

Robustness In the Supplementary Appendix we provide some necessary sensitivity checks, showing that the ethnic partitioning, discrimination and ethnic war link is present: (1) when we use an alternative ethnic partitioning index that takes the value of one for groups where at least 5% of their ancestral

³⁹In total 64 out of the 593 groups (10.8%) have experienced discrimination from the national government and have engaged in an ethnic civil war. Yet the likelihood that a partitioned ethnicity will be subject to both discrimination and ethnic war is 15.4% (36 out of 234), while the corresponding likelihood for non-split groups is 7.8% (28 out of 359).

homeland falls into more than one country and (2) when we drop iteratively a different African region.

7 Ethnic Partitioning and Individual Well-Being

So far our evidence suggests that partitioned ethnic groups are more likely to engage in conflict (predominantly against the government), experience violence against the civilian population, and political discrimination. Nevertheless, it does not shed light on the well-being of members of split groups. This is precisely the dimension along which we augment our analysis in this section. Specifically, we employ micro-level data from the Demographic and Health Surveys (DHS) to examine how individuals belonging to partitioned groups fare economically compared to citizens from non-split groups in the same country.

Exploiting individual-level variation has some straightforward advantages that complement the analysis at the ethnic homeland level. First, we can directly assess whether individuals identifying with split groups under-perform compared to those from non-split ethnicities using direct measures of well-being and self-reported ethnic affiliation. Second, we can account for a host of individual level characteristics, so as to better isolate the importance of ethnic partitioning. Third, since we observe people residing both within and outside their group’s historical homeland, we can evaluate whether ethnic partitioning has negative repercussions for those identifying with partitioned groups irrespective of their actual residence or whether the negative effects are limited to individuals residing in partitioned territories; finding the former would be more consistent with identity-based rather than location-specific discrimination.

7.1 Data and Specification

The DHS are based on nationally representative samples and include information on households’ wealth, education, occupation, and health. We use *all* available surveys with gps (global positioning system) coordinates and information on ethnic identification of the respondents. Our sample comprises 20 countries and covers 88,171 male respondents.⁴⁰ We focus on two outcome variables. First, we use the composite wealth index constructed by the DHS team (range from 0 to 5) that reflects a household’s access to basic public goods (electricity, sewage system, and piped water) and economic status, including ownership of various assets (such as television and radio). Second, we use an ordered education variable that assigns the respondent’s level of schooling into six categories; no education, incomplete primary, complete primary, incomplete secondary, complete secondary and higher.

Our empirical specification reads:

$$y_{i,e,r,c} = a_c + \beta SPLIT_e + X'_{i,e,r,c} \Phi + Z'_{i,r,c} \Gamma + \zeta_{i,e,r,c}. \quad (2)$$

⁴⁰The countries and respective survey years are Benin in 2001, Burkina Faso in 2010, Central African Republic in 1994, Ethiopia in 2011, Ghana in 2008, Guinea in 2005, Kenya in 2008, Mali in 2006, Mozambique in 2011, Malawi in 2010, Namibia in 2000, Niger in 1998, Senegal in 2010, Sierra Leone in 2008, Togo in 1998, Uganda in 2011, the Democratic Republic of Congo in 2007, and Zambia in 2007.

The dependent variable, $y_{i,e,r,c}$, reflects economic conditions and education of individual i ; that self-identifies with ethnic group e and resides in enumeration area (village/town/city) r in country c . $X'_{i,e,r,c}$ is a vector of individual characteristics; in most specifications we include a set of (62) year of birth dummies, a set of (6) marital-status fixed effects, and a set of (7) religion fixed effects. $Z'_{i,r,c}$ includes location controls (at the enumeration area). We also include a dummy variable that identifies respondents residing outside their ethnicity's ancestral homeland. All specifications include country-specific constants, a_i , that capture among other things survey differences across countries. $SPLIT_e$, the variable of interest, is an indicator that takes on the value of one if individual i identifies with ethnicity, e , that has been partitioned across different countries. Overall 38,887 individuals come from partitioned ethnicities (44%) while 49,284 individuals (56%) identify with non-split ones (see Appendix Table 36). To account for spatial correlation and the fact that the split indicator takes on the same value for individuals belonging to the same split (or non-split group) we cluster standard errors at the ethnic identity and ethnic homeland level.

7.2 Baseline Estimates

Table 8 columns (1)-(6) report the baseline country fixed-effects estimates linking the composite DHS wealth index to ethnic partitioning. The coefficient on the partitioning index in (1), that only includes a set of country-specific constants and an indicator reflecting whether the individual currently resides outside his ancestral homeland, is negative and highly significant. This implies that individuals who identify with split ethnicities have on average lower access to public goods and worse living conditions. The coefficient retains significance when we condition on individual characteristics (in (2)). In column (3) we control for distance to the national border, the capital, and the coastline. We also include an indicator for households residing in capital cities ("location controls").⁴¹ The coefficient on the ethnic partitioning index retains significance at standard confidence levels. The estimate implies that the composite wealth index is -0.20 points lower for individuals of split ethnic groups, as compared to individuals from non-split ethnicities in the same country; this translates into a standardized "beta" coefficient of 0.07, quite similar to that reported by Nunn and Wantchekon (2011), who estimate the negative repercussions of the slave trades on trust using a similar to ours approach. Another way to assess the economic importance of ethnic partitioning on individual well being is to compare its magnitude with another significant covariate. The "standardized" beta coefficient on the capital city indicator is around 0.11; this implies that in absolute magnitude the impact of ethnic partitioning is comparable though somewhat smaller to that of residing in the capital. In columns (4)-(6) we limit our attention to enumeration areas close to the border, using the median value of distance to the border (80 kilometers). In all specifications the coefficient on the ethnic partitioning index is negative and significant at the 99% confidence level. The estimate is similar to the analogous estimates in the full sample, reassuring that the results do not capture an overall border effect.

⁴¹The coefficients on distance to the capital and distance to the sea are negative and significant; the coefficient on distance to the border is positive and (marginally) significant.

In columns (7)-(12) education serves as the dependent variable.⁴² The estimates imply that - conditional on location and various individual characteristics- individuals from partitioned ethnicities have on average less formal education than individuals from non-split groups. The standardized "beta" coefficient on the ethnic partitioning index in these specifications is around 0.05, implying an economic effect as large as that of residing in capitals (the "beta" coefficient on the capital city dummy is 0.04).

7.3 Ethnic Identity and Ethnic Homelands

Taking stock so far, we have two main findings. First, political violence is more frequent in the homelands of split ethnicities and, second, the scars of partitioning can be traced in the livelihoods of members of partitioned groups. Weaving these two observations together, begets the question whether ethnic partitioning depressed standards of living for everybody currently residing in split homelands (i.e., residents in split homelands are worse off irrespective of their ethnic affiliation) or whether it is the individuals belonging to split ethnicities that experience disproportionately lower standards of living irrespective of their place of residence. The narrative in African studies suggests that it is the latter that is going on. To shed light on this hypothesis we augmented the baseline specification with a location-based indicator variable of ethnic partitioning that takes on the value of one for individuals residing in the homeland of partitioned ethnicities. Doing so allows us to disentangle the importance of ethnically identifying with a split group from that of residing in the homeland of a partitioned ethnicity. Note that for individuals residing in their ancestral homelands these two indexes (origin-based and location-based) coincide.⁴³

Table 9 reports the results. In the simple country-fixed-effects specification in (1) both the origin-based split indicator and the location-based one enter with negative and significant estimates (-0.31 and -0.23 , respectively). In (2) we control for individual and location characteristics. The coefficient on the location-based partitioning index declines and becomes insignificant; in contrast, the estimate on the ethnicity-based partitioning index retains its statistical and economic significance. In columns (3)-(4) we restrict estimation to areas close to the border. The coefficient on the identity-based index remains negative and is stable (ranging between -0.29 and -0.245); this suggests that even when we focus on border areas and control for numerous individual and location features, members of partitioned ethnicities have on average worse living conditions compared to those identifying with non-partitioned ones. In columns (5) and (6) we introduce an interaction term between the partitioned ethnic identity indicator and partitioned ethnic location dummy. So this dummy variable identifies those individuals that both reside in partitioned homelands and self-identify with split groups. The interaction enters with an insignificant coefficient

⁴²For clarity and to minimize the "incidental parameter" problem we report OLS estimates (see Nunn and Wantchekon (2011) for a similar approach). Ordered probit and ordered logit estimates that explicitly account for the nature of the outcome variable yield similar estimates.

⁴³The correlation of the two partitioning measures is 0.40. Out of 88,171 individuals, 25,631 (29%) self-identify with a split group and reside in partitioned homelands; 36,694 (41.5%) individuals neither identify with a split ethnicity nor do they reside in partitioned homelands; 13,256 individuals (15%) reside in partitioned homelands but identify with a non-split group; and 12,590 (14.5%) individuals reside in non-partitioned ethnic homelands but identify with split groups.

suggesting that the negative impact of partitioning is not magnified or mitigated for members of split groups residing in split homelands.

The pattern is similar with education (in (7)-(12)). The identity-based measure of partitioning enters all permutations with a significantly negative estimate (around -0.18), implying that individuals from partitioned ethnicities have on average lower levels of educational attainment. In contrast, the location-based index of partitioning is not systematically linked to differences in educational outcomes.

Further Evidence In the Supplementary Appendix we investigate further the inferior economic status of individuals identifying with split groups. First, to further account for location effects we estimated enumeration-area fixed-effects specifications. Within the 7,898 surveyed villages/towns/cities members of split groups have systematically worse access to utilities and appear less educated. Second, to further account for unobserved differences between "movers" and "non-movers" we estimated the link between ethnic partitioning and economic performance focusing only on non-indigenous currently residing in partitioned and non-partitioned homelands separately. The link between ethnic partitioning and public goods is present for all types of individuals. Third, we examined persistence and inertia examining the impact of ethnic partitioning for "young" and "old" respondents. The negative impact of partitioning on individual outcomes does not differ across young and old generations pointing to its ongoing importance.

8 Conclusion

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

In the first part of our paper we formally explore the nature of African political boundaries. Utilizing information on the spatial distribution of ethnicities at the time of colonization, we associate ethnic partitioning to various geographic, ecological, and natural resource indicators as well as social and institutional ethnic-specific traits including proxies of pre-colonial conflict and early development. With the exception of the size of the historical homeland and the presence of lakes, there are no significant differences between partitioned and non-partitioned ethnicities. Hence, our results offer support to the claim of the African historiography on the accidental drawing of colonial and, consequently, national borders, at least with respect to ethnic partitioning.

Second, we examine the effect of ethnic partitioning on civil conflict, as this has been conjectured as the major consequence of the Scramble for Africa. We exploit a new dataset spanning the universe of African ethnic areas that reports precisely geocoded information for 64,650 conflict events over the period 1997–2013. The database is useful in examining the long-run impact of ethnic partitioning, as it reports the location of battles between government forces, militias, and rebel groups and incidents involving violence against civilians. Exploiting within-country variation and focusing on ethnic homelands is appropriate for

Africa given the salience of ethnicity. We uncover that battles between armed groups, as well as violence against civilians are higher in the homelands of split groups. Ethnic partitioning is also associated with more deadly incidents over prolonged periods of violence.

Third, we exploit the richness of the data to shed some light on the mechanisms at work. We present evidence suggesting that neighboring countries often use the homelands of partitioned groups to stage military interventions on the other side of the border. Moreover, ethnic partitioning is strongly associated with state-driven conflict where government troops, state-backed militias and rebels opposing the government. In contrast, there is no link between partitioning and conflict between non-state actors as well as riots and protests that are mostly concentrated in the capital centers.

Fourth, using data from the Ethnic Power Relations database that reports information on ethnic-based discrimination from the government and civil wars with an explicit ethnic angle we examine in detail the ethnic partitioning - discrimination - conflict nexus. The evidence is clear-cut. Partitioned ethnicities are significantly more likely to experience political discrimination at various government levels and are more likely to participate in major ethnic-based civil wars.

Fifth, we shift our attention to the individual. Using micro-data from the Demographic and Health Surveys, including more than 85,000 respondents in 20 African countries, we show that individuals identifying with partitioned groups have fewer household assets, poorer access to public goods, and lower education, as compared to individuals from non-split ethnicities. This pattern is not due to a generalized decline in standards of living of *all* households residing in split homelands; rather it is driven by the poorer economic performance of members of split ethnicities irrespective of their actual residence. Taken together, the evidence from the EPR and DHS point out that the consequences of ethnic partitioning are not circumscribed by the contours of a given ancestral ethnic homeland but have significant repercussions for the members of partitioned groups irrespective of their whereabouts.

Our work calls for future research examining the impact of ethnic partitioning on other aspects of economic and institutional development and on the precise mechanisms via which the Scramble for Africa has affected long-run countrywide economic performance.⁴⁴ Moreover, since border artificiality and ethnic partitioning are not an exclusive African phenomenon subsequent works could also study their effect in other world regions, such as the Middle East and the Caucasus, where besides ethnicity religion is also salient.

⁴⁴For example, ethnic partitioning may offer some economic benefit inasmuch as ethnic networks facilitate cross-border trade. As more bilateral border-specific trade data become available one may be able to quantify this dimension, see Aker, Klein, O'Connell, and Yang (2010) for such evidence from the Niger–Nigeria border.

References

- ACEMOGLU, D., S. JOHNSON, AND J. A. ROBINSON (2001): “The Colonial Origins of Comparative Development: An Empirical Investigation,” *American Economic Review*, 91(5), 1369–1401.
- (2002): “Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution,” *Quarterly Journal of Economics*, 107(4), 1231–1294.
- (2005): “Institutions as a Fundamental Cause of Long-Run Growth,” in *Handbook of Economic Growth*, ed. by P. Aghion, and S. N. Durlauf, pp. 109–139. Elsevier North-Holland, Amsterdam, The Netherlands.
- ACEMOGLU, D., T. REED, AND J. A. ROBINSON (2014): “Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone,” *Journal of Political Economy*, 122(2), 319–368.
- AGHION, P., J. V. REENEN, AND L. ZINGALES (2013): “Innovation and Institutional Ownership,” *American Economic Review*, 103(1), 277–304.
- AKER, J. C., M. W. KLEIN, S. A. O’CONNELL, AND M. YANG (2010): “Borders, Ethnicity and Trade,” Working Paper, NBER, No. 15960.
- ALESINA, A., W. EASTERLY, AND J. MATUSZESKI (2011): “Artificial States,” *Journal of the European Economic Association*, 9(2), 246–277.
- ALLISON, P. D., AND R. P. WATERMAN (2002): “Fixed-Effects Negative Binomial Regression Models,” *Sociological Methodology*, 32(1), 247–265.
- ASIWAJU, A. (1985): “The Conceptual Framework,” in *Partitioned Africans*, pp. 1–18. St. Martin Press, New York.
- BATES, R. H. (1981): *States and Markets in Africa*. University of California Press, Berkeley, CA.
- BERMAN, N., M. COUTTENIER, D. ROHNER, AND M. THOENIG (2014): “This Mine is Mine! How Minerals Fuel Conflicts in Africa,” working paper, Department of Economics, University of Lausanne.
- BESLEY, T., AND T. PERSSON (2011): *Pillars of Prosperity. The Political Economics of Development Clusters*. Princeton University Press, Princeton, NJ.
- BESLEY, T., AND M. REYNAL-QUEROL (2014): “The Legacy of Historical Conflicts. Evidence from Africa,” *American Political Science Review*, 108(2), 319–336.
- BLATTMAN, C., AND E. MIGUEL (2010): “Civil War,” *Journal of Economic Literature*, 48(1), 3–57.

- BLOOM, N., M. SCHANKERMAN, AND J. V. REENEN (2013): “Identifying Technology Spillovers and Product Market Rivalry,” *Econometrica*, 81(4), 1347–1393.
- BOSKER, M., AND J. DE REE (2014): “Ethnicity and the Spread of Civil War,” *Journal of Development Economics*, 108, 206–221.
- BRECKE, P. (1999): “Violent Conflicts 1400 A.D. to the Present in Different Regions of the World,” Paper prepared for the 1999 Meeting of the Peace Science Society in Ann Arbor, Michigan.
- CAMERON, A. C., AND P. K. TRIVEDI (2013): *Regression Analysis of Count Data Book*, vol. Econometric Society Monograph No. 53. Cambridge University Press, Cambridge, UK, 2nd edition edn.
- CAMERON, C. A., J. GELBACH, AND D. MILLER (2011): “Robust Inference with Multi-Way Clustering,” *Journal of Business and Economic Statistics*, 28(2), 238–249.
- CHANDLER, T. (1987): *Four Thousand Years of Urban Growth: An Historical Census*. Edwin Mellon Press, Lewiston, NY.
- CICCONE, A., AND M. JAROCINSKI (2010): “Determinants of Economic Growth: Will Data Tell?,” *American Economic Journal Macroeconomics*, 2(2), 222–246.
- CIOFFI-REVILLA, C. (1996): “Origins and Evolution of War and Politics,” *International Studies Quarterly*, 40(1), 1–22.
- COLLIER, P. (2007): *The Bottom Billion*. Oxford University Press, Oxford, United Kingdom.
- COLLIER, P., AND A. HOFFLER (2007): “Civil War,” in *Handbook of Defense Economics, Volume 2, Defense in a Globalized World*, ed. by T. Sandler, and K. Hartley, pp. 711–740. Elsevier, North Holland, Amsterdam and Oxford.
- COLLIER, P., AND N. SAMBANIS (2005): *Understanding Civil War. Volume 1: Africa*. World Bank, Washington, DC.
- CONLEY, T. G. (1999): “GMM Estimation with Cross Sectional Dependence,” *Journal of Econometrics*, 92(1), 1–45.
- DJANKOV, S., AND M. REYNAL-QUEROL (2010): “The Causes of Civil Wars,” mimeo UPF.
- DOWDEN, R. (2008): *Africa: Altered States, Ordinary Miracles*. Portobello Books Ltd, London, UK.
- EASTERLY, W., AND R. LEVINE (2012): “The European Origins of Economic Development,” NBER Working Paper No. 18162.

- ECK, K. (2012): “In Data We Trust? A Comparison of UCDP GED and ACLED Conflict Events,” *Cooperation and Conflict*, 47(1), 124–141.
- ENGLEBERT, P., S. TARANGO, AND M. CARTER (2002): “Dismemberment and Suffocation: A Contribution to the Debate on African Boundaries,” *Comparative Political Studies*, 35(10), 1093–1118.
- ESTEBAN, J., L. MAYORAL, AND D. RAY (2012): “Ethnicity and Conflict: An Empirical Study,” *American Economic Review*, 102(4), 1310–1342.
- ESTEBAN, J., AND D. RAY (2008): “On the Saliency of Ethnic Conflict,” *American Economic Review*, 98(6), 2185–2202.
- (2011): “A Model of Ethnic Conflict,” *Journal of the European Economic Association*, 9(3), 496–521.
- ETHNOLOGUE (2005): *Languages of the World*, SIL International; Fifteenth edition.
- FEARON, J. D., AND D. D. LAITIN (2003): “Ethnicity, Insurgency and Civil War,” *American Political Science Review*, 97, 75–90.
- (2011): “Sons of the Soil, Migrants, and Civil War,” *World Development*, 39(2), 199–211.
- (2012): “How Persistent Is Armed Conflict?,” mimeo, Stanford University.
- FISMAN, R., AND E. MIGUEL (2007): “Corruption, Norms, and Legal Enforcement: Evidence from Diplomatic Parking Tickets,” *Journal of Political Economy*, 115(6), 1020–1048.
- FRANCOIS, P., I. RAINER, AND F. TREBBI (2015): “How Is Power Shared In Africa?,” *Econometrica*, Working Paper, NBER, No. 18425.
- FUCHS-SCHUNDELN, N., AND T. A. HASSAN (2015): “Natural Experiments in Macroeconomics,” in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and H. Uhlig, vol. 2.
- 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.
- GREENE, W. H. (2005): *Econometric Analysis*. Prentice Hall, Upper Saddle River, NJ, fourth edition edn.
- GUIMARAES, P. (2008): “The Fixed Effects Negative Binomial Model Revisited,” *Economic Letters*, 99(1), 63–66.

- HARARI, M., AND E. LA FERRARA (2014): “Conflict, Climate and Cells: A Disaggregated Analysis,” mimeo MIT and Bocconi University.
- HAUSMAN, J., B. H. HALL, AND Z. GRILICHES (1984): “Econometric Models for Count Data with an Application to the Patents-RD Relationship,” *Econometrica*, 52(4), 909–938.
- HEGRE, H., AND N. SAMBANIS (2006): “Sensitivity Analysis of Empirical Results on Civil War Onset,” *Journal of Conflict Resolution*, 50(4), 508–535.
- HERBST, J. (2000): *States and Power in Africa*. Princeton University Press, Princeton, NJ.
- HOROWITZ, D. L. (1985): *Ethnic Groups in Conflict*. University of California Press, Berkeley and Los Angeles, California.
- HUBER, J. D., AND L. MAYORAL (2014): “Inequality, Ethnicity and Civil Conflict,” Working Paper, Department of Political Science, Columbia University.
- HUILLERY, E. (2009): “History Matters: The Long Term Impact of Colonial Public Investments in French West Africa,” *American Economic Journal - Applied Economics*, 1(2), 176–215.
- HUMPHREYS, M., AND H. A. MOHAMED (2005): “Senegal and Mali: A Comparative Study of Rebellions in West Africa,” in *Understanding Civil War: Evidence and Analysis Vol 1 Africa*, ed. by P. Collier, and N. Sambanis. The World Bank, Washington DC, USA.
- JEDWAB, R., AND A. MORADI (2015): “The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa,” *Review of Economics and Statistics*, forthcoming.
- 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.
- MERCYCORPS (2011): “Uganda Conflict and Market Assessment - Karamoja,” Report.
- MEREDITH, M. (2005): *The State of Africa: A History of the Continent since Independence*. The Free Press, London, United Kingdom.
- MICHALOPOULOS, S. (2012): “The Origins of Ethnolinguistic Diversity,” *American Economic Review*, 102(4), 1508–1539.
- MICHALOPOULOS, S., AND E. PAPAIOANNOU (2013): “Pre-colonial Ethnic Institutions and Contemporary African Development,” *Econometrica*, 81(1), 113–152.
- (2014): “National Institutions and Subnational Development in Africa,” *Quarterly Journal of Economics*, 129(1), 151–213.

- MONTALVO, J.-G., AND M. REYNAL-QUEROL (2005): “Ethnic Polarization, Potential Conflict, and Civil Wars,” *American Economic Review*, 95(3), 796–816.
- MURDOCK, G. P. (1959): *Africa: Its Peoples and their Culture History*. McGraw-Hill.
- (1967): *Ethnographic Atlas: A Summary*. University of Pittsburgh Press, PA.
- NUNN, N. (2008): “The Long Term Effects of Africa’s Slave Trades,” *Quarterly Journal of Economics*, 123(1), 139–176.
- (2013): “Historical Development,” in *Handbook of Economic Growth, Volume 2*, ed. by P. Aghion, and S. Durlauf, vol. 2. Elsevier, Amsterdam, Netherlands.
- NUNN, N., AND D. PUGA (2012): “Ruggedness: The Blessing of Bad Geography in Africa,” *Review of Economics and Statistics*, 94(1), 20–36.
- NUNN, N., AND L. WANTCHEKON (2011): “The Slave Trade and the Origins of Mistrust in Africa,” *American Economic Review*, 101(7), 3221–3252.
- O’BRIEN, P. (1999): *Atlas of World History*. Oxford University Press, Oxford, United Kingdom.
- 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.
- POSNER, D. N. (2005): *Institutions and Ethnic Politics in Africa*. Cambridge University Press, New York, NY.
- RALEIGH, C., A. LINKE, AND C. DOWD (2014): “Armed Conflict Location and Event Dataset (ACLED) Codebook Version 3,” Working Paper, Centre for the Study of Civil War, International Peace Research Institute, Oslo (PRIO).
- ROSS, M. L. (2012): *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton University Press, Princeton, NJ.
- SUNDBERG, R., M. LINDGREN, AND A. PADSKOCIMAITE (2010): “UCDP GED Codebook version 1.0-2011,” Department of Peace and Conflict Research, Uppsala University.
- SUNDBERG, R., AND E. MELANDER (2013): “Introducing the UCDP Georeferenced Event Dataset,” *Journal of Peace Research*, 50(4), 523–532.
- THOMSON, A. (2010): *An Introduction to African Politics*. Routledge, London and Nyew York.
- TOMAS, J. (2006): “The Traditional Authorities Cross the National Border: Opposing Views on the Role of the Religious Leaders of the Jola Huluf and Ajamaat of the Lower Casamance (1886-1909),” *Africana Studia*, 9(1), 73–97.

- TOUVAL, S. (1967): “The Organization of African Unity and African Borders,” *International Organization*, 21(1), 102–127.
- VAUGHAN, S. (2003): “Ethnicity and Power in Ethiopia,” Ph.D. thesis, The University of Edingurgh.
- WANTCHEKON, L., M. KLASNJA, AND N. NOVTA (2015): “Education and Human Capital Externalities: Evidence from Colonial Benin,” *Quarterly Journal of Economics*, 130(2).
- WESSELING, H. L. (1996): *Divide and Rule: The Partition of Africa, 1880-1914*. Praeger, Amsterdam, Netherlands.
- WILKENFELD, J. K. B. J., D. BRANCATI, J. FEARON, T. GURR, D. LAITIN, A. PATE, AND S. SAIDEMAN (2014): “A-MAR (All-Minorities at Risk): Addressing the Selection Bias Issue,” mimeo.
- WIMMER, A., L.-E. CEDERMAN, AND B. MIN (2009): “Ethnic Politics and Armed Conflict. A Configurational Analysis of a New Global Dataset,” *American Sociological Review*, 74(2), 316–337.
- WOOLDRIDGE, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.

9 Data Appendix

9.1 Ethnic Partitioning

Ethnic Partitioning Index (SPLIT): Indicator variable that equals 1 if at least 10% of the historical homeland of an ethnic group is partitioned into different countries. We also construct an alternative partitioning index that equals 1 if at least 5% of the historical homeland of an ethnic group is partitioned into different countries. *Source: Calculated intersecting Murdock's (1959) ethnic map of Africa with the Digital Chart of the World (DCW) shapefile. The latter contains the polygons delineating the international boundaries in 2000.*

Ethnic Partitioning Spillovers (SPIL): The share of adjacent groups that are partitioned to the total number of adjacent (neighboring) ethnic homelands.

9.2 Civil Conflict Variables (ACLED)

All Civil Conflict Incidents: Sum of all high-precision civil conflict incidents for each of a total of 1,212 country-ethnic homelands over the period 1997–2013. There are 8 event types. (1) Battles without change of control; (2) Battles where rebel groups gain control of the location; (3) Battles where the government regains control of a location; (4) Headquarter of base establishments, where rebel groups establish (via violent or non-violent means) their base; (5) Non-violent conflict events where rebel groups, militias or government forces proceed in non-violent actions (without active fighting) that are however within the context of an ongoing civil conflict and dispute (e.g., recruitment drives, incursions or rallies); (6) Riots and protests; (7) Violence against civilians, where armed groups (rebels, militias or government forces) attack unarmed civilians; (8) Non-Violent transfer of control. See Section 2 for details. *Source: ACLED 4.*

Conflict Indicator: Dummy variable that takes on the value of one if a country-ethnic area has experienced at least one high-precision conflict incident (of any type) over the period 1997–2013. *Source: ACLED 4.*

Deadly Civil Conflict Incidents: Sum of all high-precision civil conflict incidents that have resulted in at least one fatality for each of a total of 1,212 country-ethnic homelands. *Source: ACLED 4.*

Deadly Conflict Indicator: Dummy variable that takes on the value of one if a country-ethnic area has experienced at least one high-precision deadly conflict incident (of any type) over the period 1997–2013. *Source: ACLED 4.*

Duration Civil Conflict: Duration in years that a country-ethnic homeland has been in conflict (using all types of conflict events) over the period 1997–2013. *Source: ACLED 4.*

Duration Deadly Civil Conflict. Duration in years that a country-ethnic homeland has been in conflict that has resulted in at least one fatality (using all types of deadly conflict events) over the period 1997–2013. *Source: ACLED 4.*

Total Fatalities: Number of fatalities for each of a total of 1,212 country-ethnic homelands using

all types of conflict incidents. *Source: ACLED 4.*

Battles: Total number of battles between two violent armed groups at the ethnic homeland (in each country for partitioned ethnicities). Battles include armed conflict where a control of the contested location does not change and conflict events resulting in a territorial change of control. We aggregate the data at the ethnic homeland level and at the country-ethnic homeland level. See Section 2 for details. *Source: ACLED 4.*

Violence against Civilians: Total number of violent events against civilians at the ethnic homeland (in each country for partitioned ethnicities). Violence against civilians occurs when any armed/violent group attacks unarmed civilians. Rebels, governments, militias, rioters can all commit violence against civilians. We aggregate the data at the at the country-ethnic homeland level. *Source: ACLED 4.*

Riots and Protesters: Total number of events corresponding to riots and protests at the ethnic homeland (in each country for partitioned ethnicities). We aggregate the data at the at the country-ethnic homeland level. *Source: ACLED 4.*

9.3 Civil Conflict Variables (UCDP)

State-driven Conflict: Number of events associated with "*use of armed force by two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths in one calendar year of a state or by a formally organized group against civilians which results in at least 25 deaths in a year*". The data cover the period 1989 – 2010 and include "*all events corresponding to years where the actors and conflicts did not exceed 25 battle-related deaths threshold required for inclusion in the aggregate datasets (i.e. includes inactive years)*." *Source: UCDP GED 1.5.*

One-sided Violence: Number of events associated with "*use of armed force by the government of a state or by a formally organized group against civilians which results in at least 25 deaths in a year*". The data cover the period 1989 – 2010 and include "*all events corresponding to years where the actors and conflicts did not exceed 25 battle-related deaths threshold required for inclusion in the aggregate datasets (i.e. includes inactive years)*." *Source: UCDP GED 1.5*

Non-state-driven Conflict: Number of events associated with "*use of armed force between two organized armed groups, neither of which is the government of a state, which results in at least 25 battle-related deaths in a year*." The data cover the period 1989 – 2010 and include "*all events corresponding to years where the actors and conflicts did not exceed 25 battle-related deaths threshold required for inclusion in the aggregate datasets (i.e. includes inactive years)*." *Source: UCDP GED 1.5.*

Conflict Indicator: For each type of conflict (state-driven, one-sided violence, and non-state-driven) we defined a conflict indicator (dummy) variable that equals one if a country-ethnic area has experienced at least one (high geo-precision) conflict event over the period 1989 – 2010 and zero otherwise. *Source: UCDP GED 1.5.*

Conflict Duration: For each type of conflict (state-driven, one-sided violence, and non-state-

driven) we defined variables measuring the number of years that a country-ethnic area has experienced at least one (high geo-precision) conflict event over the period 1989 – 2010. *Source: UCDP GED 1.5.*

9.4 Control Variables

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

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

Lake Indicator: Dummy variable that takes on the value one for (country) ethnic homelands with a major lake and zero otherwise. *Source: Constructed using the "Inland water area features" dataset from Global Mapping International, Colorado Springs, Colorado, USA.*

River Indicator: Dummy variable that takes on the value one for (country) ethnic homelands with a major river and zero otherwise. *Source: Constructed using the "Inland water area features" dataset from Global Mapping International, Colorado Springs, Colorado, USA.*

Elevation: Average value of elevation in kilometers. *Source: National Oceanic and Atmospheric Administration (NOAA) and U.S. National Geophysical Data Center, TerrainBase, release 1.0 (CD-ROM), Boulder, Colorado.*

Land Suitability for Agriculture: Average value of land (soil) quality for cultivation. The index is the product of two components reflecting the climatic and soil suitability for cultivation. *Source: Michalopoulos (2012); Original Source: Atlas of the Biosphere.*

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

Distance to the National Border: The geodesic distance of the centroid of the historical homeland of each ethnic group (or each country-ethnic area) from the nearest national border, measured in 1000s of km's. *Source: Global Mapping International, Colorado Springs, Colorado, USA.*

Distance to the Capital: The geodesic distance of the centroid of the historical homeland of each ethnic group (or each country-ethnic area) from the capital city, measured in 1000s of km's. *Source: Global Mapping International, Colorado Springs, Colorado, USA.*

Distance to the Sea: The geodesic distance of the centroid of the historical homeland of each ethnic group (or each country-ethnic area) from the nearest coastline, measured in 1000s of km's. *Source: Global Mapping International, Colorado Springs, Colorado, USA.*

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

Coastal Indicator: Dummy variable that takes on the value one for country-ethnic homelands that are adjacent to the coast and zero otherwise (when the area is landlocked).

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

Diamond: Indicator variable that takes on the value of one if a diamond mine is in the historical homeland of an ethnic group and zero otherwise. *Source: Map of Diamond Resources. Peace Research Institute of Oslo (PRIO).*

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 (or each country-ethnic area) and zero otherwise. *Source: Chandler (1987)*

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

9.5 Pre-colonial Features

Pre-colonial Conflict: Using data on the centroid (latitude-longitude) of 91 major conflict incidents in Africa in the pre-colonial period (between 1400 – 1700) we define a dummy variable that takes on the value of one for ethnic homelands that experienced such a conflict in their territory, as defined by Murdock’s map. Following Besley and Reynal-Querol we also use the distance of each ethnic homeland to the centroid of the closest pre-colonial conflict. *Source: Besley and Reynal-Querol (2014); original source: Brecke (1999) and Cioffi-Revilla (1996).*

Pre-colonial Kingdoms and Empires: We define a dummy variable that takes on the value of one for ethnic homelands that were part of a large pre-colonial kingdom and empire. We also use the distance of each ethnic homeland to the centroid of the closest pre-colonial kingdom/empire. *Source: Besley and Reynal-Querol (2014); original source: O’Brien (1999).*

Slavery: Number of persons of each ethnic group that were shipped during the trans-Atlantic and Indian Ocean slave trades. We define a dummy variable that takes on the value of one for ethnic homelands that were directly affected by the slave trades and zero otherwise. Also following Nunn (2008) we use the log of one plus the number of slaves per 1000 of square kilometers and a dummy variable that takes on the value of one for ethnic groups that were affected by the slave trades (and zero otherwise). *Source: Nunn (2008) and Nunn and Wantchekon (2011).*

9.6 DHS Data

Composite Wealth Index: The wealth index is a composite measure of almost all household assets and utility services including country-specific items. The wealth index is calculated using easy-to-collect data on a household’s ownership of selected assets, such as televisions and bicycles; materials used for housing construction; and types of water access and sanitation facilities. Generated with a statistical procedure known

as principal components analysis, the wealth index places individual households on a continuous scale of relative wealth within a country. More details are available here: <http://www.measuredhs.com/topics/Wealth-Index.cfm> and here: <http://www.measuredhs.com/pubs/pdf/CR6/CR6.pdf>. *Source: Demographic and Health Surveys (<http://www.measuredhs.com/>)*.

Education: Ordered, ranging from 0 to 5 education variable that assigns the respondent's level of formal schooling into six categories. A score of 0 indicates "no education"; a score of 1 indicates "incomplete primary"; a score of 2 indicates "complete primary"; a score of 3 indicates "incomplete secondary"; a score of 4 indicates "complete secondary"; and a score of 5 indicates "higher education". *Source: Demographic and Health Surveys. <http://www.measuredhs.com/>*.

Ethnic Partitioning - Identity: Indicator variable that equals 1 for individuals that self-identify with a partitioned ethnic group. To construct this dummy variable we link the ethnic affiliation from the DHS to the ethnic groups in Murdock's map.

Ethnic Partitioning - Location: Indicator variable that equals 1 if an individual resides in an ethnic homeland that at least 10% of it is partitioned into different countries.

Non-Indigenous (Mover) Indicator: Dummy variable that identifies individuals residing outside their ethnicity's ancestral homeland.

Marital Status: A vector of six variables capturing marital status. The categories are: *Source: Demographic and Health Surveys. <http://www.measuredhs.com/>*

Age: A vector of age bracket constants (fixed-effects) of household head. The 10 categories are: 15-19; 20-24; 25-29; 30-34; 35-39; 40-44; 45-49; 50-54; 55-59; 60-64. *Source: Demographic and Health Surveys. <http://www.measuredhs.com/>*

Religion: A vector of seven religion constants (fixed effects). The 7 categories are: Traditional, Islam, Catholic, Protestants, Other Christian, Other, None. *Source: Demographic and Health Surveys. <http://www.measuredhs.com/>*

Distance to the Capital City: The geodesic distance from the location (gps coordinates) of each household to the capital city of the country it belongs to. *Source: Calculated using the Haversine formula.*

Distance to the Sea Coast: The geodesic distance from the location (gps coordinates) of each household to the nearest coastline. *Source: Global Mapping International, Colorado Springs, Colorado, USA. Series name: Global Ministry Mapping System. Series issue: Version 3.0.*

Distance to the National Border: The geodesic distance from the location (gps coordinates) of each household to the nearest national border. *Source: Calculated using ArcGis.*

Capital Indicator: Dummy variable that takes on the value one when the household is located in the ethnic homeland that hosts the capital city of the country and zero otherwise.

Coastal Indicator: Dummy variable that takes on the value one for country-ethnic homelands that are adjacent to the coast and zero otherwise (when the area is landlocked).

Table 1 - Border (Ethnic Partitioning) Artificiality
Panel A: Geography, Ecology, Natural Resources and Ethnic Partitioning

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|----------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Log Land Area | 0.0556*** | 0.0529*** | 0.0618*** | 0.0554*** | 0.0489*** | 0.0528*** | 0.0572*** |
| Double-clustered s.e. | (0.0153) | (0.0159) | (0.0175) | (0.0186) | (0.0150) | (0.0156) | (0.0169) |
| Lake Indicator | | 0.0961* | 0.0963 | 0.0965* | 0.0971* | 0.0933* | 0.0941 |
| Double-clustered s.e. | | (0.0565) | (0.0645) | (0.0567) | (0.0558) | (0.0557) | (0.0629) |
| River Indicator | | -0.0053 | -0.0164 | -0.0092 | -0.0065 | -0.0057 | -0.0149 |
| Double-clustered s.e. | | (0.0337) | (0.0324) | (0.0325) | (0.0338) | (0.0346) | (0.0338) |
| Elevation | | | -0.0411 | | | | -0.0475 |
| Double-clustered s.e. | | | (0.0709) | | | | (0.0727) |
| Suitability for Agriculture | | | 0.1239 | | | | 0.1324 |
| Double-clustered s.e. | | | (0.0974) | | | | (0.1103) |
| Malaria Stability Index | | | | 0.0195 | | | -0.0378 |
| Double-clustered s.e. | | | | (0.0982) | | | (0.1086) |
| Distance to the Coast | | | | 0.0000 | | | 0.0001 |
| Double-clustered s.e. | | | | (0.0001) | | | (0.0001) |
| Diamond Mine Indicator | | | | | 0.0359 | | 0.0293 |
| Double-clustered s.e. | | | | | (0.0651) | | (0.0655) |
| Oil Indicator | | | | | 0.0285 | | 0.017 |
| Double-clustered s.e. | | | | | (0.0565) | | (0.0577) |
| Nearby Groups in the Same Family | | | | | | -0.0727 | -0.0711 |
| Double-clustered s.e. | | | | | | (0.0579) | (0.0632) |
| Adjusted R-square | 0.050 | 0.055 | 0.059 | 0.056 | 0.056 | 0.057 | 0.062 |
| Region Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 825 | 825 | 825 | 825 | 825 | 825 | 825 |

Table 1 - Panel A reports linear probability model (LPM) estimates associating ethnic partitioning (SPLIT) with geographical, ecological and natural resource variables. 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 (using the 2000 Digital Chart of the World). All specifications include a set of (five) region fixed effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. The Supplementary Appendix reports summary statistics for all variables. Standard errors 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 1 - Border (Ethnic Partitioning) Artificiality**Panel B: Historical (Pre-colonial) Features and Ethnic Partitioning**

| | (1) | (3) | (3) | (4) | (5) | (6) | (7) |
|---|-----------|-----------|-----------|--------------------|-----------|-----------|-----------|
| Log Land Area | 0.0551*** | 0.0540*** | 0.0527*** | 0.0527*** | 0.0485*** | 0.0529*** | 0.0524*** |
| Double-clustered s.e. | (0.0158) | (0.0153) | (0.0155) | (0.0160) | (0.0170) | (0.0174) | (0.0158) |
| Lake Indicator | 0.0984* | 0.0915 | 0.0956* | 0.0942* | 0.0878 | 0.0962 | 0.0967* |
| Double-clustered s.e. | (0.0555) | (0.0589) | (0.0558) | (0.0559) | (0.0582) | (0.0590) | (0.0561) |
| River Indicator | -0.0049 | -0.0097 | -0.0058 | -0.0077 | -0.0067 | -0.0054 | -0.006 |
| Double-clustered s.e. | (0.0337) | (0.0351) | (0.0348) | (0.0347) | (0.0337) | (0.0337) | (0.0331) |
| Pre-colonial Conflict Indicator | -0.0663 | | | | | | |
| Double-clustered s.e. | (0.0733) | | | | | | |
| Distance to Pre-colonial Conflict | | -0.0444 | | | | | |
| Double-clustered s.e. | | (0.0839) | | | | | |
| Slave Trades Indicator | | | 0.0045 | | | | |
| Double-clustered s.e. | | | (0.0322) | | | | |
| Log Number of Slaves (normalized by land area) | | | | 0.0063 (0.0080) | | | |
| Pre-colonial Kingdom Indicator | | | | | 0.0466 | | |
| Double-clustered s.e. | | | | | (0.0469) | | |
| Distance to Pre-colonial Kingdom | | | | | | 0.0009 | |
| Double-clustered s.e. | | | | | | (0.1235) | |
| Major City in 1400AD | | | | | | | 0.0233 |
| Double-clustered s.e. | | | | | | | (0.0652) |
| Adjusted R-square | 0.056 | 0.056 | 0.055 | 0.056 | 0.057 | 0.055 | 0.055 |
| Region Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 825 | 825 | 825 | 825 | 825 | 825 | 825 |

Table 1 - Panel B reports linear probability model (LPM) estimates associating ethnic partitioning (SPLIT) with historical variables capturing pre-colonial conflict, kingdoms, and slavery (in Panel B). 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 (using the 2000 Digital Chart of the World). All specifications include a set of (five) region fixed effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. The Supplementary Appendix reports summary statistics for all variables. Standard errors 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: Ethnic Partitioning and Civil Conflict.
Baseline Country Fixed-Effects Estimates

| | All Ethnicity-Country Homelands | | | | | | Ethnicity-Country Homelands Close to the National Border | | | | | |
|--|---------------------------------|-----------|-----------|-----------|-------------------|-------------------|--|-----------|-----------|-------------------|-------------------|-----------|
| | All Observations | | | | Excl. Outliers | Excl. Capitals | All Observations | | | Excl. Outliers | Excl. Capitals | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Panel A. Negative Binomial ML Estimates | | | | | | | | | | | | |
| SPLIT (Partitioning) | 0.4513*** | 0.3329** | 0.4495*** | 0.4626*** | 0.4494*** | 0.4565*** | 0.9247*** | 0.8050*** | 0.6014*** | 0.5906*** | 0.5906*** | 0.5806*** |
| Double-clustered s.e. | (0.1611) | (0.1851) | (0.1254) | (0.1201) | (0.1172) | (0.1236) | (0.1704) | (0.2372) | (0.2226) | (0.2176) | (0.2176) | (0.2146) |
| SPIL (Adjacent Split) | 0.0481 | 0.391 | 0.4619* | 0.4920* | 0.4834* | 0.4256* | 0.0879 | 0.5679 | 0.4328 | 0.4514 | 0.4514 | 0.3928 |
| Double-clustered s.e. | (0.2789) | (0.3430) | (0.2626) | (0.2628) | (0.2686) | (0.2760) | (0.5748) | (0.4733) | (0.3818) | (0.3565) | (0.3565) | (0.3640) |
| Log Likelihood | -4506.794 | -4280.172 | -4119.95 | -4108.723 | -3993.148 | -3781.286 | -1697.469 | -1561.61 | -1517.999 | -1510.73 | -1510.73 | -1445.62 |
| Panel B. Linear Probability Model (LPM) Estimates | | | | | | | | | | | | |
| SPLIT (Partitioning) | 0.0562** | 0.0660*** | 0.0783*** | 0.0819*** | 0.0839*** | 0.0789*** | 0.0874** | 0.0835* | 0.0821 | 0.0903** | 0.0903** | 0.0893* |
| Double-clustered s.e. | (0.0241) | (0.0238) | (0.0258) | (0.0266) | (0.0266) | (0.0266) | (0.0399) | (0.0484) | (0.0523) | (0.0457) | (0.0457) | (0.0461) |
| SPIL (Adjacent Split) | 0.0571 | 0.1146*** | 0.1284*** | 0.1443*** | 0.1487*** | 0.1468*** | 0.1787*** | 0.2246*** | 0.2297*** | 0.2444*** | 0.2444*** | 0.2347*** |
| Double-clustered s.e. | (0.0486) | (0.0394) | (0.0397) | (0.0408) | (0.0402) | (0.0408) | (0.0594) | (0.0604) | (0.0631) | (0.0562) | (0.0562) | (0.0575) |
| adjusted R-square | 0.304 | 0.43 | 0.44 | 0.445 | 0.446 | 0.446 | 0.315 | 0.463 | 0.475 | 0.489 | 0.489 | 0.486 |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | No | No | Yes | Yes | Yes | Yes | No | No | Yes | Yes | Yes | Yes |
| Geographic Controls | No | No | No | Yes | Yes | Yes | No | No | No | Yes | Yes | Yes |
| Country Fixed Effects | No | Yes | Yes | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 1199 | 1165 | 579 | 579 | 579 | 579 | 579 | 568 |

The table reports Negative Binomial Maximum Likelihood (NB-ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in Panel A is the total number of civil conflict incidents at each country-ethnic homeland over the period 1997-2013. The dependent variable in Panel B is an dummy variable that takes on the value of one for country-ethnic homelands that have experienced conflict and zero otherwise over the period 1997-2013. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (2)-(6) and (8)-(12) include country fixed effects (constants not reported). The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). The specifications in columns (5) and (11) exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (6) and (12) exclude country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethnolinguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 3: Ethnic Partitioning and Civil Conflict Intensity. Baseline Country Fixed-Effects Estimates

| | All Ethnicity-Country Homelands | | | | | Ethnicity-Country Homelands Close to the National Border | | | | |
|-----------------------|---------------------------------|----------------------------|------------------|------------------------|---------------------------|--|----------------------------|------------------|------------------------|---------------------------|
| | Deadly Incidents | Deadly Incidents Indicator | Total Casualties | Duration All Incidents | Duration Deadly Incidents | Deadly Incidents | Deadly Incidents Indicator | Total Casualties | Duration All Incidents | Duration Deadly Incidents |
| | NB-ML | LPM | NB-ML | Poisson - ML | Poisson - ML | NB-ML | LPM | NB-ML | Poisson - ML | Poisson - ML |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| SPLIT (Partitioning) | 0.3356** | 0.0599** | 0.4843*** | 0.2015*** | 0.1658** | 0.4651** | 0.0820* | 0.8489*** | 0.2784** | 0.4350*** |
| Double-clustered s.e. | (0.1357) | (0.0287) | (0.1651) | (0.0622) | (0.0740) | (0.2037) | (0.0428) | (0.3787) | (0.1216) | (0.1521) |
| SPIL (Adjacent Split) | 0.3948 | 0.1461*** | 0.1161 | 0.2478** | 0.3648** | 0.2745 | 0.2378*** | 0.3573 | 0.3731** | 0.4674* |
| Double-clustered s.e. | (0.2465) | (0.0463) | (0.3121) | (0.1174) | (0.1629) | (0.3110) | (0.0611) | (0.5155) | (0.1804) | (0.2415) |
| Log Likelihood | -2910.906 | — | -4516.44 | -2759.21 | -2232.74 | -1028.82 | — | -1657.27 | -1057.16 | -805.15 |
| adjusted R-square | — | 0.411 | — | — | — | — | 0.449 | — | — | — |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1199 | 1212 | 1212 | 579 | 579 | 575 | 579 | 579 |

The table reports estimates associating civil conflict with ethnic partitioning at the country-ethnicity homeland level. In columns (1) and (6) the dependent variable is the total number of deadly civil conflict incidents at each country-ethnic homeland over the sample period (1997-2013). These models are estimated with the negative binomial ML model. In columns (2) and (7) the dependent variable is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced at least one deadly conflict incident over the period 1997-2013 and zero otherwise. These columns give linear probability model estimates. In columns (3) and (8) the dependent variable is the total number of fatalities at each country-ethnic homeland over 1997-2013. These models are estimated with the negative binomial ML model. For the estimation we exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. In columns (4) and (9) the dependent variable is the number of years that each country-ethnic homeland has experienced conflict over the period 1997-2013. These columns give Poisson ML estimates. In columns (5) and (10) the dependent variable is the number of years that each country-ethnic homeland has experienced deadly conflict (at least one casualty) over the period 1997-2013. These columns give Poisson ML estimates. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include country fixed effects (constants not reported) and a rich set of controls. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 4: Ethnic Partitioning and Civil Conflict Aspects

| | All Ethnicity-Country Homelands | | | Ethnicity-Country Homelands Close to the National Border | | |
|--|---------------------------------|--------------------------|-----------------------------|--|--------------------------|-----------------------------|
| | <u>Battles</u> | <u>Civilian Violence</u> | <u>Riots & Protests</u> | <u>Battles</u> | <u>Civilian Violence</u> | <u>Riots & Protests</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Negative Binomial ML Estimates | | | | | | |
| SPLIT (Partitioning) | 0.4428*** | 0.4328*** | 0.0747 | 0.5238** | 0.4980*** | 0.0453 |
| Double-clustered s.e. | (0.1489) | (0.1229) | (0.1526) | (0.2818) | (0.1949) | (0.2402) |
| SPIL (Adjacent Split) | 0.4846 | 0.3816 | 0.4119 | 0.4372 | -0.0188 | 0.9385* |
| Double-clustered s.e. | (0.3060) | (0.3523) | (0.2574) | (0.3765) | (0.3662) | (0.4926) |
| Log Likelihood | -2918.506 | -2876.564 | -2203.732 | -1068.327 | -1000.611 | -648.381 |
| Panel B: Linear Probability Estimates | | | | | | |
| SPLIT (Partitioning) | 0.0912** | 0.0517 | 0.0193 | 0.0902* | 0.0647 | 0.0066 |
| Double-clustered s.e. | (0.0375) | (0.0320) | (0.0305) | (0.0462) | (0.0447) | (0.0540) |
| SPIL (Adjacent Split) | 0.0631 | 0.1749*** | 0.0773 | 0.1724*** | 0.1839*** | 0.0745 |
| Double-clustered s.e. | (0.0442) | (0.0577) | (0.0533) | (0.0615) | (0.0705) | (0.0775) |
| Adjusted R-squared | 0.465 | 0.422 | 0.439 | 0.457 | 0.435 | 0.417 |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 579 | 579 | 579 |

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating the main categories of civil conflict with ethnic partitioning at the country-ethnicity homeland level. Columns (1) and (4) focus on battles. Columns (2) and (5) focus on violence against the civilian population. Columns (3) and (6) focus on riots and protests. In Panel A the dependent variable is the total number of battles (in columns (1) and (4)), violent events against the civilian population (in columns (2) and (5)) and riots and protests events (in columns (3) and (6)). In Panel B the dependent variable is an indicator (dummy) variable for country-ethnic homelands that have experienced at least one battle (in columns (1) and (4)), at least one violent event against the civilian population (in columns (2) and (5)) and at least one event of riots and protests (in columns (3) and (6)) over the period 1997-2013 (and zero otherwise). SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 5: Ethnic Partitioning and Conflict Actors

| | All Ethnicity-Country Homelands | | | | Ethnicity-Country Homelands Close to the National Border | | | |
|--|---------------------------------|-------------------|-----------------|----------------|--|-------------------|-----------------|----------------|
| | Government Forces | Rebels & Militias | Nearby External | Other External | Government Forces | Rebels & Militias | Nearby External | Other External |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Panel A. Negative Binomial ML Estimates | | | | | | | | |
| SPLIT (Partitioning) | 0.5247*** | 0.4908*** | 1.1280*** | 0.244 | 0.8198*** | 0.6083*** | 1.1310*** | 0.8889* |
| Double-clustered s.e. | (0.1394) | (0.1381) | (0.2577) | (0.2534) | (0.2212) | (0.2434) | (0.2242) | (0.5275) |
| SPIL (Adjacent Split) | 0.496 | 0.3258 | 0.1629 | -0.519 | 0.2893 | 0.0667 | -0.0037 | -1.1611 |
| Double-clustered s.e. | (0.3108) | (0.3089) | (0.4327) | (0.4765) | (0.3840) | (0.3620) | (0.3405) | (0.9901) |
| Log Likelihood | -3213.30 | -3538.28 | -1088.25 | -571.59 | -1127.39 | -1278.77 | -418.72 | -170.35 |
| Panel B. Linear Probability Model (LPM) Estimates | | | | | | | | |
| SPLIT (Partitioning) | 0.1089*** | 0.0663** | 0.0658** | 0.0065 | 0.1240*** | 0.0861* | 0.0693** | 0.0349 |
| Double-clustered s.e. | (0.0281) | (0.0327) | (0.0325) | (0.0228) | (0.0426) | (0.0497) | (0.0342) | (0.0298) |
| SPIL (Adjacent Split) | 0.1300** | 0.1059** | 0.0737 | -0.009 | 0.1905*** | 0.1671*** | 0.0074 | -0.0625 |
| Double-clustered s.e. | (0.0530) | (0.0482) | (0.0466) | (0.0292) | (0.0625) | (0.0619) | (0.0487) | (0.0415) |
| adjusted R-square | 0.453 | 0.472 | 0.345 | 0.378 | 0.467 | 0.485 | 0.384 | 0.425 |
| Observations | 1212 | 1212 | 1212 | 1212 | 579 | 579 | 579 | 579 |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating civil conflict by actor with ethnic partitioning at the country-ethnicity homeland level. Columns (1) and (5) focus on conflict where government forces participate. Columns (2) and (6) focus on conflict where rebels and militias participate. Columns (3) and (7) focus on military interventions of adjacent (nearby) African countries. Columns (4) and (8) focus on foreign interventions by peace-keeping forces (UN, African Union, etc.). In Panel A the dependent variable is the total number of events of each category across country-ethnic homelands over the period 1997-2013. In Panel B the dependent variable is an indicator (dummy) variable for country-ethnic homelands that have experienced at least one event from each type of civil conflict over the period 1997-2013 (and zero otherwise). SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (5)-(8) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 6: Ethnic Partitioning and Civil Conflict Types. UCDP GED

| | All Ethnic Homelands | | | Ethnic Homelands close to the Border | | |
|--|----------------------|------------------|-----------------|--------------------------------------|------------------|-----------------|
| | <u>All Events</u> | <u>Indicator</u> | <u>Duration</u> | <u>Events</u> | <u>Indicator</u> | <u>Duration</u> |
| | NB-ML (1) | LPM (2) | NB-ML (3) | NB-ML (4) | LPM (5) | NB-ML (6) |
| Panel A: State (Government Forces) Civil Conflict | | | | | | |
| SPLIT (Partitioning) | 0.4978** | 0.0487* | 0.3390** | 0.8053*** | 0.0799** | 0.5469** |
| Double-clustered s.e. | (0.2411) | (0.0294) | (0.1422) | (0.2335) | (0.0393) | (0.2389) |
| SPLIT (Adjacent Split) | 1.1577*** | 0.0902* | 0.6868** | 0.4340 | 0.0424 | 0.255 |
| Double-clustered s.e. | (0.4761) | (0.0518) | (0.2964) | (0.5468) | (0.0534) | (0.3744) |
| Log Likelihood | -1453.054 | — | -1046.922 | -528.002 | — | -383.392 |
| adjusted R-square | — | 0.471 | — | — | 0.441 | — |
| Panel B: One-Sided Violence against Civilian Population | | | | | | |
| SPLIT (Partitioning) | 0.3468 | 0.0269 | 0.2750* | 0.3288 | 0.0331 | 0.2925 |
| Double-clustered s.e. | (0.2416) | (0.0292) | (0.1474) | (0.2615) | (0.0404) | (0.2237) |
| SPLIT (Adjacent Split) | 0.4708 | 0.0829* | 0.4935** | 0.0901 | 0.0161 | 0.1659 |
| Double-clustered s.e. | (0.4549) | (0.0481) | (0.2277) | (0.6886) | (0.0626) | (0.3416) |
| Log Likelihood | -1499.837 | — | -1099.667 | -556.790 | — | -396.804 |
| adjusted R-square | — | 0.404 | — | — | 0.434 | — |
| Panel C: Non-State Civil Conflict | | | | | | |
| SPLIT (Partitioning) | -0.2087 | -0.0459 | 0.026 | -0.4122 | -0.0351 | -0.1797 |
| Double-clustered s.e. | (0.4062) | (0.0317) | (0.3374) | (0.5178) | (0.0283) | (0.4917) |
| SPLIT (Adjacent Split) | -0.8703 | -0.0344 | -0.5089 | -0.6593 | -0.0268 | -0.6964 |
| Double-clustered s.e. | (0.7193) | (0.0302) | (0.4423) | (0.8728) | (0.0415) | (0.7086) |
| Log Likelihood | -841.675 | — | -644.791 | -243.970 | — | -199.677 |
| adjusted R-square | — | 0.320 | — | — | — | — |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 579 | 579 | 579 |

The table reports estimates associating three types of civil conflict with ethnic partitioning at the country-ethnicity homeland level using data from the UCDP GED 1.5 project. Panel A gives estimates focusing on state conflict (where government forces, troops, and militias participate). Panel B gives estimates focusing on one-sided violence, mostly against the civilian population. Panel C gives estimates focusing on conflict between (at least) two non-state actors (where the government is not involved). The dependent variable in columns (1) and (4) is the total number of civil conflict incidents at each country-ethnic homeland over the period 1989-2010. These models are estimated with the negative binomial maximum likelihood model. The dependent variable in columns (2) and (5) is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced each conflict type over the period 1989-2010 (and zero otherwise). The dependent variable in columns (3) and (6) is the number of years that each country-ethnic homeland has experienced each type of conflict over the period 1989-2010. These models are estimated with the negative binomial maximum likelihood model. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands.

The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include a country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7: Ethnic Partitioning, Ethnic-based Political Discrimination, and Major Ethnic Civil Wars.

| | Ethnic Discrimination | | | | Ethnic War | | | | |
|--------------------------|-----------------------|-----------|-----------|----------|------------|-----------|-----------|-----------|----------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| SPLIT (Partitioning) | 0.1030*** | 0.0778*** | 0.0766*** | 0.0718** | 0.1155** | 0.1402*** | 0.1354*** | 0.1101*** | 0.0663* |
| Double-clustered s.e. | (0.0342) | (0.0276) | (0.0289) | (0.0282) | (0.0509) | (0.0381) | (0.0415) | (0.0390) | (0.0351) |
| SPIL (Adjacent Split) | | | | 0.0169 | | | | 0.2570** | 0.2446** |
| Double-clustered s.e. | | | | (0.0764) | | | | (0.1155) | (0.1091) |
| Political Discrimination | | | | | | | | | 0.6099** |
| Double-clustered s.e. | | | | | | | | | (0.1411) |
| Adjusted R-square | 0.018 | 0.475 | 0.495 | 0.521 | 0.019 | 0.423 | 0.426 | 0.487 | 0.503 |
| Observations | 593 | 593 | 593 | 593 | 593 | 593 | 593 | 593 | 593 |
| Countries | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 |
| Country Fixed Effects | No | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes |
| Simple Controls | No | No | Yes | Yes | No | No | Yes | Yes | Yes |
| Location Controls | No | No | Yes | Yes | No | No | Yes | Yes | Yes |
| Geographic Controls | No | No | No | Yes | No | No | No | Yes | Yes |

The table reports linear probability model estimates, associating ethnic-based political discrimination and major ethnic wars with ethnic partitioning. The dependent variable in columns (1)-(4) is a dummy variable that takes on the value of one if an ethnicity has experienced discrimination from the central government for at least one year over the period 1960-2010. The dependent variable in columns (5)-(9) is a dummy variable that takes on the value of one if an ethnicity has experienced a major or minor civil war with an explicit ethnic dimension over the period 1960-2010. Data on ethnic wars and ethnic-based political discrimination from the national government come from the Ethnic Power Relations (EPR) database (Wimmer, Cederman and Min (2009)).

SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (2)-(5) and (6)-(9) include a vector of country fixed effects (constants not reported). The specifications in columns (3)-(4) and (7)-(9) include log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers (simple controls). The specifications in columns (4), (8) and (9) include a set of location and geographic controls. The specification in column (9) conditions on the political discrimination dummy that takes on the value of one if an ethnicity has experienced discrimination from the central government for at least one year over the period 1960-2010. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 8: The Long-Run Effects of Ethnic Partitioning on Individual Well-Being and Education. DHS Data
Baseline Estimates**

| | DHS Composite Wealth Index | | | | | | Education | | | | | |
|-----------------------|--------------------------------|------------|-----------|----------------------------------|------------|------------|--------------------------------|-----------|----------|----------------------------------|----------|----------|
| | All Observations (Individuals) | | | Observations close to the Border | | | All Observations (Individuals) | | | Observations close to the Border | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Ethnic Partitioning | -0.3853*** | -0.3227*** | -0.1978** | -0.2818*** | -0.2679*** | -0.2248*** | -0.3166** | -0.2238** | -0.1482* | -0.1726* | -0.1493* | -0.1242 |
| Double-clustered s.e. | (0.1129) | (0.0987) | (0.0919) | (0.1018) | (0.0961) | (0.0791) | (0.1297) | (0.0938) | (0.0811) | (0.0976) | (0.0869) | (0.0758) |
| Non-Indigenous | 0.1936** | 0.1942** | 0.1825*** | 0.1359* | 0.1397* | 0.1942*** | 0.1082 | 0.1163 | 0.1024 | 0.0568 | 0.0587 | 0.0813 |
| Double-clustered s.e. | (0.0887) | (0.0844) | (0.0690) | (0.0783) | (0.0741) | (0.0688) | (0.0838) | (0.0766) | (0.0704) | (0.0691) | (0.0632) | (0.0549) |
| Adjusted R-square | 0.028 | 0.072 | 0.167 | 0.053 | 0.090 | 0.162 | 0.151 | 0.222 | 0.255 | 0.132 | 0.188 | 0.210 |
| Observations | 88171 | 88171 | 88171 | 44090 | 44090 | 44090 | 88171 | 88171 | 88171 | 44090 | 44090 | 44090 |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Individual Controls | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| Location Controls | No | No | Yes | No | No | Yes | No | No | Yes | No | No | Yes |

The table reports OLS estimates, associating the DHS composite wealth index (in columns (1)-(6)) and an education index (in columns (7)-(12)) with ethnic partitioning at the individual level. The composite wealth index is calculated by the DHS team in each country via a principal component method using easy-to-collect data on a household's ownership of selected assets (e.g., televisions and bicycles), materials used for housing construction and public good access (e.g., type of water access, electrification, and sanitation). The ordered education index (range from 0 to 5) assigns the respondent's level of formal schooling into six categories. A score of 0 indicates "no education"; a score of 1 indicates "incomplete primary"; a score of 2 indicates "complete primary"; a score of 3 indicates "incomplete secondary"; a score of 4 indicates "complete secondary"; and a score of 5 indicates "higher education". The ethnic partitioning index takes on the value of one for individuals that identify with a partitioned ethnicity. The non-indigenous indicator takes on the value of one for individuals residing outside their ethnicity's ancestral homeland and takes on the value of zero for individuals residing in their ethnicity's ancestral homeland ("movers"). All specifications include a vector of country fixed effects (constants not reported). The set of individual controls in columns (2), (3), (5), (6), (8), (9), (11), and (12) includes a vector of 10 age-bracket fixed effects, a vector of 6 marital-status fixed effects, and a vector of 7 religion fixed effects. The set of location controls in columns (3), (6), (9), and (12) includes the distance of each individual to the capital city, the distance to the sea, the distance to the national border and an indicator that takes on the value of one if the individual resides in the capital city. The specifications in columns (4)-(6) and (10)-(12) focus on individuals residing close to the national border (using as a cut-off the median distance; 80 kilometers). The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the ethnicity and the ethnic homeland dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

The table reports OLS estimates, associating DHS composite wealth index (in columns (1)-(6)) and an education index (in columns (7)-(12)) with ethnic partitioning at the individual level. The composite wealth index is calculated by the DHS team in each country via a principal component method using easy-to-collect data on a household's ownership of selected assets (e.g., televisions and bicycles), materials used for housing construction and public good access (e.g., type of water access, electrification, and sanitation). The ordered education index (range from 0 to 5) assigns the respondent's level of formal schooling into six categories. A score of 0 indicates "no education"; a score of 1 indicates "incomplete primary"; a score of 2 indicates "complete primary"; a score of 3 indicates "incomplete secondary"; a score of 4 indicates "complete secondary"; and a score of 5 indicates "higher education". The ethnic partitioning identity index (SPLIT-ID) takes on the value of one for individuals that identify with a partitioned ethnicity. The location based ethnic partitioning index (SPLIT-LOC) takes on the value of one for individuals that reside in ethnic homelands that have been partitioned by the national border and zero otherwise. The partitioned location and interaction takes on the value of one for individuals that both identify with a partitioned ethnicity and reside in partitioned ethnic homelands and zero otherwise. The non-indigenous indicator takes on the value of one for individuals residing outside their ethnicity's ancestral homeland and takes on the value of zero for individuals residing in their ethnicity's ancestral homeland ("movers"). All specifications include a vector of country fixed effects (constants not reported). The set of individual controls includes a vector of 10 age-bracket fixed effects, a vector of 6 marital-status fixed effects, and a vector of 7 religion fixed effects. The set of location controls includes the distance of each individual to the capital city, the distance to the sea, the distance to the national border and an indicator that takes on the value of one if the individual resides in the capital city. The specifications in columns (3)-(4) and (9)-(10) focus on individuals residing close to the national border (using as a cut-off the median distance; 80 kilometers). The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the ethnicity and the ethnic homeland dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

On-line Supplementary Appendix

(not for publication)

The Long-Run Effects of the Scramble for Africa*

Stelios Michalopoulos
Brown University, Harvard University, NBER and CEPR

Elias Papaioannou
London Business School, NBER and CEPR

March 23, 2015

Abstract

The Supplementary Appendix is structured into seven sections. Section 1 gives descriptive evidence, summary statistics and presents the key patterns on the spatial distribution of all types of conflict across Africa. Section 2 reports additional results showing that ethnic partitioning is not systematically related to various ethnic-specific pre-colonial economic, social, and political traits using data from Murdock's Ethnographic Atlas (1967). This section also reports "balancedness tests" showing that within countries there are no systematic differences between split and non-split ethnic homelands across numerous geographical, ecological, natural resources and location characteristics. Section 3 gives further evidence and numerous sensitivity checks on the impact of ethnic partitioning on various aspects of civil conflict (intensity, likelihood, duration, fatalities). Section 4 presents a case-study illustrating the effect of ethnic partitioning on conflict in Central-Eastern Africa. Section 5 explores potential heterogeneous effects of ethnic partitioning on conflict. Section 6 reports robustness checks on the link between partitioning, ethnic-based discrimination from the government and participation on major ethnic wars using data from the Ethnic Power Database. Section 7 reports further evidence on the negative impact of identifying with a partitioned ethnicity using individual-level data from the Demographic and Health Surveys.

*We thank 4 anonymous referees for proposing many of the useful sensitivity checks and additional results. We also thank Sebastian Hohmann and Fabrizio Dell' Acqua for excellent research assistance. All errors are our sole responsibility.

1 Descriptives

This Section reports summary statistics of the main variables employed in the empirical analysis and discusses the main data patterns.

1.1 Summary Statistics

Appendix Table *A* reports the name of partitioned ethnicities (as coded by Murdock (1959)) and the percentage of the homeland of each group falling into more than one country.

Appendix Table 1 gives summary statistics for all variables at the ethnic homeland level, which is the unit of analysis in the section examining the correlates of ethnic partitioning (Section 3). Panel *A* looks at all homelands (825 observations) whereas in Panel *B* we focus on those homelands whose centroid is close (border distance below or equal the median) to the national boundaries (413 observations).

Appendix Table 2 reports summary statistics for all variables at the country-ethnic homeland level, which is the unit of analysis in Sections 4 – 5; these sections examine the impact of ethnic partitioning on civil conflict. Panel *A* gives summary statistics for the full sample (1, 212 observations) and Panel *B* reports statistics in the border sample, using again the median value of distance to the national border (606 observations).

1.2 Data Patterns

1.2.1 ACLED

Appendix Tables 3 – 6 give details on the main conflict data (ACLED). Appendix Table 3 gives the number and share of all and deadly conflict events by year. On average there are around 2, 500 – 3, 000 events per year over the period 1997 – 2010. Approximately 900 of these events have resulted in at least one fatality. Conflict events increase considerably in 2011 (5, 261) and especially in 2012 (8, 753) and in 2013 (12, 565).¹

Appendix Table 4 reports the number (and share) of all and deadly events for each ACLED category. The sample includes 23, 381 battles, mostly involving government forces, militias and rebel groups and 20, 409 events of violence against the civilian population. These two categories are by far the most deadly (on average 43% of these events have resulted in at least one fatality). ACLED also includes 16, 147 events associated with riots and protests which in the overwhelming majority (93.5%) are not deadly. ACLED also reports 4, 727 non-violent events (such as base and headquarter establishments) conducted by a conflict actor. Less than 1% of these events result in fatalities.

¹Since this increase is most likely driven by improved reporting, we report below estimates dropping 2011, 2012, and 2013. If anything the impact of ethnic partitioning strengthens.

Appendix Table 5 reports the distribution of conflict events by the interaction of conflict actors. Panel *A* tabulates the data for all events, while Panel *B* tabulates the data for deadly events. Close to 10,000 events involve fighting between government troops and rebels that have an explicit agenda to counter state violence. If one adds conflict between government troops and (ethnic and political) militias, we have 16,138 events. Approximately 40% of these events result in fatalities. Attacks against civilians by government forces and rebels are also numerous (and quite deadly). There are 4,312 and 4,222 civilian attacks by government forces and rebels, respectively. The data further show that militias mostly target civilians, as ACLED reports 11,415 such events (40% of those result in at least one death).²

Appendix Table 6 reports conflict by country with information on the type and the actors involved. Somalia and the Democratic Republic of Congo are by far the most conflict-prone countries in Africa; in both countries we observe pervasive violence against civilians and many battles between government forces and rebel groups. Nigeria, Uganda, Burundi, Sudan, Angola, and Ethiopia are also conflict prone. Zimbabwe, South Africa, Egypt and Algeria look quite violent when examining total ACLED events; yet this is mainly driven by riots and protests. Burkina Faso, Gambia, Benin, Djibouti, Lesotho, Western Sahara, and Equatorial Guinea are the least conflict prone countries (less than 100 events of any type).

Figure 1 reports the spatial distribution of conflict events by main actor, while Figures 2*a – f* map events associated for each conflict actor.

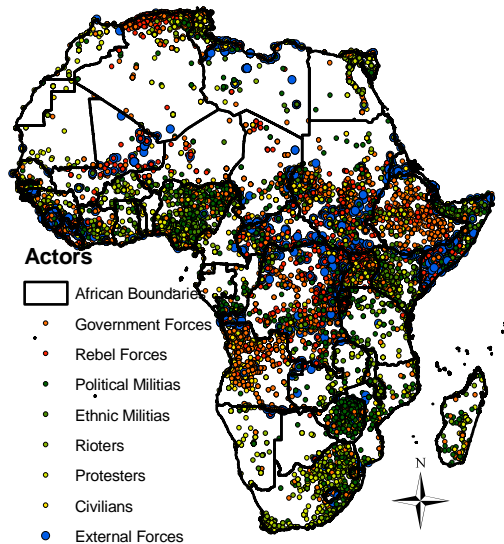


Figure 1

²So violence against civilians and riots-protests are reported twice, as part of main categories (that also includes battles and non-violent events associated with a conflict actor) and as part of actors (that also includes rebels, government troops, etc.). While there are some minor discrepancies across the two classifications, this does not affect the results (correlation around 0.99). Below we report estimates with using the category classification (e.g., Appendix Tables 13 and 16) and the "actor-based" classification (e.g., Appendix Tables 14 and 17).

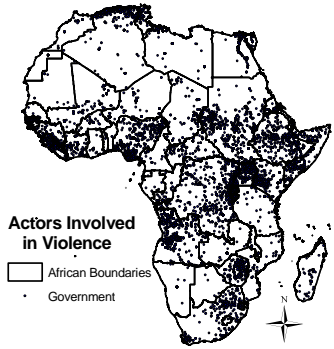


Figure 2a

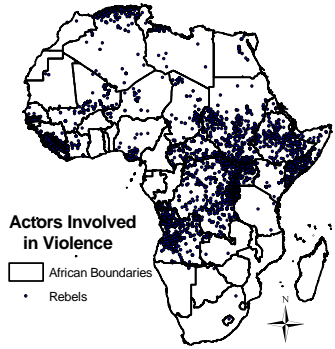


Figure 2b

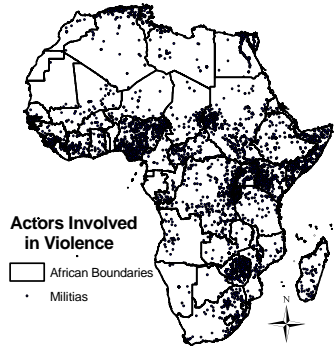


Figure 2c

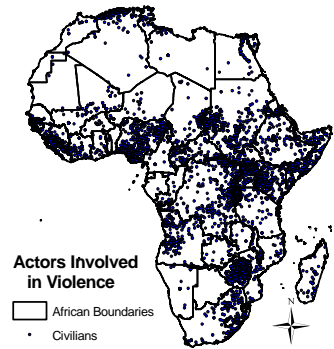


Figure 2d

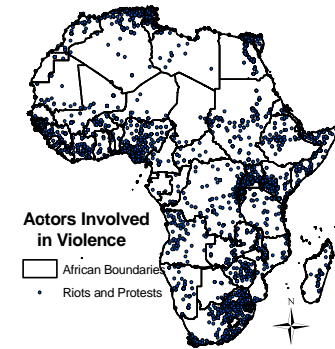


Figure 2e

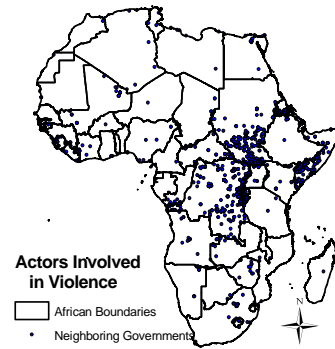


Figure 2f

Figures 3a and 3b plot the number of fatal events and fatal battles across the 1,212 country-ethnic homelands. Conflict intensity, as reflected in the number of deaths, is high in Eastern Congo at the border with Rwanda, Burundi and Uganda, in Somalia, Sudan and Angola.

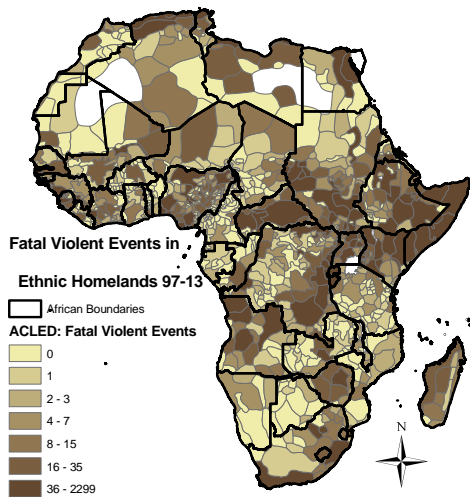


Figure 3a

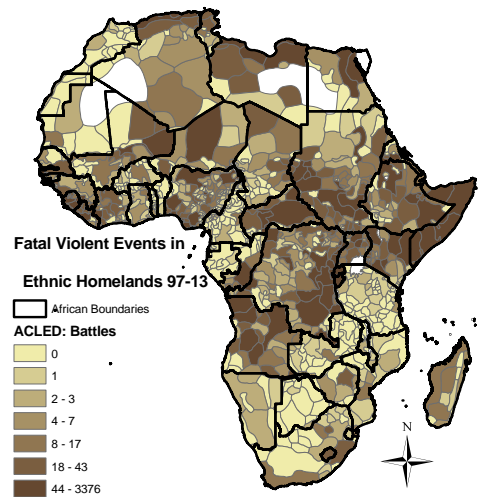


Figure 3b

1.2.2 UCDP

Figure 4 plots the spatial distribution of all conflict events in the UCDP GED 1.5 database (Sundberg, Lindgren, and Padsokocimaite (2010), Sundberg and Melander (2013)). This database focuses on deadly events (at least one fatality) that are associated with a major or minor civil war, as classified by the widely-used PRIO Armed Conflict Database; this database identifies minor and major civil war when conflict results into more than 25 fatalities in a given year. An event is defined as “*the incidence of the use of armed force by an organized actor against another organized actor, or against civilians, resulting in at least 1 direct death in either the best, low or high estimate categories at a specific location and for a specific temporal duration.*”

The database covers the period 1989 – 2010. UCDP consists of three databases, each focusing on different aspects of warfare:

(1) state conflict where government troops and state-backed militias fight either against rebel group and other anti-government forces (7,512 events). UCDP defines these incidents as “*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 in one calendar year*”.

(2) one-sided violence against civilians perpetrated either by state forces, rebels or militias (5,219 events). UCDP defines one-sided violence as “*use of armed force by the government of a state or by a formally organized group against civilians which results in at least 25 deaths in a year*”.

(3) non-state conflict where two (or more) rebel groups or militias fight with each other, without government’s participation (3,645 events). UCDP defines this type of conflict as “*conflict between two organized armed groups, neither of which is the government of a state, which results in at least 25 battle-related deaths in a year.*”

Figures 5a – 5c plot the events associated with each (mutually exclusive) type of conflict. Algeria and Somalia have experienced most state-related conflict events, followed by Angola, Ethiopia, Sierra Leone, and Burundi. The Democratic Republic of Congo, Sierra Leone, South Africa and Sudan are the countries where one-sided civilian violence takes the maximum value. South Africa is by far the country with most non-state conflicts (2,125 events), reflecting the intense fighting of ANC with other antiapartheid movements before and during the democratic transition in the early/mid 1990s. Somalia is the second-runner with 505 non-state conflict events, while in the other countries UCDP records less than 200 events. Appendix Table 6 gives the number of state conflict, one-sided violence and non-state conflict by country.

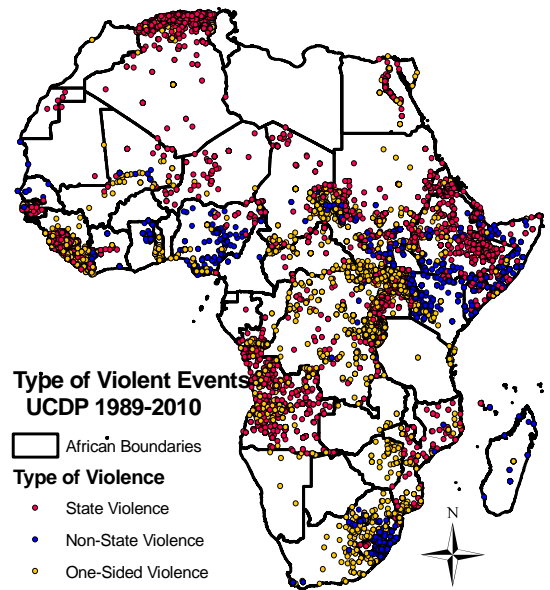


Figure 4

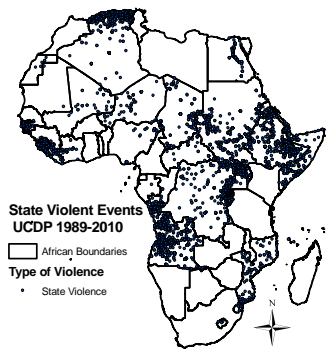


Figure 5a

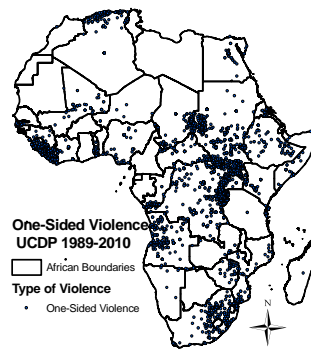


Figure 5b

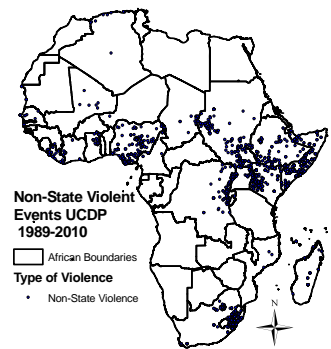


Figure 5c

1.2.3 ACLED and UCDP

Appendix Table 7 gives the correlation matrix of the various conflict measures. The correlation between battles and civilian violence (in ACLED) is high (0.64), but far from perfect. The correlogram further shows that riots and protests are distinct from battles and violence against civilians, as the correlation is low (0.13 and 0.30, respectively). Comparing ACLED to UCDP GED reveals a strong correlation between UCDP's state-based conflict and ACLED's battles (0.72); this is illustrated in Figures 6a – 6b below. There is a much weaker association between state-based conflict and civilian violence (0.47) and riots-protests (0.19). Non-state-based conflict is uncorrelated with UCDP state conflict (0.09) as well as ACLED-based battles (0.15) and civilian violence (0.13).

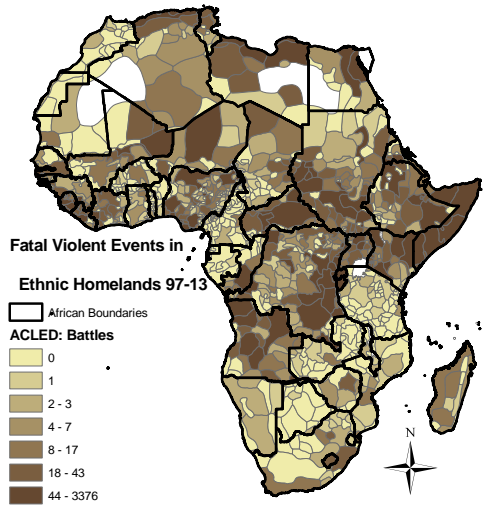


Figure 6a

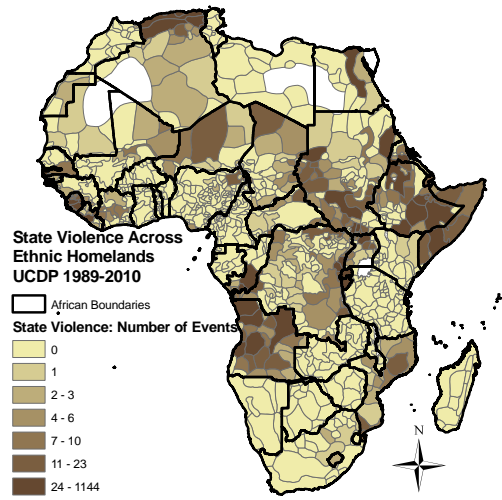


Figure 6b

2 Border Artificiality (w.r.t Ethnic Partitioning)

In this section we provide further evidence on border artificiality with respect to ethnic partitioning. First, we report cross-sectional specifications linking ethnic partitioning to pre-colonial features. Second, we report "balancedness tests" that examine whether within countries split and non-split homelands differ systematically across various geographic, ecological, location, and natural resource characteristics.

2.1 Pre-colonial Ethnic Traits

As in Section 3, we estimate simple cross-sectional specifications associating the binary ethnic partitioning index to numerous ethnic-specific variables from Murdock (1967) *Ethnographic Atlas* that provides information on pre-colonial political, economic, and societal features. While Murdock's (1967) coding does not cover all groups, examining the role of various ethnic-specific pre-colonial features on partitioning sheds light on the (absence of) considerations among Europeans while drawing the colonial borders in late 19th century.

Appendix Tables 8 reports linear probability model (LPM) estimates across all ethnic homelands with information on pre-colonial traits. As in the main text, below the coefficient estimates we report double-clustered standard errors at the ethnic family and the country level using the method of Cameron, Gelbach, and Miller (2011). We always include the log of ethnic homeland's surface area and indicators for the presence of lakes and rivers, as Table 1 shows that larger homelands were more likely to be split and that there is some (weak) association between partitioning and the presence of water bodies. We also include five regional constants

to account for the different timing of colonization across Africa.³

In columns (1), (2), and (3) we examine the role of pre-colonial economic organization using a measure capturing the presence of complex and densely populated local communities, an index quantifying the importance of agriculture for subsistence, and a measure of pastoralism (based on the share of animal husbandry for subsistence), respectively.⁴ All three proxies of pre-colonial economic structure enter with small and statistically insignificant estimates.

In columns (4) and (5) we examine whether there is a link between ethnic partitioning and pre-colonial institutions using a dummy variable that identifies societies where local leaders were elected and a dummy variable that equals one for societies with inheritance rules for property, respectively.⁵ Again there is no systematic link between ethnic partitioning and these ethnic-specific institutional features.

In column (6) we link ethnic partitioning to the binary index of political centralization of Gennaioli and Rainer (2007) that equals zero when Murdock (1967) assigns an ethnicity as "*stateless*" or "*a petty chiefdom*" (e.g., Xam or the Tiv) and 1 when the ethnicity is part of a "*large paramount chiefdom*" or a "*state*" (e.g., Ganda and Zulu).⁶ The coefficient on the political centralization is statistically and economically insignificant.

In column (7) we examine the association between ethnic partitioning and a class stratification index that ranges from zero, for societies without any class distinctions, to four, for groups with significant class and wealth distinctions. There is no systematic association between partitioning and the proxy of within group inequality.

In column (8) we examine the role of family institutions, as captured by a binary polygyny variable.⁷ The association is again weak and statistically insignificant.

Overall the results in Appendix Table 8 reveal that ethnic partitioning does not correlate systematically with various proxies of pre-colonial ethnic-specific features. Across all specifications, the implied economic effect (as reflected in the increase in R^2 vis-a-vis the model with region constants and log land area and the water indicators) is negligible.⁸ And while these variables do contain noise (and thus attenuation is a concern), measurement error cannot fully explain the lack of significance, as several recent works document robust associations between societal traits in Murdock's Ethnographic Atlas and contemporary outcomes, such as regional

³The results are similar if we omit these variables.

⁴Michalopoulos, Putterman, and Weil (2015) show that these measures correlate with contemporary proxies of human capital and well-being.

⁵Giuliano and Nunn (2013) show that societies that used to elect local chiefs via elections have more favorable to democracy attitudes today.

⁶Gennaioli and Rainer (2007) and Michalopoulos and Papaioannou (2013) show that pre-colonial political centralization correlates with contemporary development both across and within-countries.

⁷Fenske (2014a) examines the origins and correlates of polygamy in Sub-Saharan Africa.

⁸The same applies with other goodness-of-fit measures.

development, public goods provision, fertility, etc. (e.g., Fenske (2014b), Alesina, Giuliano, and Nunn (2013), Michalopoulos and Papaioannou (2013), among many others).

2.2 Balancedness Tests

The crucial issue (identifying assumption) for the causal interpretation of our estimates linking ethnic partitioning to conflict and discrimination is not that *all* African borders were drawn in an arbitrary manner (something that is definitely not the case). Inference requires that partitioned and non-split homelands do not differ with respect to factors that may independently affect contemporary conflict *within* countries. While this assumption cannot be directly tested, we can examine whether split and non-split homelands differ across various observable characteristics, as this (indirectly) sheds light on whether there may be differences in unobservable features (see for example, Wantchekon, Klasnja, and Novta (2015) and Dell (2012), among many others).

In Appendix Table 9 we report country fixed-effects (LS) specifications associating geographical, ecological, natural resource, and location features to ethnic partitioning. As in our empirical analysis linking partitioning to conflict (Sections 4 – 5), the unit of analysis is an ethnicity-country observation. So, these specifications are test of means once we net out the impact of country-specific factors. Following the structure of our empirical framework, Panel *A* reports estimates in the full sample (1,212 observations), while Panel *B* focuses on the border sample using as a cutoff the median distance to the national border from the centroid of each region (606 observations). The evidence in Panel *A* shows that the two sets of country-ethnic homelands are comparable with respect to land area, water bodies, elevation, soil quality, malaria, and natural resources. In columns (1)-(12) *SPLIT* enters with a small and statistically indistinguishable from zero coefficient. At the same time, the marginal (within) R^2 that captures the explanatory power of ethnic partitioning once we net the impact of the country fixed effects, is zero. Unsurprisingly, the only differences between split and non-split groups are with regards to distance to the border (column (13)) and distance to the capital (column (14)).

Yet, as the results in Panel *B* show, by focusing at the border sample we neutralize these differences; in the border sample there are no differences between partitioned and non-split groups on distance from the centroid of each country-ethnic homeland to the border and the capital. The coefficient on *SPLIT* in (13) and (14) is small and statistically indistinguishable from zero. Column (1) shows that across border groups partitioned ones are larger in size compared to non-split groups (echoing the finding in Table 1 and Appendix Table 8). Moreover, although there seem to be some marginal statistical difference with respect to the presence of

rivers in this sample the implied economic effect is tiny (the coefficient is 0.0864 while the mean value of the dependent variable is 0.46) and the marginal (within) R^2 is 0.5%. Across all other specifications, the economic impact of ethnic partitioning -as implied by the marginal-within R^2 and the comparison of the coefficient estimate with the mean value of the dependent variable- on these observable characteristics is tiny; the marginal R^2 is less than 0.005 in all specifications in columns (2)-(14).

3 Ethnic Partitioning and Civil Conflict: Further Evidence and Sensitivity Analysis

3.1 Preliminary Evidence across Ethnic Homelands

Appendix Table 10 reports test of means (and medians) for the main civil conflict variables across the 825 ethnic homelands. For each variable we report summary statistics for partitioned and non-split ethnic homelands and a corresponding test of mean and median equality. Panel *A* focuses on all homelands, while Panel *B* reports statistics for ethnic regions close to the national border (using the median value of distance from the centroid of a group; 102 kilometers).⁹

All Civil Conflict Incidents (ACLED) 83% of all ethnic areas have experienced at least one conflict event (including riots and protests) over the period 1997 – 2013. 88% of partitioned ethnicities experienced conflict, while the likelihood of a civil conflict incidence for non-partitioned ethnicities is 6 percentage points lower (81.5%). When we focus on groups close to the border, the difference is starker; 72.5% of non-partitioned ethnic homelands experienced conflict, while 87.3% of split groups suffered from some type of conflict. This 15 percentage point difference is statistically significant at standard confidence levels. Partitioned groups have also experienced more violent events with an average of 65 incidents, while the average for non-split groups is 47.6. This difference is not statistically significant, because there are some extreme observations both in the sample of partitioned and non-split ethnicities (see Appendix Table 2). Yet the median difference (in (3)) is significant, as the median number of events for partitioned groups is 13 and for non-split groups 3. To account for outliers in (4) and (5) we exclude ethnic homelands where capitals fall and in (6) and (7) homelands where the number of conflict events exceeds the 99th percentile. The mean differences for these two samples now become starker (17 and 22, respectively) and statistically significant. There are also large differences between partitioned and non-split ethnicities across the median value of conflict events (13 compared to 3). The differences in conflict intensity between partitioned

⁹Standard errors in the test of means are adjusted for correlation across countries (using a homeland’s centroid) and across ethnic families using the multi-way clustering method of Cameron, Gelbach, and Miller (2011).

and non-split ethnic groups are also sizable when we focus on areas close to the border (Panel *B*). While the average (median) number of all civil conflict incidents for partitioned ethnicities is 66 (12), for non-split ones the average (median) is 32.3 (1).

Battles between government troops, rebels, and militias (ACLED) On average 72.5% of partitioned ethnic homelands have experienced at least one battle between government forces, rebels, or militias. When we focus on border groups the likelihood is similar (72%). The corresponding likelihood for non-split groups, however, in the full sample is 65% and in the border sample 55%. These differences are statistically significant at standard confidence levels. On average partitioned ethnic homelands have experienced ten more battles as compared to non-split groups (30 versus 20); and while due to outliers this difference is not statistically significant, once we exclude the top 1% of the conflict distribution or regions where capitals fall, the difference becomes significant. Likewise, the median value of battles for non-split ethnic groups close to the national border is zero, while the corresponding median value for partitioned ethnic homelands is 3.

Violence against Civilians (ACLED) The summary statistics of violence against civilians reveal considerable and statistically significant differences between partitioned and non-split groups. The likelihood that a partitioned homeland has experienced at least one violent event against its civilian population is 0.78, while the corresponding likelihood for non-split ethnicities is 0.68. The difference is larger when we focus on ethnic homelands close to the national border (0.15). On average partitioned ethnic homelands close to the border experience 24 violent incidents against the civilian population, while the average for non-split ethnicities is 11. The median value of violence against civilians across partitioned ethnic homelands is three times the median value across non-split groups (3 versus 1 or 0).

Riots and Protests (ACLED) In contrast to battles and violence against the civilian population, the test of means and medians with riots and protests do not reveal significant differences between partitioned and non-split homelands. While the likelihood of riots and protests is higher for partitioned groups, the number of riots and protests is higher for non-split groups. And when we exclude homelands where capitals fall or outliers, the mean and median differences are small and statistically indistinguishable from zero.

State Conflict (UCDP) The tests of means reveal sizable differences on both the likelihood and intensity of state-based conflict (as recorded in the UCDP GED database). The likelihood of state conflict (over 1989 – 2010) in split homelands is 44.6%, while in non-split

homelands is 15 percentage points lower, 29.9%. Again the difference is larger (close to 24 percentage points). when we focus on border areas. Mean differences in the intensity of state conflict are also large, especially when we drop outliers and observations in the capital and even more so when we focus on border areas. In the border sample and when we exclude capitals or outliers we observe on average 2.4 events in non-split homelands and 6.5 events in partitioned homelands.

One-Sided Violence (UCDP) The likelihood of one-sided violence according to the UCDP that focuses on deadly incidents associated with major and minor civil wars in the homelands of partitioned ethnic homelands is 0.48 (in both the full and the border sample). The corresponding values for non-split homelands are 0.33 and 0.245 for the full and the border sample, respectively. The intensity of violence against the civilian population is also significantly higher for split homelands. On average we observe 10.5 and 11 events in the homelands of partitioned groups in the full and border sample. In contrast there are on average 4.7 events in the sample of non-split groups (in both samples).

Non-State Conflict (UCDP) The tests of means with non-state conflict indicate that there are no major differences between partitioned and non-split ethnic homelands in this type of conflict. The differences in the likelihood of non-state conflict are very small. And conflict intensity as captured in the number of events is somewhat higher in non-split ethnic areas; yet even this pattern is driven by outliers. When we drop outliers (top 1%) the average number of non-state civil conflict events is around 1.1 – 1.3 for both split and non-split ethnicities.

3.2 Sensitivity Analysis in the Ethnicity/Country Sample

We have performed a comprehensive set of sensitivity checks to assess the robustness of our results linking ethnic partitioning to conflict. In particular:

3.2.1 Excluding 2011 – 2013

First, we examined the stability of the estimates excluding the period 2011 – 2013, when ACLED’s coverage increased considerably (see Appendix Table 3). During the period 1997 – 2010, there are on average 2,720 events per year. The median value is 2,756 events and the range is 1,952 (in 2006) - 3,509 (in 2010). The total number of events reported increased considerably in 2011 (5,261 events) and especially in 2012 (8,753 events) and 2013 (12,565 events). Since this increase may reflect both a change in reporting over the past years and/or

an increase in the true incidence of conflict we repeated estimation focusing only on the period 1997 – 2010.

Appendix Table 11 reports unconditional negative binomial (NB) maximum likelihood estimates with country constants (in Panel *A*) and country fixed-effects linear probability model (LPM) estimates (in Panel *B*). [The table "mirrors" Table 2 in the main part of the paper.] The ethnic partitioning index is always statistically significant (usually at the 99% confidence level). The coefficient in the specifications with a rich set of controls in the full sample is 0.51 implying that conflict is approximately 65% higher in the homelands of partitioned ethnicities, as compared to non-split homelands. The estimate in the border sample is comparable - though somewhat larger (0.60).¹⁰ The LPM specifications suggest that there is an 8% - 10% increased likelihood of conflict in the homelands of split (as compared to non-split) groups. These effects are similar -and if anything somewhat larger- to the baseline estimates over the full sample period (1997 – 2010).

Appendix Table 12 replicates our core analysis linking ethnic partitioning to conflict intensity, as reflected in the number and incidence of deadly events, fatalities, and conflict duration. [This table "mirrors" Table 3 in the main part of the paper]. There is a strong link between ethnic partitioning and deadly conflict, as well as deaths and duration. The linear probability estimates suggest that the likelihood of deadly conflict is 10 percentage points higher for partitioned (as compared to non-split) ethnic homelands. Moreover, on average conflict duration is at least 25% longer in the homelands of partitioned ethnic groups.

Appendix Table 13 associates battles, violence against civilians and riots and protests with ethnic partitioning. There is strong link between battles and ethnic partitioning both in the full sample (columns (1)-(2)) and when we restrict estimation to ethnic areas near the national border (in (7)-(8)). Both the NB and the LPM estimates in columns (3) and (4) show that civilian violence occurs more often in the homelands of partitioned ethnicities. The results are similar in the border sample, though in the LPM the coefficient on ethnic partitioning is marginally insignificant as the standard error increases (*t*-stat 1.6). In line with our results in the main text, ethnic partitioning is not systematically linked to riots and protests, which are usually non-deadly events taking place mostly in the capitals.

Appendix Table 14 repeats our analysis linking conflict by actor to ethnic partitioning focusing on the period 1997 – 2010. [The table "mirrors" Table 5 in the main part of the paper.] The table summarizes our key findings. There is a strong impact of ethnic partitioning

¹⁰The estimates in columns (10) and (11) are identical because all outliers (observations where all conflict incidents exceed the 99th percentile) are not in the border sample. The border sample is somewhat smaller than 606 observations, because there is no variability on ethnic partitioning for some countries when we zoom in the border.

on civil conflict by government forces or rebels that oppose the state. Ethnic partitioning is also linked to a higher incidence and intensity of violence against the civilian population. In contrast, partitioning is unrelated to riots and protests. The specifications further show that foreign interventions from neighboring countries are way more common in the homelands of split ethnicities, while in contrast peace-keeping interventions and non-African interventions (mostly by NATO and European powers) are unrelated to ethnic partitioning.

3.2.2 Conditional Negative Binomial ML Estimates

Second, we estimated the baseline econometric specification with the conditional negative binomial model of Hausman, Hall, and Griliches (1984) to control for country unobservables. This technique accounts for unobserved country heterogeneity, by parametrizing the over-dispersion parameter (to have a country-specific component). As such, this is not a "pure" fixed-effects approach (based on mean parametrization) and this is why it is not commonly used.

Appendix Table 15 - Panel *A* gives the results. The conditional NB estimates show that ethnic partitioning is systematically linked to civil conflict. The coefficient is around 0.25 in the full sample and around 0.40 in the border sample, implying economically large effects. The spillover variable (*SPIL*) enters also with a significant estimate indicating that conflict is higher in the homelands that neighbor partitioned groups. In Appendix Table 16 - Panel *A* we also report conditional NB estimates linking each type of conflict (battles, civilian violence, and riots-protests) to ethnic partitioning. The results show a strong link between ethnic partitioning and battles between government forces, rebels and militias and also show a significant impact of ethnic partitioning on violence against the civilians.

3.2.3 Accounting for Outliers. Poisson ML Estimates

Third, to account for outliers in the number of conflict events (see Appendix Table 2) and the highly non-linear nature of the outcome variable, we estimated country-fixed-effects Poisson (ML) estimates dropping the top 5% of the dependent variable; when we do so, the Poisson model assumption of mean and variance equality approximately holds.

Appendix Table 15 - Panel *B* reports the estimates. *SPLIT* enters with a highly positive estimate, reassuring that our results are not driven by outliers. *SPIL* enters also with a positive coefficient, which however is not always significant at standard confidence levels. The results in Appendix Table 16 - Panel *B* further show that the link between partitioning and battles (mostly between government troops, rebels and militias) and civilian violence is quite robust and not-driven by outliers. Appendix Table 17 reports country-fixed-effects Poisson ML estimates focusing on the number of events by conflict actors. The results are similar to the NB estimates

in the main part of the paper (Table 5). Ethnic partitioning is systematically linked to conflict involving government troops or state-backed militias (columns (1) and (7)) and rebels with an explicit agenda to counter the state via violent means (columns (2) and (8)). There is also a significant link between ethnic partitioning and violence against civilians (columns (4) and (10)). Perhaps more importantly the Poisson specifications show that foreign interventions from nearby countries are more frequent in partitioned homelands (columns (5) and (11)). This suggests that partitioned ethnic groups are often used by governments of neighboring countries to wage conflict on the other side of the border. In contrast, there is no link between ethnic partitioning and interventions from non-adjacent nations (usually international peace-keeping missions) (columns (6) and (12)).

3.2.4 Not Accounting for Spillovers

Forth, while the analysis reveals considerable spillovers of ethnic partitioning to nearby regions (especially when focusing on the likelihood of conflict and duration), we also estimated specifications without accounting for the share of nearby ethnic groups that have been partitioned (*SPLIT*). Appendix Table 18 reports the results. *SPLIT* enters with a positive and significant coefficient in all permutations both when we we examine conflict intensity (in Panel *A*) and the likelihood of conflict (in Panel *B*). The LPM estimates imply that the likelihood of conflict is 8 percentage points higher in the homelands of partitioned as compared to non-split groups.

3.2.5 Alternative Index of Ethnic Partitioning

Fifth, we repeated estimation with an alternative measure of ethnic partitioning using a 5% threshold to identify split groups. Appendix Table 19 gives the results using as the dependent variable the number of main conflict events (in Panel *A*) and an indicator that switches to one when a country-ethnic homeland has experienced a main conflict over the sample period (in Panel *B*). The ethnic partitioning measure enters with a positive and highly significant coefficient, implying that a higher degree of ethnic partitioning is associated with a higher likelihood of civil conflict. The linear probability model estimates imply that compared to non-split homelands, the likelihood of a main conflict is approximately 8% higher in the ancestral homelands of split ethnicities.

3.2.6 Further Accounting for Overall Border Effects

Sixth, to further isolate the impact of ethnic partitioning from an overall border effect (which, nevertheless, could still be driven by ethnic partitioning itself) we augment the specification

with higher-order polynomials in distance to the national border. Appendix Tables 20 and 21 report unconditional NB ML specifications with country constants (in Panel *A*) and country fixed-effects linear probability model (LPM) estimates (in Panel *B*) adding a third-order and a fourth-order polynomial on distance from the centroid of each country-ethnic area to the border, respectively. The ethnic partitioning index enters with a highly significant coefficient both in the NB-ML and the LPM specifications in the full sample of country-ethnic homelands (in column (1)). The estimate retains significance when we restrict estimation to the border sample (in (5)). *SPIL* also enters with a positive (and in many specifications significant) coefficient implying sizable spillovers of ethnic partitioning to nearby regions. The LPM specifications imply an 8% – 9% increased likelihood of conflict in the homelands of split ethnicities. And the probability of conflict further increases by 8% when half of the neighboring ethnic homelands are partitioned (as compared to ethnic homelands where none of the nearby groups is partitioned). In line with our results in the main text, ethnic partitioning has strong effects on battles between government forces, rebels and militias and violence against the civilian population. In contrast, there is no link between partitioning and riots and protests.

3.2.7 Ethnic Family Fixed-Effects Specifications

Seventh, to minimize concerns that our estimates are driven by unobservable features -related perhaps to ethnic social and political features or hard-to-measure geographic factors (such as presence of palm trees or other agricultural endowments that were important during colonization- we also estimated specifications adding on top of the country constants ethnic-family fixed-effects. While these models are very restrictive (as we have 96 ethnic families), they further account for local conditions and broad cultural, institutional, and other hard-to-observe ethnic-family-specific factors. Examples of ethnic families include the Bedouin Arabs, the Tuareg, and the Southwestern Bantu. Appendix Table 22 reports NB ML specifications with the number of conflict events on the LHS (in (1) and (4)), linear probability model estimates with the binary index of conflict as the dependent variable (in (2) and (5)), and Poisson ML specifications with the duration of conflict (in years) as the dependent variable. The NB estimate of the ethnic partitioning index is highly significant, reassuring that the baseline results are not driven by some hard-to-account-for ethnic family factor. The LPM estimates suggest that, even when we solely examine within-country, within-ethnic-family variation, civil conflict is significantly more likely to occur in border areas where partitioned ethnicities reside. The Poisson fixed-effects specifications further show that ethnic partitioning is associated with more prolonged warfare. Moreover, the coefficient on *SPIL* is positive (and in the LPM significant at the 99% confidence level), implying sizable spillovers.

Perhaps more importantly, the coefficient on the ethnic partitioning index remains stable across these restrictive specifications. The NB estimates in the country-fixed-effects specifications are 0.34 and 0.46 with the simple and the rich set of controls, respectively (see Table 2). While standard errors rise, adding the ethnic family fixed effects yields a comparable estimate, 0.36. The same applies to the LPM specifications. The coefficient on *SPLIT* in the specifications with only country fixed effects range from 0.066–0.089. This is similar to the specification that adds the 96 ethnic family constants, 0.057. This implies that it is unlikely that unobservable features -correlated with both partitioning and conflict- can explain the strong negative association (see Altonji, Elder, and Taber (2005)).

3.2.8 Regional Effects

Eighth, we dropped iteratively ethnic homelands from each of the five main African regions so as to investigate whether the results are driven by a particular part of the continent.¹¹ Appendix Table 23 reports NB-ML (in Panel *A*) and LPM estimates (in Panel *B*). In (1)-(2) we exclude North Africa to account for the fact that Europeans had contacts with the northern part of the continent since the ancient times. In (3)-(4) we drop Southern African countries. In columns (5)-(6) we drop Western African countries, because a few of the contemporary African borders in this region correspond to internal administrative borders of the Federation of the French West Africa. In (7)-(8) and (9)-(10) we exclude ethnic areas in East Africa and Central Africa, respectively. This allows us to examine the robustness of our results to influential observations, as the most deadly and prolonged conflicts have taken place in Ethiopia, Somalia, Sudan, Rwanda, and the Democratic Republic of Congo. Moreover, since most large pre-colonial African states were in Ethiopia and other parts of Eastern Africa (Besley and Reynal-Querol (2014)), we further account for the role of pre-colonial conflict and political development. The results show that the strong positive effect of ethnic partitioning on civil conflict is not driven by a particular region though the link at the intensive margin appears stronger for split groups in Central African countries.

3.2.9 Spatial Models

Ninth, we estimated spatial models to account formally for spatial spillovers and for factors in nearby ethnic regions.¹² Appendix Table 24 reports spatial lag specifications (estimated by maximum-likelihood). Since spatial lag models are somewhat sensitive to the specification details, we report estimates using two different types of (inverse) weighting matrixes (linear in

¹¹The regional classification follows Nunn (2008).

¹²Harari and La Ferrara (2014) conduct a thorough exploration of the link between weather shocks and conflict across Africa using such state-of-the-art spatial econometric empirical models.

distance and quadratic in distance) and three different models. In Panel *A* the dependent variable is an indicator that switches to one for country-ethnic homelands that have experienced a main conflict incident over the sample period and zero otherwise. In Panel *B* the dependent variable is the natural logarithm of one plus the total number of main conflict events. Accounting for spatial spillovers in conflict as well as the effect of the independent variables in the homelands of neighboring groups (in the Durbin and Generalized Spatial models) has little effect on the coefficient of ethnic partitioning that retains its economic and statistical significance. In line with our baseline estimates, the LPM estimates reveal that the probability of conflict is 7% – 9% higher in the homelands of partitioned (as compared to non-split) ethnicities.

3.2.10 Modelling Spillovers Along Ethnic Family and Country Lines

Tenth, in Appendix Table 25 we further explore the impact of spatial spillovers running specifications linking conflict to ethnic partitioning conditioning on the total number of conflicts in the same country (netting out conflict in each ethnic homeland) and conditioning on total conflict of each ethnolinguistic family (netting out conflict of each ethnic homeland). The cross-sectional estimates (with regional constants) reveal that conflict is higher when an ethnic homeland falls in a conflict-prone country and when there is a lot of conflict in groups from the same family and groups from the same country. These results are quite interesting as they show that conflict and violence against civilians spread *both* along country and ethnic family lines. Yet, this has little impact on the ethnic partitioning index that retains its economic and statistical significance.

3.2.11 Accounting for Pre-colonial Conflict

Eleventh, we estimated specifications linking contemporary conflict to ethnic partitioning, controlling for the historical legacy of violence. Appendix Table 26 reports the results. Following Besley and Reynal-Querol (2014), in columns (1) and (4) we augment the empirical specification with the distance of the centroid of each ethnic homeland to the center of the closest pre-colonial conflict. Given the lack of association between pre-colonial conflict and ethnic partitioning (see Table 1 in the main part of the paper), it should not come as a surprise that the estimate on the partitioning index retains its economic and statistical significance across all permutations. In columns (2) and (5) we control for the impact of conflict during the slave trades augmenting the specification with the estimates of Nunn (2008) and Nunn and Wantchekon (2011) on the number of slaves at the ethnicity level. The coefficient on the partitioning index remains intact suggesting that the latter does not capture conflict during the pre-colonial period. Since many

wars before colonization took place in the territories of large centralized kingdoms, in columns (3) and (6) we include in the specification a dummy variable that takes on the value of one if an ethnic homeland was part of a large pre-colonial kingdom (using data from Besley and Reynal-Querol (2014)). The coefficient on the ethnic partitioning index remains positive and highly significant.

3.2.12 Accounting for Income Differences

Twelfth, given the significant negative association between conflict and economic development (see Collier and Hoeffler (2007) and Blattman and Miguel (2010) for reviews and Miguel, Satyanath, and Sergenti (2004) for a seminal contribution on identification), we examined whether ethnic partitioning affects civil conflict, conditional on income or whether the partitioning-conflict nexus operates via income. In Appendix Table 27 we report specifications associating all conflict, battles, violence against civilians, and riots and protests with ethnic partitioning, conditioning on the log of per capita GDP at the ethnic homeland level using data from the *G-Econ* project (Chen and Nordhaus (2011)). While in some specifications income enters with a significant negative coefficient, this has no effect on the estimate on the split indicator that retains its significance.¹³ The coefficient on the ethnic partitioning index is almost identical to the analogous estimates in Tables 2 – 3. Moreover, ethnic partitioning itself is not systematically linked to regional income (results not shown).¹⁴ This set of results, therefore, suggests that the effect of ethnic partitioning on conflict does not operate through an *overall* decline in economic performance of partitioned areas. Appendix Table 28 reports linear probability model estimates examining the incidence of conflict by actor and ethnic partitioning. Ethnic partitioning is associated with a much higher likelihood of civil conflict with active government troops participation and with rebel fights. Moreover foreign interventions from neighboring countries' troops is much more likely to take place in the homelands of split, as compared to non-split, ethnicities.

3.2.13 Measurement Error in the Conflict Data

Finally, to account for error-in-variables on ACLED and UCDP, we defined a trichotomous and a binary conflict measure combining the two datasets. We defined an ordered variable ranging from 0 – 2 that takes on the value of two when both databases record conflict in a given ethnic

¹³In these specifications we lose one observation (Tajakant in Mauritania), as the *G-Econ* project does not give local GDP estimates.

¹⁴We also estimated specifications proxying regional income with satellite images on light density at night (as in Michalopoulos and Papaioannou (2014)) finding similar results. Overall ethnic partitioning is negatively associated with luminosity at the country-ethnic homeland level, though the correlation is insignificant in most (though not all) specifications.

area, a value of one when country-ethnic areas have experienced conflict based on only one of the two databases, and a value of zero when both databases indicate the absence of conflict. We also defined a binary index which equals to zero when either of the two (or both) databases indicate that no conflict has occurred. We exclude 2011, 2012, and 2013 as UCDP stops in 2010 and focus on main conflict incidents, excluding riots and protests, that are covered only by ACLED.

Appendix Table 29 reports the results. The ethnic partitioning index enters the specifications in the full sample with the trichotomous conflict measure with a positive and highly significant coefficient; the estimate (0.125) implies that the likelihood to observe conflict in both databases in a split country-ethnic area is approximately 25% higher as compared to non-split homelands. The coefficient is similar when we restrict estimation to border groups (0.13 – 0.14) though standard errors increase rendering the estimate insignificant (t -stat 1.5 – 1.6). The linear probability model estimates imply that the likelihood of conflict being reported in both civil conflict databases is 6.5% – 9% higher for partitioned as compared to non-split ethnic homelands.

Likewise, we defined binary and trichotomous indicators combining ACLED’s info of conflict involving government forces with UCDP state-based conflict and we also defined binary and trichotomous indicators of civilian violence combining ACLED and UCDP corresponding categories. Appendix Table 30 gives the results. In line with the evidence in the main part of the paper, the link between ethnic partitioning and civil conflict where government forces and state-backed militias are involved is strong. There is also a somewhat weaker (though still significant) association between violence against the civilian population and ethnic partitioning.

4 Conflict in East-Central Africa: A Case Study

In this section we briefly discuss conflict in East-Central Africa, one of the most conflict-prone regions in the world, as it illustrates the link between ethnic partitioning, violence against civilians and battles between government forces, rebels, and militias (see Figure 7).

Let us start from Tanzania, a country with little conflict overall; in the 69 ethnic regions of Tanzania ACLED records 250 main conflict incidents. Namely, 46 battles and 204 events of violence against civilians over the period 1997 – 2013. The mean (median) conflict per ethnic homeland is 3.6 (0). Most conflict occurs at the border with Rwanda and Burundi where the partitioned Rundi group resides. While in the Rundi homeland only 1.3% of Tanzania’s total population lives (in both 1960 and in 2000), we observe 21 conflict incidents (8.5%). There are zero and two conflict events in the non-split homelands of the Bende and the Fipa, respectively, although both groups reside on the border with the Democratic Republic of Congo,

the country with the highest conflict intensity in Africa. This is because lake Tanganyika serves as the natural border between the two countries.

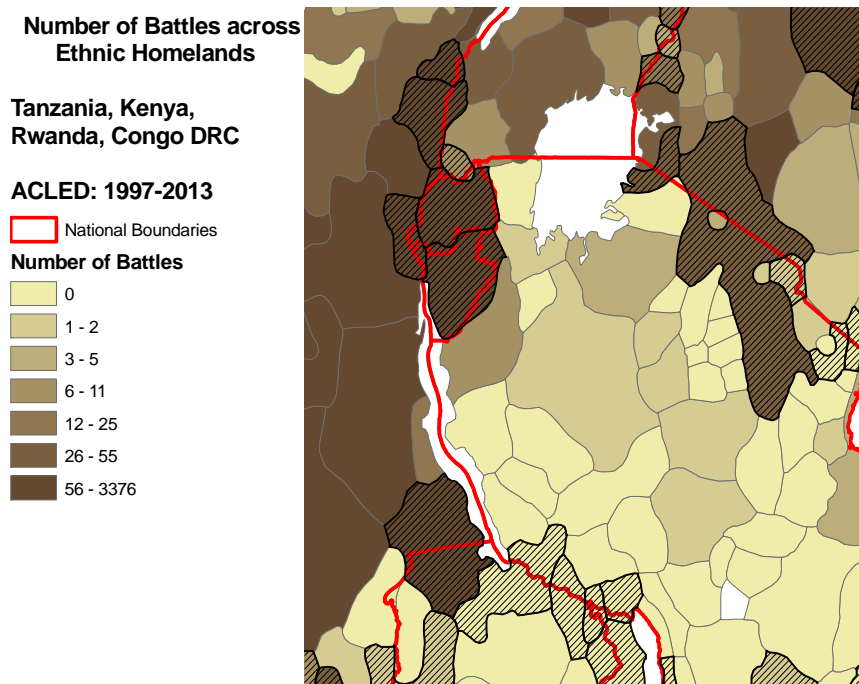


Figure 7

Focusing now on the Democratic Republic of Congo (DRC), there are 5,473 main conflict events (excl. riots and protests) across the 102 ethnic regions (mean=52.6; median=8). In the homelands of the three partitioned Rwandan ethnicities (of the Interlacustrine Bantu - Ruanda family) we observe 1,239 incidents (more than 20% of all conflict), while the share of population residing in the homelands of the Bashi, the Konjo, and the Ruanda is around 6%; and in the adjacent non-split, but ethnically similar Hunde we have 293 battles and 159 events of civilian violence. So, despite the small share of DRC's population living in the Hunde homeland (is around 2.2%), we observe close to 10% of conflict. The UCDP data reveal a similar picture. In the three Rundi homelands we observe 19% of all state conflict and 34% of one-sided violence, while in the nearby ethnic region of the Hunde 14% of all state conflict and 19% of all one-sided violence events occurs. Moreover in the homelands of the Rundi tribes we observe many foreign interventions from the government troops of Rwanda and Uganda.

So, while the origins of conflict in Eastern Congo are complex (involving also within-ethnic group conflict) and by no means solely driven by partitioning, nevertheless, it involves ethnic militias (such as the FDLR) constantly moving across the border between Rwanda, Democratic Republic of Congo and Uganda within partitioned groups to re-group, re-organize

and re-arm.¹⁵ Moreover partitioned ethnic groups are used by governments in nearby countries (Uganda, Rwanda) to intervene. In contrast, there is little conflict at the (non-split) Holoholo ethnic homeland on the Western (Zairian) bank of the Tanganyika lake (in total 28 main events in ACLED and 14 events in UCDP), which seems to be driven by spillovers (as half of this group's neighbors have been split). In the adjacent to the Holoholo homeland in the South, in the Ta(a)bwa region we observe 121 conflict events (2.2%) and 5% of all state-conflict in UCDP, while the share of DRC's population in this region is just 0.8%. This is not surprising since, although the Eastern border of Congo with Tanzania is organic (Tanganyika lake), the Southern one with Zambia follows a straight (latitudinal) line that splits the homeland of the Ta(a)bwa almost equally between DRC and Zambia.

A Note It is important to note that the map we use for the location of ethnicities was published in 1959. According to the documentation the distribution of ethnic homelands corresponds to the late decades of 1800s and early decades of 1900s. Hence, one cannot entirely rule out the possibility that conflict-prone groups might have crossed the borders after their delineation in the various conferences in Europe, but before Murdock's map was drawn. Take the Banyamulenge, for example. After the borders were decided in 1890 in Brussels, there were indeed several migration flows of people of Rwandan origin towards the Kivus including Rwandese who came to the Democratic Republic of the Congo to work in the mines in Katanga as well as those who fled to the Democratic Republic of the Congo during the Rwandan Hutu revolution (1959 – 1962). However, it is important to keep in mind that peoples' movements often follow settlement patterns that predate the colonial era. This is most likely the case for the Banyamulenge. According to the historian Alexis Kagame and the famous French-American political scientist Rene Lemarchand the pre-colonial arrival of Tutsis in the Kivus meets general agreement among historians but is "*vehemently contested, however, by many Congolese intellectuals*" (Lemarchand (2008)).

5 Ethnic Partitioning and Civil Conflict. Heterogeneous Effects

Our results suggest that ethnic partitioning has a strong effect on civil conflict, on average. Yet it may be the case that ethnic partitioning is especially harmful in certain environments.

¹⁵The FDLR and other Hutu-based militias, which fled Rwanda after the 1994 genocide and sought shelter in their homeland in Eastern Congo, had played a major role in the recent civil conflict in Congo. For example, in early 2007 in just one event, FDLR groups raided two villages killing 17 civilians and wounding 19. In the next day FDLR militias looted 18 houses in a nearby town. Also in a single event in Fendula in Eastern Congo, (at least) 30 civilians were burned alive and 50 wounded by Rwanda militias. Prunier (2009) provides a detailed narrative of how the partitioning of the Rwandan tribes and the genocide in Rwanda spread to Congo.

Based on previous works on the origins of African development and the historical narrative of the influence of the Scramble for Africa on conflict, in Appendix Tables 31 and 32 we explore potential heterogeneous effects of ethnic partitioning. Searching for heterogeneous effects is also helpful as it sheds light on how and under which circumstances ethnic partitioning spurs conflict.

5.1 Heterogeneity w.r.t. Group Characteristics

In Appendix Table 31 we examine whether specific ethnicity-country-specific characteristics mitigate or accentuate the adverse effect of ethnic partitioning on civil conflict. Panel *A* reports negative binomial maximum likelihood estimates (NB) with country constants, whereas in Panel *B* we focus on the extensive margin of conflict estimating linear probability models (LPM) with country fixed effects. In each specification we allow the coefficient on the binary ethnic partitioning index to vary depending on the intensity of the underlying characteristic.

In column (1) we explore whether the relationship between ethnic partitioning and conflict depends on whether a partitioned ethnic region is surrounded by a high or a low share of other split groups.¹⁶ The estimated coefficients are similar in the two sets of split groups. This applies both in NB and the LPM specifications. So the adverse impact of partitioning on local conflict is not systematically influenced by the intensity of adjacent split groups.

In column (2) we examine whether relatively large partitioned ethnic groups have managed to escape the negative repercussions of partitioning. In this regard we allow the impact of ethnic partitioning to differ for country-ethnic groups that are above and below the median population share in the respective country. The coefficient on *SPLIT* is positive and highly significant for both sets of split ethnicities.

In column (3) we investigate whether the partitioning civil conflict link is affected by the (population) size of the group's partition on the other side of the border. These specifications ask whether conflict is more frequent in partitioned regions whose brethren on the other side of the border is relatively large or small in the neighboring country.¹⁷ This test is related the argument of Esteban and Ray (2008, 2011) that large within-ethnic group disparities in income and political power increase the likelihood of civil war (Huber and Mayoral (2014) present supportive to this idea cross-country evidence). The results show that the partitioning - civil conflict nexus is independent of the relative size of the co-ethnics in neighboring countries.

In column (4) we explore whether the impact of partitioning is mediated by the size of ethnically similar groups in the same country. In this regard we allow the ethnic partitioning

¹⁶These models can be thought as examining possible interactive effects between the ethnic partitioning index (*SPLIT*) and the spillovers measure (*SPIL*).

¹⁷In case of multiple (more than two) partitions we use the population share of the largest partitioned area.

index to differ for groups where the share of groups in the same ethnic family (cluster) is relatively high and relatively low in each country. Ethnic partitioning is systematically linked to conflict for both sets of ethnicities and there is not much heterogeneity.

In column (5) we address a similar question. Yet we now investigate the role of ethnic similarity in the immediate neighborhood of each partitioned group, allowing *SPLIT* to differ for ethnicities where most of nearby ethnicities are from the same family and those split groups where most neighbors are from different ethnic families. The NB estimates in Panel *A* reveal that split groups surrounded by a large share of ethnically similar entities (i.e., groups that belong to the same ethnolinguistic family) experience on average more conflict.¹⁸ Yet the LPM estimates in Panel *B* do not reveal any differences on the impact of ethnic partitioning with respect to the share of ethnically similar neighboring groups. So overall there is not much heterogeneity.

In column (6) we examine whether being split by a squiggly or a straightline-like border affects the relationship between partitioning and conflict. Following Alesina, Easterly, and Matuzeski (2011) we construct the fractal dimension of each border segment using the box-count method that identifies straight versus squiggly borders. The intuition behind this measure is that straight-line like borders are more likely to be arbitrary compared to squiggly ones. We allowed the impact of *SPLIT* to differ for (two-way) partitioned groups that are split by relatively straight borders (fractal measure below the median) and by more wavy boundaries (fractal measure above the median). The NB estimates show that ethnic partitioning is systematically linked to civil conflict irrespective of the shape of the boundary. The LPM estimates yield weak evidence that the impact of ethnic partitioning is slightly larger for groups split across straight-line borders. Yet the coefficients are not statistically different from each other. Moreover reasonable permutations (dropping outliers, estimating Poisson ML models, etc.) do not yield major differences on the coefficient of *SPLIT* in the two set of partitioned groups.

In column (7) we allow the coefficient on the ethnic partitioning index to differ for groups split across national borders separating colonial powers (e.g., Niger-Nigeria border that separated the French and the British colonies) or separating countries that were ruled by the same colonial power (e.g., Niger and Mali that were part of French West Africa). Roughly 65% – 70% of split groups in Africa are separated by borders that were separating colonial powers. The remaining 30% – 35% of partitioned groups are split by political boundaries that were internal administrative borders (as those in French West Africa). The NB estimates in

¹⁸This pattern may look puzzling at first-glance, as one may expect more heterogeneous neighbors to intensify the impact of partitioning on conflict. Nevertheless, this finding is in line with Spolaore and Wacziarg (2014) who show in a cross-country setting that genetically similar populations are more prone to go to war with each other (even when one conditions on distance, ties, and other similarities).

(Panel *A*) suggest that there is a differential pattern between these two broad sets of groups. Ethnicities split between colonial powers are significantly more likely to engage in conflict compared to groups that came to be split within the borders of a given colonizer. Nevertheless, this differential pattern disappears when we look at the probability of conflict (in Panel *B*).

In column (8) we examine whether being split in two countries or in three or more countries affects the ethnic partitioning conflict nexus. In our sample, 69% of partitioned groups are split across 2 contemporary countries with the remaining 31% of the sample consisting of groups split between three, four, five and six countries. The estimates do not reveal major differences between the two sets of partitioned ethnicities. Ethnic partitioning is associated with more civil conflict incidents and a higher likelihood of conflict for both sets of split groups.

Overall the results in Panels *A* and *B* of Appendix Table 31 do not reveal much heterogeneity with respect to group characteristics related to their size, the ethnic similarity of their neighbors, the straightness of borders or the degree of partitioning.

5.2 Heterogeneity w.r.t. Country Characteristics

The Scramble for Africa has led to the creation of some very large and heterogeneous states. Moreover, the border design resulted in Africa having the most landlocked countries across all continents. In Appendix Table 32 we examine whether these aspects of the Scramble for Africa interact with ethnic partitioning. Specifically, we explore how various aspects of country-level diversity, polarization, location and size modify the role of partitioning. Panel *A* reports negative binomial maximum likelihood estimates looking at conflict events, whereas Panel *B* gives linear probability model estimates focusing on the extensive margin of conflict.

In columns (1)-(3) we investigate whether the impact of ethnic partitioning differs with respect to the degree of ethnic fractionalization; we do so allowing the coefficient on the ethnic partitioning index to differ for countries scoring above and below the median of an index of ethnolinguistic fractionalization (from Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003)), ethnic fragmentation and polarization at a very fine ethnic level (using data from Desmet, Ortuño-Ortín, and Wacziarg (2012)). While the NB estimates in Panel *A* show that the coefficient on *SPLIT* is more significant in fragmented countries, the LPM specifications (that are less sensitive to outliers) do not reveal any differences. So overall the estimates show that the impact of ethnic partitioning is quite homogeneous and does not depend on the degree of fractionalization.¹⁹

¹⁹The results are similar when we group countries based on the degree of cultural fractionalization (using data from Fearon (2003)), religious polarization (using the index of Montalvo and Reynal-Querol (2005b)), ethnic segregation (using data from Alesina and Zhuravskaya (2011)) or ethnic inequality (using data from Alesina, Michalopoulos, and Papaioannou (2015)).

In column (4) we examine whether ethnic partitioning has a differential impact on conflict on landlocked countries (e.g., Chad, Burkina Faso, Rwanda, Mali, the Central African Republic). The NB and the LMP results suggest that ethnic partitioning is particularly deleterious for ethnicities groups in landlocked countries.

In column (5) we allow the impact of the ethnic partitioning index (*SPLIT*) to differ for large in terms of size (land area) and small countries (using the median value of land area as a cutoff). The NB-ML specifications reveal that the impact of ethnic partitioning is more pronounced in large countries, such as Zaire, Sudan, and Angola, as compared to smaller one. Yet the LPM estimates on *SPLIT* are similar for large and small countries.

6 EPR-based Analysis: Sensitivity Checks

Appendix Tables 33, 34 and 35 report further evidence illustrating the link between ethnic partitioning, political discrimination, and participation in ethnic wars using the Ethnic Power Relations database of Wimmer, Cederman, and Min (2009); this database provides assessments of formal and informal degrees of political participation and exclusion at the ethnic level for all politically relevant groups over the period 1960 – 2010.

6.1 Alternative Index of Ethnic Partitioning

In Appendix Table 33 we report linear probability model (with country fixed-effects) estimates associating political discrimination (in (1)-(4)) and participation in ethnic civil wars (in (5)-(9)) at the ethnicity-country level to ethnic partitioning using the 5% threshold of a group's homeland to identify split groups. In line with the baseline estimates (in Table 7) the binary ethnic partitioning index enters with a positive and highly significant coefficient. The estimate suggests that the likelihood of being discriminated from the national government is approximately 8% – 10% higher for partitioned ethnicities. Moreover, the probability to engage in a major civil war with an explicit ethnic dimension is 10 percentage points higher for split groups. This estimate is quite similar to the baseline estimates linking ethnic partitioning to state-based conflict (in UCDP) and conflict where government troops participate (in ACLED).

6.2 Regional Effects

Similar to our analysis linking partitioning with conflict, we examined whether a particular region drives the association between ethnic partitioning, discrimination, and ethnic wars dropping iteratively homelands from each African region. Appendix Tables 34 and 35 report the results using the baseline ethnic partitioning index and the alternative partitioning index

based on the 5% cutoff. The results show that ethnic partitioning is associated with an increased likelihood of political discrimination from the national government and an increased participation at ethnic-based civil wars across all African regions.

7 DHS-Based Analysis: Further Evidence and Sensitivity Checks.

7.1 Descriptives

Appendix Table 36 - Panel *A* presents descriptives on the Demographic and Health Surveys (DHS) sample. In total we have information on 88,171 individuals from twenty Sub-Saharan African countries.²⁰ 25,631 respondents reside in partitioned homelands and self-identify with split groups (29%). 36,694 individuals reside outside the homelands of split groups and do not self-identify with partitioned ethnicities (41.6%). 13,256 respondents reside in partitioned ethnic homelands but do not identify with partitioned groups (15%), while 12,590 identify with split ethnicities, while residing in non-partitioned ethnic homelands (14.3). Appendix Table 36 - Panel *B* gives summary statistics for the main variables employed in the empirical analysis. Appendix Figure 8 gives the spatial distribution of respondents.

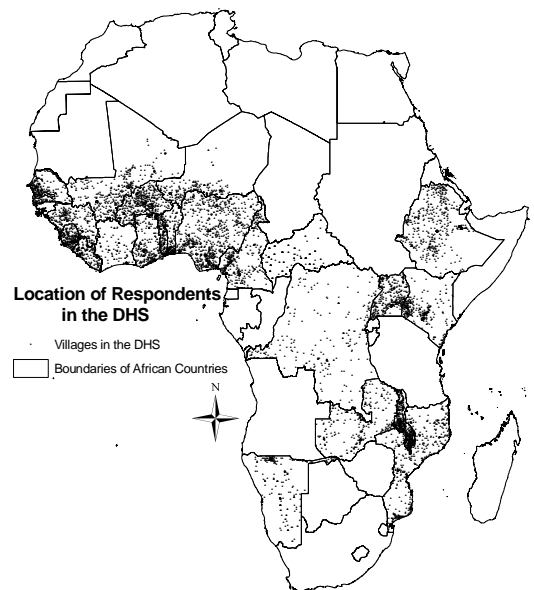


Figure 8

²⁰The countries and interview years are Benin in 2001, Burkina Faso in 2010, Central African Republic in 1994, Ethiopia in 2011, Ghana in 2008, Guinea in 2005, Kenya in 2008, Mali in 2006, Mozambique in 2011, Malawi in 2010, Namibia in 2000, Niger in 1998, Senegal in 2010, Sierra Leone in 2008, Togo in 1998, Uganda in 2011, the Democratic Republic of Congo in 2007, and Zambia in 2007.

7.2 Enumeration Area Fixed-Effects Estimates

We repeated estimation exploiting within DHS enumeration-area (village/town/city) variation. Estimating enumeration-area-fixed effects specifications allows controlling for (observed and unobserved) local factors at a very fine level; as such the ethnic partitioning indicator isolates the impact of respondent's ethnic origin. Appendix Table 37 reports the within-enumeration-area estimates. These extremely restrictive specifications -that include more than 4,500 city/town constants- assess whether differences in wealth and education across ethnic lines are related to ethnic partitioning among individuals residing in the same location. In spite of the efficiency loss (as most of the variation in both the explanatory variables and the ethnic partitioning index is absorbed by the village fixed effects), the coefficient on the ethnic partitioning indicator is negative and significant in all but one specification. These estimates point out that even when we compare respondents currently residing in the same village, those of a partitioned ethnic background have lower access to utility services and are less educated.

7.3 Looking at Movers

The specifications in Table 9 of the main text leverage the presence of "movers" (i.e., members of split groups residing outside their ancestral homelands) for identification. However, while in all specifications we control for whether an individual resides outside his ancestral homeland, there may be unobserved or hard-to-account for differences between "movers" and "non-movers". Hence, we estimated the link between public goods and ethnic partitioning looking solely at individuals that reside outside their group's ancestral homeland. Appendix Table 38 reports the results. In columns (1)-(2) and (7)-(8) we focus on the full sample of individuals that currently reside outside their tribal ancestral homelands. Even if one looks only at individuals who no longer reside in their groups' ancestral homelands, members of partitioned groups are systematically worse off. In an effort to further account for "selection" into location, we also examine the role of ethnic partitioning for movers who currently reside either in partitioned or non-partitioned homelands. In columns (3)-(4) and (9)-(10) we compare migrants from split ethnicities living in non-partitioned homelands to migrants from non-split groups also living in non-partitioned homelands (different from theirs). Individuals identifying with partitioned groups have lower access to public goods and education as compared to individuals identifying with ethnicities whose ancestral homeland boundaries were not disrupted by the colonial border design. A similar pattern emerges when instead of looking at "movers" within non-partitioned homelands, we look on "movers" residing in partitioned homelands (in (5)-(6) and (11)-(12)). Although the sample drops considerably, the negative impact of ethnic partitioning remains quite robust. These results that further account for unobservables and location further show

that ethnic partitioning is systematically linked to under-provision of public goods and lower levels of education.

7.4 Persistence

The results linking ethnic partitioning to conflict and education/public goods point out that the Scramble for Africa has had sizable long-run effects on economic development. Since the mid-nineties many African countries have made considerable steps towards democratic institutions, while over the past decade growth has been high across the continent. Hence, it is important from a policy standpoint to assess whether the recent economic and political modernization efforts have been associated with convergence in economic conditions between members of split and non-split groups. To shed light on this question we exploit the fact that the DHS reports respondents' age cohort and repeat estimation separately for "old" and "young" cohorts. Appendix Table 39 reports the results. In odd-numbered columns we restrict estimation to "young" respondents, while in even-numbered columns we restrict estimation to "old" respondents; as a cut-off value we use those born before and after 1977 which is the median date of birth across respondents (the pattern is similar when we use 1975, 1980, or 1985). The results suggest that the negative effect of ethnic partitioning is strong both across "young" and "old" cohorts. Moreover, the estimate on the ethnic partitioning index is similar across the two sub-samples, revealing that the adverse effects of partitioning are quite persistent.

7.5 Summary

The evidence presented above that partitioned groups are more likely to have been victims of active discrimination by the government and more likely to have staged ethnic wars complements the findings from the DHS that documents a persistent economic disadvantage among those individuals identifying with partitioned groups. Taken together the empirical regularities uncovered, point out that the consequences of the Scramble for Africa, as manifested in the event of partitioning, are not circumscribed by the contours of a given ancestral ethnic homeland, but have significant repercussions for the members of partitioned groups irrespective of their whereabouts.

8 Data Sources and Variable Definitions

8.1 Main

Partitioning Index (SPLIT): Indicator variable that equals 1 if at least 10% of the historical homeland of an ethnic group is partitioned into different countries. We also construct an alternative partitioning index that equals 1 if at least 5% of the historical homeland of an ethnic group is partitioned into different countries. *Source: Calculated intersecting Murdock's (1959) ethnic map of Africa with the Digital Chart of the World (DCW) shapefile. The latter contains the polygons delineating the international boundaries in 2000. Appendix Table 1 reports partitioned ethnicities.*

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

Conflict Indicator: Dummy variable that takes on the value of one if a country-ethnic area has experienced at least one high-precision conflict incident (of any type) over the period 1997 – 2013. *Source: ACLED 4.*

Deadly Civil Conflict Incidents: Sum of all high-precision civil conflict incidents that have resulted in at least one fatality for each of a total of 1,212 country-ethnic homelands. *Source: ACLED 4.*

Deadly Conflict Indicator: Dummy variable that takes on the value of one if a country-ethnic area has experienced at least one high-precision deadly conflict incident (of any type) over the period 1997 – 2013. *Source: ACLED 4.*

Duration Civil Conflict: Duration in years that a country-ethnic homeland has been in conflict (using all types of conflict events) over the period 1997 – 2013. *Source: ACLED 4.*

Duration Deadly Civil Conflict. Duration in years that a country-ethnic homeland has been in conflict that has resulted in at least one fatality (using all types of deadly conflict events) over the period 1997 – 2013. *Source: ACLED 4.*

Total Fatalities: Number of fatalities for each of a total of 1,212 country-ethnic homelands using all types of conflict incidents. *Source: ACLED 4.*

Battles: Total number of battles between two violent armed groups at the ethnic homeland (in each country for partitioned ethnicities). Battles include armed conflict where a control of the contested location does not change and conflict events resulting in a territorial change of control. We aggregate the data at the ethnic homeland level and at the country-ethnic homeland level. See Section 2 for details. *Source: ACLED 4.*

Violence against Civilians: Total number of violent events against civilians at the ethnic homeland (in each country for partitioned ethnicities). Violence against civilians occurs when any armed/violent group attacks unarmed civilians. Rebels, governments, militias, rioters can all commit violence against civilians. We aggregate the data at the at the country-ethnic homeland level. *Source: ACLED 4.*

Riots and Protesters: Total number of events corresponding to riots and protests at the ethnic homeland (in each country for partitioned ethnicities). We aggregate the data at the at the country-ethnic homeland level. *Source: ACLED 4.*

Income per capita: Log of per capita income (GDP) in 2000 at the country-ethnic homeland level. *Source: G-Econ Database. available at: <http://gecon.yale.edu/>*

8.2 Pre-colonial Ethnic Features from Murdock (1967)

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 (indicating nomadic, semi-nomadic, and semi-sedentary). *Source: Murdock (1967); variable code v30.*

Dependence on Agriculture: 0 – 9 scale index reflecting the intensity of agriculture. "It includes penetration of the soil, planting, tending the growing crops, and harvesting but not subsequent food preparation". The index equals 0 when there 0% – 5% dependence; 1 when there is 6% – 15% dependence; 2 when there is 16% – 25% dependence; 3 when there is 26% – 35% dependence; 4 when there is 36% – 45% dependence; 5 when there is 46% – 55% dependence; 6 when there is 56% – 65% dependence; 7 when there is 66% – 75% dependence; 8 when there is 76% – 85% dependence; and 9 when there is 86% – 100% dependence. *Source: Murdock (1967); variable code v5.*

Animal Husbandry: 0 – 9 index reflecting dependence on pastoralist activities, animal husbandry. The index equals 0 when there 0% – 5% dependence; 1 when there is 6% – 15% dependence; 2 when there is 16% – 25% dependence; 3 when there is 26% – 35% dependence; 4 when there is 36% – 45% dependence; 5 when there is 46% – 55% dependence; 6 when there

is 56% – 65% dependence; 7 when there is 66% – 75% dependence; 8 when there is 76% – 85% dependence; and 9 when there is 86% – 100% dependence. *Source: Murdock (1967); variable code v4.*

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

Inheritance Rule for Property: Indicator that equals 1 when some form of inheritance rule of real property (land) is present; the binary indicator equals zero when there is "absence of individual property rights". *Source: Murdock (1967); variable code v74.*

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

Class Stratification: Binary class stratification variable. A zero score indicates "*absence of significant class distinctions among freemen, ignoring variations in individual reputes achieved through skill, valor, piety, or wisdom.*" A score of 1 indicates either "*the presence of wealth distinctions, based on possession or distribution of property, which however have not crystallized into distinct and hereditary social classes*" or "*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*" or "*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*" or "*complex stratification into social classes correlated in large measure with extensive differentiation of occupational statuses.*" *Source: Murdock (1967); variable code v67.*

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 or when large extended families are present (and zero otherwise). *Source: Murdock (1967); variable code v8.*

8.3 Country-level Variables Used in Heterogeneous Effects Section

Land Area: Log of country's surface/land area. *Source: Nunn and Puga (2012).*

Ethnolinguistic Fractionalization: Index of ethnic/linguistic heterogeneity. It reflects the probability that two randomly selected individuals belong to different ethnolinguistic/religious groups. For completeness we use two measures, one from Alesina *et al.* (2003), which in turn is based on CIA Factbook and Encyclopedia Britannica and one from Desmet *et al.* (2012), which is based on *Ethnologue* (level 15). *Source: Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003) and Desmet, Ortuño-Ortín, and Wacziarg (2012).*

Ethno-linguistic Polarization: Index of ethnolinguistic polarization that achieves a maximum score when a country is occupied by two groups of the same population. *Source: Desmet, Ortuño-Ortín, and Wacziarg (2012), following Montalvo and Reynal-Querol (2005a).*

Landlocked: Indicator for countries without access to the sea coast. These countries are: Botswana, Burkina Faso, Burundi, Central African Republic, Chad, Democratic Republic of Congo, Lesotho, Mali, Malawi, Niger, Rwanda, Swaziland, Uganda, Zambia, Zimbabwe. *Source: Global Development Network Growth Database.*

References

- ALESINA, A., A. DEVLEESCHAUWER, W. EASTERLY, S. KURLAT, AND R. WACZIARG (2003): “Fractionalization,” *Journal of Economic Growth*, 8(2), 155–194.
- ALESINA, A., W. EASTERLY, AND J. MATUSZESKI (2011): “Artificial States,” *Journal of the European Economic Association*, 9(2), 246–277.
- ALESINA, A., P. GIULIANO, AND N. NUNN (2013): “On the Origins of Gender Roles: Women and the Plough,” *Quarterly Journal of Economics*, 128(2), 169–530.
- ALESINA, A., S. MICHALOPOULOS, AND E. PAPAIOANNOU (2015): “Ethnic Inequality,” *Journal of Political Economy*, forthcoming.
- ALESINA, A., AND E. ZHURAVSKAYA (2011): “Segregation and the Quality of Government in a Cross-Section of Countries,” *American Economic Review*, 101(5), 1872–1911.
- ALTONJI, J., T. E. ELDER, AND C. R. TABER (2005): “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools,” *Journal of Political Economy*, 113(1), 151–184.
- BESLEY, T., AND M. REYNAL-QUEROL (2014): “The Legacy of Historical Conflicts. Evidence from Africa,” *American Political Science Review*, 108(2), 319–336.
- BLATTMAN, C., AND E. MIGUEL (2010): “Civil War,” *Journal of Economic Literature*, 48(1), 3–57.
- CAMERON, C. A., J. GELBACH, AND D. MILLER (2011): “Robust Inference with Multi-Way Clustering,” *Journal of Business and Economic Statistics*, 28(2), 238–249.
- CHEN, X., AND W. D. NORDHAUS (2011): “Using Luminosity Data as a Proxy for Economic Statistics,” *Proceedings of the National Academy of Sciences*, 108(21), 8589–8594.
- COLLIER, P., AND A. HOEFFLER (2007): “Civil War,” in *Handbook of Defense Economics, Volume 2, Defense in a Globalized World*, ed. by T. Sandler, and K. Hartley, pp. 711–740. Elsevier, North Holland, Amsterdam and Oxford.
- DELL, M. (2012): “Path Dependence in Development: Evidence from the Mexican Revolution,” mimeo, Harvard University, Department of Economics.
- DESMET, K., I. ORTUÑO-ORTÍN, AND R. WACZIARG (2012): “The Political Economy of Ethnolinguistic Cleavages,” *Journal of Development Economics*, 97(2), 322–338.

- ESTEBAN, J., AND D. RAY (2008): “On the Saliency of Ethnic Conflict,” *American Economic Review*, 98(6), 2185–2202.
- (2011): “A Model of Ethnic Conflict,” *Journal of the European Economic Association*, 9(3), 496–521.
- FEARON, J. D. (2003): “Ethnic Structure and Cultural Diversity by Country,” *Journal of Economic Growth*, 8, 195–222.
- FENSKE, J. (2014a): “African polygamy: Past and present,” working paper, Oxford University.
- (2014b): “Ecology, Trade, and States in Pre-colonial Africa,” *Journal of the European Economic Association*, 12(3), 612–640.
- GENNAIOLI, N., AND I. RAINER (2007): “The Modern Impact of Precolonial Centralization in Africa,” *Journal of Economic Growth*, 12(3), 185–234.
- GIULIANO, P., AND N. NUNN (2013): “The Transmission of Democracy: From the Village to the Nation State,” *American Economic Review, Papers and Proceedings*, 103(3), 86–92.
- HARARI, M., AND E. LA FERRARA (2014): “Conflict, Climate and Cells: A Disaggregated Analysis,” mimeo MIT and Bocconi University.
- HAUSMAN, J., B. H. HALL, AND Z. GRILICHES (1984): “Econometric Models for Count Data with an Application to the Patents-RD Relationship,” *Econometrica*, 52(4), 909–938.
- HUBER, J. D., AND L. MAYORAL (2014): “Inequality, Ethnicity and Civil Conflict,” Working Paper, Department of Political Science, Columbia University.
- LEMARCHAND, R. (2008): *The Dynamics of Violence in Central Africa (National and Ethnic Conflict in the 21st Century)*. University of Pennsylvania Press, Philadelphia, PA.
- MICHALOPOULOS, S., AND E. PAPAIOANNOU (2013): “Pre-colonial Ethnic Institutions and Contemporary African Development,” *Econometrica*, 81(1), 113–152.
- (2014): “National Institutions and Subnational Development in Africa,” *Quarterly Journal of Economics*, 129(1), 151–213.
- MICHALOPOULOS, S., L. PUTTERMAN, AND D. N. WEIL (2015): “The Influence of Ancestral Lifeways on Individual Economic Outcomes in Sub-Saharan Africa,” mimeo, Brown University, Department of Economics.

- MIGUEL, E., S. SATYANATH, AND E. SERGENTI (2004): “Economic Shocks and Civil Conflict: An Instrumental Variables Approach,” *Journal of Political Economy*, 112(4), 725–753.
- MONTALVO, J.-G., AND M. REYNAL-QUEROL (2005a): “Ethnic Diversity and Economic Development,” *Journal of Development Economics*, 76(2), 293–323.
- (2005b): “Ethnic Polarization, Potential Conflict, and Civil Wars,” *American Economic Review*, 95(3), 796–816.
- MURDOCK, G. P. (1959): *Africa: Its Peoples and their Culture History*. McGraw-Hill.
- (1967): *Ethnographic Atlas: A Summary*. University of Pittsburgh Press, PA.
- NUNN, N. (2008): “The Long Term Effects of Africa’s Slave Trades,” *Quarterly Journal of Economics*, 123(1), 139–176.
- NUNN, N., AND D. PUGA (2012): “Ruggedness: The Blessing of Bad Geography in Africa,” *Review of Economics and Statistics*, 94(1), 20–36.
- NUNN, N., AND L. WANTCHEKON (2011): “The Slave Trade and the Origins of Mistrust in Africa,” *American Economic Review*, 101(7), 3221–3252.
- PRUNIER, G. (2009): *Africa’s World War. Congo, the Rwandan Genocide, and the Making of a Continental Catastrophe*. Oxford University Press, Oxford, United Kingdom.
- SPOLAORE, E., AND R. WACZIARG (2014): “War and Relatedness,” NBER Working Paper 15095.
- SUNDBERG, R., M. LINDGREN, AND A. PADSKOCIMAITE (2010): “UCDP GED Codebook version 1.0-2011,” Department of Peace and Conflict Research, Uppsala University.
- SUNDBERG, R., AND E. MELANDER (2013): “Introducing the UCDP Georeferenced Event Dataset,” *Journal of Peace Research*, 50(4), 523–532.
- WANTCHEKON, L., M. KLASNJA, AND N. NOVTA (2015): “Education and Human Capital Externalities: Evidence from Colonial Benin,” *Quarterly Journal of Economics*, 130(2).
- WIMMER, A., L.-E. CEDERMAN, AND B. MIN (2009): “Ethnic Politics and Armed Conflict. A Configurational Analysis of a New Global Dataset,” *American Sociological Review*, 74(2), 316–337.

Appendix Table 1: Summary Statistics at the Ethnic Homeland Level

| | Obs. | mean | st. dev. | median | min | max |
|--|------|----------|-----------|---------|-------|-------------|
| Panel A: All Ethnic Homelands | | | | | | |
| Population around Independence | 825 | 329,432 | 1,070,569 | 118,424 | 58 | 25,800,000 |
| Land Area | 825 | 34.213 | 59.204 | 14.500 | 0.235 | 604.903 |
| Lake Indicator | 825 | 0.135 | 0.341 | 0.000 | 0.000 | 1.000 |
| River Indicator | 825 | 0.552 | 0.498 | 1.000 | 0.000 | 1.000 |
| Mean Elevation | 825 | 0.621 | 0.436 | 0.490 | 0.000 | 2.170 |
| Land Suitability for Agriculture | 825 | 0.411 | 0.240 | 0.418 | 0.001 | 0.979 |
| Malaria Stability Index | 825 | 0.752 | 0.360 | 0.976 | 0.000 | 1.000 |
| Distance to the Sea Coast | 825 | 598.2 | 431.9 | 556.3 | 0.2 | 1,721.3 |
| Diamond Mine Indicator | 825 | 0.125 | 0.331 | 0.000 | 0.000 | 1.000 |
| Oil Indicator | 825 | 0.126 | 0.399 | 0.000 | 0.000 | 4.000 |
| Precolonial Conflict Indicator | 825 | 0.048 | 0.215 | 0.000 | 0.000 | 1.000 |
| Distance to Precolonial Conflict | 825 | 0.403 | 0.344 | 0.314 | 0.000 | 2.241 |
| Slave Trades Indicator | 825 | 0.361 | 0.481 | 0.000 | 0.000 | 1.000 |
| Slave Trade Impact | 825 | 13,428.4 | 143,720.1 | 0.0 | 0.0 | 3,838,953.0 |
| Precolonial Kingdom-Empire Indicator | 825 | 0.377 | 0.485 | 0.000 | 0.000 | 1.000 |
| Distance to Precolonial Kingdom-Empire | 825 | 0.173 | 0.226 | 0.073 | 0.000 | 1.236 |
| Major City in 1400 Indicator | 825 | 0.038 | 0.190 | 0.000 | 0.000 | 1.000 |
| Distance to Explorers Routes | 825 | 204.8 | 223.0 | 126.7 | 0.2 | 1,280.5 |
| Number of Distinct Ethnic Families of Adjacent Groups | 825 | 2.804 | 1.266 | 3.000 | 1.000 | 11.000 |
| Share of Adjacent Groups in the Same Ethnic Family | 825 | 0.454 | 0.294 | 0.429 | 0.000 | 1.000 |
| Panel B: Homelands close to the National Border | | | | | | |
| Population around Independence | 413 | 248,522 | 504,731 | 100,027 | 141 | 7,019,231 |
| Land Area | 413 | 26.881 | 50.342 | 11.616 | 0.235 | 565.597 |
| Lake Indicator | 413 | 0.162 | 0.369 | 0.000 | 0.000 | 1.000 |
| River Indicator | 413 | 0.545 | 0.499 | 1.000 | 0.000 | 1.000 |
| Mean Elevation | 413 | 0.608 | 0.425 | 0.475 | 0.000 | 1.813 |
| Land Suitability for Agriculture | 413 | 0.430 | 0.220 | 0.452 | 0.001 | 0.970 |
| Malaria Stability Index | 413 | 0.805 | 0.310 | 0.987 | 0.000 | 1.000 |
| Distance to the Sea Coast | 413 | 575.3 | 401.8 | 566.3 | 0.2 | 1,721.3 |
| Diamond Mine Indicator | 413 | 0.126 | 0.332 | 0.000 | 0.000 | 1.000 |
| Oil Indicator | 413 | 0.099 | 0.403 | 0.000 | 0.000 | 4.000 |
| Precolonial Conflict Indicator | 413 | 0.031 | 0.175 | 0.000 | 0.000 | 1.000 |
| Distance to Precolonial Conflict | 413 | 0.376 | 0.308 | 0.277 | 0.000 | 1.966 |
| Slave Trades Indicator | 413 | 0.378 | 0.485 | 0.000 | 0.000 | 1.000 |
| Slave Trade Impact | 413 | 18,790.1 | 197,525.3 | 0.0 | 0.0 | 3,838,953.0 |
| Precolonial Kingdom-Empire Indicator | 413 | 0.412 | 0.493 | 0.000 | 0.000 | 1.000 |
| Distance to Precolonial Kingdom-Empire | 413 | 0.153 | 0.209 | 0.040 | 0.000 | 0.880 |
| Major City in 1400 Indicator | 413 | 0.017 | 0.129 | 0.000 | 0.000 | 1.000 |
| Distance to Explorers Routes | 413 | 201.3 | 231.7 | 125.3 | 0.6 | 1,280.5 |
| Number of Distinct Ethnic Families of Adjacent Groups | 413 | 2.862 | 1.290 | 3.000 | 1.000 | 11.000 |
| Share of Adjacent Groups in the Same Ethnic Family | 413 | 0.450 | 0.295 | 0.429 | 0.000 | 1.000 |

The table gives summary statistics for the main variables across African ethnic homelands (this is the unit of analysis in Section 3 that examines the correlates of ethnic partitioning). Panel A reports summary statistics across all ethnic homelands (N=825). Panel B gives summary statistics across ethnic homelands that are close to the national border (using as a cut-off the median distance from the centroid of each ethnic homeland to the national border; 102 kilometers). The Data Appendix gives detailed variable definitions and data sources.

Appendix Table 2: Summary Statistics
ACLED. Conflict Variables at the Country-Ethnic Homeland Level

| | Obs. | mean | st. dev. | median | min | p99 | max |
|--|------|--------|----------|--------|-----|-----|--------|
| Panel A: All Ethnic Homelands | | | | | | | |
| All Conflict Events | 1212 | 53.25 | 238.78 | 4 | 0 | 102 | 5423 |
| All Conflicts Indicator | 1212 | 0.73 | 0.44 | 1 | 0 | 1 | 1 |
| Duration All Conflicts | 1212 | 4.47 | 5.01 | 3 | 0 | 13 | 17 |
| All Fatalities | 1212 | 317.60 | 3306.86 | 3 | 0 | 435 | 107554 |
| Fatal Conflict Events | 1212 | 15.99 | 88.51 | 1 | 0 | 28 | 2299 |
| Fatal Conflict Indicator | 1212 | 0.61 | 0.49 | 1 | 0 | 1 | 1 |
| Duration Deadly Events | 1212 | 2.70 | 3.82 | 1 | 0 | 8 | 17 |
| Main Conflict Events | 1212 | 39.62 | 202.71 | 3 | 0 | 74 | 5268 |
| Main Conflict Indicator | 1212 | 0.69 | 0.46 | 1 | 0 | 1 | 1 |
| Duration Main Conflicts | 1212 | 3.87 | 4.63 | 2 | 0 | 11 | 17 |
| Battles | 1212 | 19.24 | 115.39 | 1 | 0 | 37 | 3376 |
| Battles Indicator | 1212 | 0.56 | 0.50 | 1 | 0 | 1 | 1 |
| Violence against the Civilians | 1212 | 16.81 | 92.81 | 1 | 0 | 30 | 2221 |
| Violence against the Civilians Indicator | 1212 | 0.60 | 0.49 | 1 | 0 | 1 | 1 |
| Riots and Protests | 1212 | 13.31 | 84.06 | 0 | 0 | 19 | 2360 |
| Riots and Protests Indicator | 1212 | 0.46 | 0.50 | 0 | 0 | 1 | 1 |
| Government Forces Events | 1212 | 22.22 | 110.16 | 2 | 0 | 42 | 2857 |
| Government Forces Indicator | 1212 | 0.63 | 0.48 | 1 | 0 | 1 | 1 |
| Rebels and Militias Events | 1212 | 34.81 | 186.60 | 3 | 0 | 63 | 4957 |
| Rebels and Militias Indicator | 1212 | 0.66 | 0.47 | 1 | 0 | 1 | 1 |
| Riots and Protests | 1212 | 13.59 | 86.36 | 0 | 0 | 19 | 2440 |
| Riots and Protests Indicator | 1212 | 0.46 | 0.50 | 0 | 0 | 1 | 1 |
| Violence against the Civilians | 1212 | 17.07 | 94.11 | 1 | 0 | 31 | 2266 |
| Violence against the Civilians Indicator | 1212 | 0.60 | 0.49 | 1 | 0 | 1 | 1 |
| Interventions (nearby countries) Events | 1212 | 3.21 | 20.98 | 0 | 0 | 5 | 583 |
| Interventions (nearby countries) Indicator | 1212 | 0.26 | 0.44 | 0 | 0 | 1 | 1 |
| Outside External Interventions Events | 1212 | 1.64 | 7.75 | 0 | 0 | 3 | 127 |
| Outside External Interventions Indicator | 1212 | 0.22 | 0.42 | 0 | 0 | 1 | 1 |

Appendix Table 2: Summary Statistics (cont.)
ACLED. Conflict Variables at the Country-Ethnic Homeland Level

| | Obs. | mean | st. dev. | median | min | p99 | max |
|--|------|--------|----------|--------|-----|------|-------|
| Panel B: Homelands close to the National Border | | | | | | | |
| All Conflict Events | 606 | 23.78 | 123.71 | 2 | 0 | 385 | 2701 |
| All Conflicts Indicator | 606 | 0.62 | 0.49 | 1 | 0 | 1 | 1 |
| Duration All Conflicts | 606 | 3.15 | 4.31 | 1 | 0 | 17 | 17 |
| All Fatalities | 606 | 145.82 | 1062.31 | 0 | 0 | 1531 | 18641 |
| Fatal Conflict Events | 606 | 8.37 | 66.23 | 0 | 0 | 127 | 1558 |
| Fatal Conflict Indicator | 606 | 0.50 | 0.50 | 0 | 0 | 1 | 1 |
| Duration Deadly Events | 606 | 1.77 | 3.02 | 0 | 0 | 15 | 17 |
| Main Conflict Events | 606 | 20.48 | 119.45 | 1 | 0 | 296 | 2654 |
| Main Conflict Indicator | 606 | 0.57 | 0.50 | 1 | 0 | 1 | 1 |
| Duration Main Conflicts | 606 | 2.69 | 3.94 | 1 | 0 | 17 | 17 |
| Battles | 606 | 9.91 | 61.59 | 0 | 0 | 140 | 1375 |
| Battles Indicator | 606 | 0.44 | 0.50 | 0 | 0 | 1 | 1 |
| Violence against the Civilians | 606 | 8.74 | 54.30 | 0 | 0 | 118 | 1196 |
| Violence against the Civilians Indicator | 606 | 0.48 | 0.50 | 0 | 0 | 1 | 1 |
| Riots and Protests | 606 | 3.22 | 12.05 | 0 | 0 | 65 | 134 |
| Riots and Protests Indicator | 606 | 0.33 | 0.47 | 0 | 0 | 1 | 1 |
| Government Forces Events | 606 | 10.25 | 59.73 | 1 | 0 | 139 | 1347 |
| Government Forces Indicator | 606 | 0.50 | 0.50 | 1 | 0 | 1 | 1 |
| Rebels and Militias Events | 606 | 17.95 | 116.10 | 1 | 0 | 319 | 2630 |
| Rebels and Militias Indicator | 606 | 0.53 | 0.50 | 1 | 0 | 1 | 1 |
| Riots and Protests | 606 | 3.31 | 12.36 | 0 | 0 | 65 | 139 |
| Riots and Protests Indicator | 606 | 0.33 | 0.47 | 0 | 0 | 1 | 1 |
| Violence against the Civilians | 606 | 8.81 | 54.40 | 0 | 0 | 120 | 1196 |
| Violence against the Civilians Indicator | 606 | 0.48 | 0.50 | 0 | 0 | 1 | 1 |
| Interventions (nearby countries) Events | 606 | 1.93 | 8.81 | 0 | 0 | 39 | 115 |
| Interventions (nearby countries) Indicator | 606 | 0.21 | 0.41 | 0 | 0 | 1 | 1 |
| Outside External Interventions Events | 606 | 1.33 | 6.35 | 0 | 0 | 25 | 92 |
| Outside External Interventions Indicator | 606 | 0.19 | 0.39 | 0 | 0 | 1 | 1 |

The table reports summary statistics for the main conflict variables from the ACLED employed in the empirical analysis (in Section 4-5). Panel A reports summary statistics across all country-ethnic homelands (1212 observations). Panel B reports summary statistics for country-ethnicity homelands close to the national border using as a cut-off the median distance from the centroid of each ethnic homeland to the national border (606 observations). The Data Appendix gives detailed variable definitions and data sources.

Appendix Table 3: Total and Fatal Civil Conflict Events by Year (ACLED)

| Year | All Events | | Fatal Events | | Percentage Fatal |
|-------|------------|------------|--------------|------------|------------------|
| | Number | Percentage | Number | Percentage | |
| 1997 | 1,992 | 3.08 | 810 | 4.16 | 40.7% |
| 1998 | 2,490 | 3.85 | 848 | 4.36 | 34.1% |
| 1999 | 3,572 | 5.52 | 1,000 | 5.14 | 28.0% |
| 2000 | 3,289 | 5.09 | 1,185 | 6.09 | 36.0% |
| 2001 | 2,811 | 4.35 | 951 | 4.89 | 33.8% |
| 2002 | 3,313 | 5.12 | 987 | 5.07 | 29.8% |
| 2003 | 2,706 | 4.18 | 1,003 | 5.16 | 37.1% |
| 2004 | 2,348 | 3.63 | 797 | 4.1 | 33.9% |
| 2005 | 1,958 | 3.03 | 528 | 2.71 | 27.0% |
| 2006 | 1,952 | 3.02 | 455 | 2.34 | 23.3% |
| 2007 | 2,212 | 3.42 | 559 | 2.87 | 25.3% |
| 2008 | 3,127 | 4.84 | 674 | 3.46 | 21.6% |
| 2009 | 2,806 | 4.34 | 873 | 4.49 | 31.1% |
| 2010 | 3,509 | 5.43 | 1,461 | 7.51 | 41.6% |
| 2011 | 5,261 | 8.14 | 1,460 | 7.5 | 27.8% |
| 2012 | 8,753 | 13.54 | 2,350 | 12.08 | 26.8% |
| 2013 | 12,565 | 19.43 | 3,514 | 18.06 | 28.0% |
| Total | 64,664 | 100 | 19,455 | 100 | 30.1% |

The table gives the distribution (number and share) of all conflict incidents and deadly conflict incidents for each year for the ACLED database (vintage 4). The Data Appendix gives detailed variable definitions and data sources.

Appendix Table 4: Total and Fatal Conflict Incidents by ACLED Category

| Conflict Category | All Events | | Fatal Events | | Percentage Fatal |
|--|------------|------------|--------------|------------|------------------|
| | Number | Percentage | Number | Percentage | |
| Battle-Government regains territory | 1,205 | 1.86 | 224 | 1.15 | 18.59% |
| Battle-No change of territory | 20,893 | 32.31 | 9,117 | 46.86 | 43.64% |
| Battle-Non-state actor overtakes territory | 1,283 | 1.98 | 255 | 1.31 | 19.88% |
| Headquarters or base established | 271 | 0.42 | 2 | 0.01 | 0.74% |
| Non-violent activity by a conflict actor | 3,913 | 6.05 | 37 | 0.19 | 0.95% |
| Non-violent transfer of territory | 543 | 0.84 | 3 | 0.02 | 0.55% |
| Riots/Protests | 16,147 | 24.97 | 1,040 | 5.35 | 6.44% |
| Violence against civilians | 20,409 | 31.56 | 8,777 | 45.11 | 43.01% |
| Total | 64,664 | 100 | 19,455 | 100 | 30.1% |

The table gives the distribution (number and percentage share) of all conflict incidents and deadly conflict incidents for each conflict category for the ACLED database (vintage 4). The Data Appendix gives detailed variable definitions and data sources.

Appendix Table 5: Conflict Events by Actors (ACLED). 1997-2013

Panel A: All Conflict Events

| Conflict Actor | Id Number | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total |
|------------------------|-----------|---------------|---------------|--------------|--------------|--------------|--------------|------------|---------------|--------------|---------------|
| Government Force | 1 | 829 | 764 | 6,891 | 3,043 | 229 | 546 | 435 | 3,743 | 198 | 16,678 |
| Rebel Force | 2 | 1,298 | 3,007 | 397 | 642 | 71 | 8 | 3 | 3,354 | 853 | 9,633 |
| Political Militia | 3 | 938 | 2,733 | 478 | 1,036 | 111 | 14 | 28 | 8,886 | 329 | 14,553 |
| Ethnic Militia | 4 | 70 | 235 | 61 | 149 | 1,422 | 1 | 1 | 1,189 | 24 | 3,152 |
| Rioters | 5 | 2,196 | 2,110 | 14 | 66 | 13 | 440 | 14 | 213 | 58 | 5,124 |
| Protesters | 6 | 8,845 | 1,272 | 14 | 29 | 5 | 26 | 44 | 0 | 30 | 10,265 |
| Civilians | 7 | 1 | 569 | 868 | 1,215 | 125 | 52 | 0 | 0 | 43 | 2,873 |
| Outside/external Force | 8 | 284 | 397 | 865 | 288 | 21 | 9 | 3 | 455 | 50 | 2,372 |
| Total | | 14,461 | 11,087 | 9,588 | 6,468 | 1,997 | 1,096 | 528 | 17,840 | 1,585 | 64,650 |

Panel B: Deadly Conflict Events

| Conflict Actor | Id | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total |
|------------------------|----|------------|--------------|--------------|--------------|--------------|------------|-----------|--------------|------------|---------------|
| Government Force | 1 | 7 | 279 | 2,426 | 1,277 | 121 | 84 | 37 | 1,229 | 50 | 5,510 |
| Rebel Force | 2 | 14 | 1,326 | 159 | 236 | 25 | 7 | 1 | 1,879 | 254 | 3,901 |
| Political Militia | 3 | 14 | 1,243 | 160 | 373 | 57 | 6 | 3 | 3,230 | 106 | 5,192 |
| Ethnic Militia | 4 | 0 | 113 | 21 | 84 | 796 | 0 | 1 | 699 | 16 | 1,730 |
| Rioters | 5 | 205 | 432 | 7 | 14 | 3 | 83 | 1 | 43 | 5 | 793 |
| Protesters | 6 | 87 | 81 | 0 | 4 | 1 | 1 | 1 | 0 | 3 | 178 |
| Civilians | 7 | 0 | 240 | 530 | 545 | 82 | 13 | 0 | 0 | 26 | 1,436 |
| Outside/external Force | 8 | 0 | 76 | 278 | 96 | 7 | 5 | 0 | 227 | 19 | 708 |
| Total | | 327 | 3,790 | 3,581 | 2,629 | 1,092 | 199 | 44 | 7,307 | 479 | 19,448 |

The table gives the distribution (number) of all conflict incidents (in Panel A) and deadly conflict incidents (in Panel B) by ACLED conflict actors. There are 8 actor categories (1: Government Forces; 2: Rebel Forces; 3: Political Militia; 4: Ethnic Militia; 5: Rioters; 6: Protesters; 7: Civilians; and 8: Outside/external Force. 0 indicates unassigned conflict actor). The Data Appendix gives detailed variable definitions and data sources.

Appendix Table 6: Civil Conflict by Country (ACLED and UCDP GED)

| Country Name | Murdock (1959) | | ACLED 4 | | | | | | | | UCDP GED 1.5 | | |
|-------------------|------------------|-----------------|------------------------|----------------------|-------------------|--------------------|-------------------|--------------|----------|-----------|----------------|--------------------|--------------------|
| | Ethnic Homelands | Split Homelands | All Conflict Incidents | All Types of Battles | Civilian Violence | Riots and Protests | Government Forces | Rebel Forces | Militias | Civilians | State Conflict | One-Sided Violence | Non-State Conflict |
| Angola | 29 | 13 | 2443 | 1848 | 337 | 113 | 1947 | 2166 | 70 | 339 | 782 | 156 | 0 |
| Burundi | 3 | 3 | 2824 | 1433 | 1250 | 49 | 1404 | 2303 | 447 | 1250 | 481 | 333 | 19 |
| Benin | 16 | 12 | 64 | 1 | 6 | 55 | 21 | 0 | 4 | 6 | 0 | 10 | 0 |
| Botswana | 30 | 14 | 204 | 18 | 23 | 156 | 85 | 1 | 23 | 23 | 0 | 0 | 0 |
| Burkina Faso | 17 | 7 | 47 | 3 | 7 | 35 | 8 | 2 | 9 | 7 | 0 | 0 | 0 |
| Central African | 26 | 18 | 1166 | 444 | 485 | 120 | 395 | 446 | 417 | 486 | 53 | 89 | 0 |
| Cote d'Ivoire | 33 | 13 | 1305 | 452 | 406 | 413 | 604 | 150 | 482 | 408 | 50 | 66 | 25 |
| Cameroon | 65 | 28 | 187 | 61 | 55 | 62 | 87 | 11 | 89 | 56 | 1 | 19 | 7 |
| Congo, Rep. | 16 | 11 | 285 | 173 | 66 | 21 | 190 | 1 | 231 | 68 | 104 | 86 | 0 |
| Djibuti | 2 | 2 | 54 | 17 | 11 | 20 | 34 | 6 | 4 | 11 | 21 | 0 | 0 |
| Algeria | 26 | 8 | 2057 | 958 | 518 | 482 | 1182 | 1052 | 482 | 524 | 1854 | 192 | 6 |
| Egypt, Arab Rep. | 10 | 3 | 3838 | 449 | 414 | 2628 | 1449 | 33 | 816 | 427 | 205 | 99 | 0 |
| Eritrea | 9 | 4 | 256 | 97 | 126 | 7 | 178 | 12 | 21 | 126 | 73 | 15 | 0 |
| Western Sahara | 5 | 5 | 80 | 2 | 20 | 55 | 51 | 0 | 5 | 16 | 4 | 0 | 0 |
| Ethiopia | 48 | 13 | 1186 | 703 | 239 | 181 | 855 | 638 | 179 | 248 | 590 | 105 | 156 |
| Gabon | 13 | 8 | 79 | 3 | 13 | 58 | 32 | 3 | 12 | 13 | 0 | 0 | 0 |
| Ghana | 31 | 17 | 205 | 61 | 70 | 67 | 53 | 2 | 118 | 71 | 0 | 0 | 31 |
| Guinea | 19 | 14 | 591 | 166 | 182 | 209 | 330 | 113 | 162 | 181 | 20 | 35 | 1 |
| Gambia, The | 3 | 2 | 81 | 9 | 39 | 27 | 20 | 7 | 32 | 40 | 1 | 1 | 0 |
| Guinea-Bissau | 9 | 6 | 190 | 102 | 23 | 48 | 129 | 18 | 18 | 26 | 18 | 0 | 0 |
| Equatorial Guine: | 5 | 4 | 31 | 7 | 13 | 6 | 24 | 5 | 5 | 15 | 0 | 0 | 0 |
| Kenya | 36 | 15 | 3095 | 753 | 1042 | 1148 | 1026 | 194 | 1577 | 1056 | 2 | 87 | 152 |
| Liberia | 14 | 8 | 917 | 580 | 163 | 121 | 609 | 659 | 91 | 162 | 108 | 231 | 46 |
| Libya | 12 | 3 | 1535 | 777 | 287 | 289 | 880 | 288 | 609 | 298 | 0 | 0 | 0 |
| Lesotho | 3 | 1 | 90 | 26 | 30 | 32 | 38 | 1 | 35 | 30 | 3 | 0 | 0 |
| Morocco | 20 | 6 | 369 | 3 | 43 | 290 | 143 | 4 | 28 | 39 | 3 | 2 | 0 |

| | | | | | | | | | | | | | |
|-----------------|-----|----|------|------|------|------|------|------|------|------|------|-----|------|
| Madagascar | 11 | 0 | 484 | 45 | 119 | 305 | 164 | 0 | 137 | 120 | 0 | 1 | 32 |
| Mali | 24 | 13 | 614 | 222 | 173 | 119 | 204 | 327 | 87 | 181 | 21 | 37 | 3 |
| Mozambique | 21 | 15 | 369 | 38 | 144 | 163 | 115 | 0 | 176 | 150 | 94 | 67 | 2 |
| Mauritania | 11 | 7 | 237 | 19 | 23 | 180 | 118 | 17 | 12 | 21 | 3 | 8 | 5 |
| Malawi | 12 | 11 | 179 | 5 | 68 | 97 | 84 | 0 | 52 | 71 | 0 | 0 | 0 |
| Namibia | 14 | 9 | 419 | 47 | 85 | 284 | 70 | 70 | 34 | 85 | 6 | 7 | 0 |
| Niger | 23 | 12 | 311 | 137 | 61 | 106 | 190 | 78 | 52 | 63 | 43 | 13 | 1 |
| Nigeria | 112 | 23 | 4309 | 1471 | 1641 | 1052 | 1412 | 133 | 2833 | 1646 | 31 | 88 | 186 |
| Rwanda | 5 | 4 | 529 | 143 | 324 | 25 | 276 | 173 | 200 | 327 | 93 | 140 | 0 |
| Sudan | 83 | 23 | 3590 | 1411 | 1431 | 483 | 1702 | 1025 | 1289 | 1444 | 480 | 467 | 148 |
| Senegal | 12 | 9 | 565 | 207 | 138 | 200 | 241 | 233 | 107 | 142 | 91 | 94 | 11 |
| Sierra Leone | 13 | 7 | 1250 | 797 | 266 | 74 | 318 | 1032 | 222 | 275 | 497 | 766 | 11 |
| Somalia | 12 | 6 | 9559 | 5309 | 2761 | 574 | 3830 | 3150 | 5558 | 2807 | 1077 | 141 | 505 |
| Swaziland | 2 | 2 | 147 | 1 | 36 | 90 | 78 | 0 | 29 | 43 | 0 | 2 | 0 |
| Chad | 45 | 19 | 446 | 252 | 161 | 17 | 274 | 155 | 196 | 165 | 91 | 73 | 8 |
| Togo | 24 | 17 | 182 | 10 | 24 | 143 | 80 | 0 | 17 | 22 | 0 | 89 | 1 |
| Tunisia | 12 | 5 | 1025 | 89 | 82 | 765 | 313 | 5 | 178 | 85 | 0 | 1 | 0 |
| Tanzania | 69 | 14 | 452 | 46 | 204 | 182 | 149 | 18 | 212 | 206 | 0 | 8 | 1 |
| Uganda | 27 | 13 | 1919 | 657 | 660 | 416 | 907 | 1032 | 330 | 674 | 310 | 220 | 34 |
| South Africa | 28 | 11 | 3342 | 85 | 574 | 2637 | 769 | 2 | 589 | 613 | 5 | 531 | 2125 |
| Congo, Dem. Rep | 104 | 30 | 5872 | 3098 | 1614 | 399 | 2470 | 3183 | 1903 | 1618 | 297 | 898 | 129 |
| Zambia | 34 | 20 | 803 | 20 | 217 | 537 | 157 | 9 | 158 | 229 | 0 | 5 | 0 |
| Zimbabwe | 14 | 10 | 4759 | 59 | 3701 | 597 | 1239 | 0 | 3101 | 3780 | 0 | 37 | 0 |

| | | | | | | | | | | | | | |
|-------|------|-----|-------|-------|-------|-------|-------|-------|-------|-------|------|------|------|
| Total | 1212 | 518 | 64541 | 23314 | 20375 | 16137 | 26934 | 18757 | 23438 | 20693 | 7512 | 5219 | 3645 |
|-------|------|-----|-------|-------|-------|-------|-------|-------|-------|-------|------|------|------|

The table gives the number of ethnic homelands, partitioned ethnic homelands, and conflict incidents for each country with the ACLED (v4) and UCDP GED (v1.5) database. ACLED covers the period 1997-2013. UCDP GED covers the period 1989-2010. The Data Appendix gives detailed variable definitions and data sources.

Appendix Table 7: Correlation Structure - Main Conflict Variables (ACLED and UCDP)

| | ACLED | | | | | | | | UCDP | | |
|------------------------|------------|---------------|------------|---------------------|-----------------|---------|-------------------|------------------|---------|--------------------|-----------|
| | All Events | Deadly Events | Fatalities | Duration All Events | Duration Deadly | Battles | Civilian Violence | Riots & Protests | State | One-Sided Violence | Non-State |
| ACLED | | | | | | | | | | | |
| All Conflict Events | 1 | | | | | | | | | | |
| Deadly Events | 0.8727* | 1 | | | | | | | | | |
| Fatalities | 0.2285* | 0.2551* | 1 | | | | | | | | |
| Duration All Events | 0.4274* | 0.3413* | 0.1392* | 1 | | | | | | | |
| Duration Deadly Events | 0.4725* | 0.4096* | 0.1778* | 0.9053* | 1 | | | | | | |
| Battles | 0.8325* | 0.9496* | 0.3050* | 0.3003* | 0.3455* | 1 | | | | | |
| Civilian Violence | 0.8585* | 0.7120* | 0.1493* | 0.3543* | 0.4087* | 0.6443* | 1 | | | | |
| Riots and Protests | 0.5678* | 0.2417* | 0.0324 | 0.3230* | 0.3259* | 0.1309* | 0.2988* | 1 | | | |
| UCDP | | | | | | | | | | | |
| State Conflict | 0.6318* | 0.7494* | 0.2672* | 0.2373* | 0.2867* | 0.7191* | 0.4695* | 0.1888* | 1 | | |
| One-Sided Violence | 0.5243* | 0.5377* | 0.2213* | 0.3900* | 0.4299* | 0.4740* | 0.4304* | 0.2765* | 0.4978* | 1 | |
| Non-State Conflict | 0.1896* | 0.1701* | 0.0152 | 0.1361* | 0.1717* | 0.1527* | 0.1318* | 0.1587* | 0.0888* | 0.3024* | 1 |

The table gives the correlation structure of the main civil conflict variables across all country-ethnic homelands (N=1212). * indicate statistical significance at the 5% level.

**Appendix Table 8: Border Artificiality
Pre-colonial Ethnic Features (using data from Murdock (1967)) and Ethnic Partitioning**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|-------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Log Land Area | 0.0748*** (0.0171) | 0.0777*** (0.0207) | 0.0725*** (0.0205) | 0.0725*** (0.0178) | 0.0679*** (0.0195) | 0.0790*** (0.0164) | 0.0818*** (0.0222) | 0.0739*** (0.0176) |
| Lake Indicator | 0.0866 (0.0678) | 0.1015 (0.0679) | 0.1057 (0.0665) | 0.0808 (0.0799) | 0.0835 (0.0771) | 0.0888 (0.0685) | 0.1176 (0.0785) | 0.0997 (0.0657) |
| River Indicator | -0.0169 (0.0430) | -0.0248 (0.0432) | -0.0172 (0.0429) | -0.0053 (0.0419) | -0.0259 (0.0472) | -0.0238 (0.0428) | -0.0003 (0.0426) | -0.0224 (0.0417) |
| Complex Settlement Patterns | -0.0004 (0.0538) | | | | | | | |
| Dependence on Agriculture | | 0.0015 (0.0112) | | | | | | |
| Animal Husbandry | | | (0.0088) (0.0108) | | | | | |
| Local Elections | | | | -0.0459 (0.0848) | | | | |
| Inheritance Rule for Property | | | | | 0.0040 (0.0891) | | | |
| Political Centralization | | | | | | -0.0719 (0.0482) | | |
| Class Stratification | | | | | | | -0.0674 (0.0444) | |
| Polygyny | | | | | | | | -0.0121 (0.0498) |
| adjusted R-squared | 0.081 | 0.081 | 0.082 | 0.095 | 0.0790 | 0.0830 | 0.098 | 0.077 |
| Region Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 451 | 451 | 437 | 437 | 394 | 394 | 487 | 487 |

The table reports linear probability model (LPM) estimates associating ethnic partitioning (SPLIT) with various variables reflecting pre-colonial economic, social and political development (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 (using the 2000 Digital Chart of the World). All specifications include a set of (five) region fixed effects (constants not reported). Standard errors in parentheses are adjusted for double clustering at the country-dimension and the ethnolinguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 9: "Balancedness Tests." Ethnic Partitioning and Geographic Characteristics within Countries (cont.)

Panel B: Country-Ethnic Homelands near the National Border

| | Dependent variable is: | | | | | | | | | | | | | |
|--------------|------------------------|---------------------|---------------------|--------------------|--------------------|---------------------|--------------------|---------------------|--------------------|--------------------|--------------------|---------------------|---------------------|---------------------|
| | Log Land Area | Lake Indicator | River Indicator | Mean Elevation | Land Suitability | Malaria Stability | Diamond Indicator | Petroleum Indicator | Major City in 1400 | Coastal Indicator | Capital Indicator | Distance Sea | Distance Border | Distance Capital |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| SPLIT | 0.7236*** (0.1762) | -0.0191 (0.0294) | 0.0864* (0.0399) | 0.0010 (0.0326) | 0.0243 (0.0154) | -0.0090 (0.0214) | 0.0123 (0.0250) | 0.0095 (0.0212) | 0.0057 (0.0045) | 0.0475 (0.0311) | 0.0118 (0.0094) | -0.0268 (0.0303) | -1.9752 (2.8585) | -0.0128 (0.0370) |
| marginal R2 | 0.052 | 0.001 | 0.005 | 0.000 | 0.002 | 0.000 | 0.001 | 0.001 | 0.002 | 0.005 | 0.001 | 0.001 | 0.002 | 0.000 |
| Mean DV | 1.2276 | 0.0941 | 0.4620 | 0.5928 | 0.4314 | 0.7714 | 0.0594 | 0.0413 | 0.0033 | 0.1089 | 0.0215 | 0.5855 | 24.8881 | 0.5163 |
| Observations | 606 | 606 | 606 | 606 | 606 | 606 | 606 | 606 | 606 | 606 | 606 | 606 | 606 | 606 |
| Country FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

The table reports OLS estimates associating various geographical, ecological, and natural resource characteristics with ethnic partitioning within countries. The unit of analysis is an ethnic territory in a country (ethnicity-country). Panel A reports estimates in the full sample of (country-ethnicity) homelands (1212 observations). Panel B gives estimates in the sample of country-ethnic homelands that are close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. All specifications include country fixed effects (constants not reported). The dependent variable in column (1) is the log of a country-ethnicity's region surface area; in column (2) is an indicator for regions with a lake; in column (3) is an indicator for regions with a river; in column (4) is average elevation; in column (5) is an index capturing land's (soil) suitability (quality) for agriculture; in column (6) is the average value of a malaria stability index; in column (7) is a binary index that takes on the value of one if a diamond mine is present; in column (8) is a binary index that takes on the value of one if an oil/petroleum field is present; in column (9) a binary index that takes on the value of one if a major city was present before European's arrival in Africa (in 1400); in column (10) is an indicator for ethnic homelands that are adjacent to the sea; in column (11) is an indicator for homelands where capital cities fall; in columns (12), (13), and (14) is the distance from the centroid of each country-ethnic areas to the closest sea-coast, the national border, and the capital, respectively. The Data Appendix gives detailed variable definitions and data sources. The table reports (in parentheses) double-clustered standard errors at the country and the ethno-linguistic family dimensions. The table reports the marginal (partial) R-square that equals the difference of the R-square of the model with the ethnic partitioning index (SPLIT) and the model without it (just country constants). The table reports the mean value of the dependent variable. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 10: Test of Means and Medians for Main Civil Conflict Measures
Panel A: All Ethnic Homelands

| | Indicator | Number of Incidents | | Excluding | | Excluding | |
|---|---------------|---------------------|---------------|---------------|---------------|-------------------|---------------|
| | Likelihood | - All | | Capitals | | Outliers (top 1%) | |
| | mean (1) | mean (2) | median (3) | mean (4) | median (5) | mean (6) | median (7) |
| <u>All Types of Conflict Events (ACLED)</u> | | | | | | | |
| non-partitioned ethnic groups (N=596) | 0.815 | 47.616 | 3.000 | 30.495 | 3.000 | 30.464 | 3.000 |
| partitioned ethnic groups (N=229) | 0.878 | 65.031 | 13.000 | 47.872 | 10.000 | 52.526 | 13.000 |
| difference | 0.062 | 17.415 | 10.000 | 17.377 | 7.000 | 22.062 | 10.000 |
| difference (p-value) | <i>(0.07)</i> | <i>(0.42)</i> | <i>(0.00)</i> | <i>(0.05)</i> | <i>(0.09)</i> | <i>(0.02)</i> | <i>(0.01)</i> |
| <u>Battles (ACLED)</u> | | | | | | | |
| non-partitioned ethnic groups (N=596) | 0.649 | 19.837 | 1.000 | 14.551 | 1.000 | 12.566 | 1.000 |
| partitioned ethnic groups (N=229) | 0.725 | 30.052 | 3.000 | 22.782 | 2.000 | 21.212 | 2.000 |
| difference | 0.076 | 10.215 | 2.000 | 8.231 | 1.000 | 8.646 | 1.000 |
| difference (p-value) | <i>(0.33)</i> | <i>(0.20)</i> | <i>(0.03)</i> | <i>(0.10)</i> | <i>(0.19)</i> | <i>(0.03)</i> | <i>(0.12)</i> |
| <u>Violence against Civilians (ACLED)</u> | | | | | | | |
| non-partitioned ethnic groups (N=596) | 0.683 | 17.648 | 1.000 | 10.502 | 1.000 | 10.007 | 1.000 |
| partitioned ethnic groups (N=229) | 0.777 | 23.258 | 3.000 | 16.758 | 2.000 | 17.079 | 3.000 |
| difference | 0.094 | 5.610 | 2.000 | 6.257 | 1.000 | 7.073 | 2.000 |
| difference (p-value) | <i>(0.05)</i> | <i>(0.53)</i> | <i>(0.03)</i> | <i>(0.05)</i> | <i>(0.10)</i> | <i>(0.05)</i> | <i>(0.03)</i> |
| <u>Riots and Protests (ACLED)</u> | | | | | | | |
| non-partitioned ethnic groups (N=596) | 0.530 | 7.792 | 0.000 | 3.757 | 0.000 | 4.640 | 0.000 |
| partitioned ethnic groups (N=229) | 0.673 | 6.590 | 1.000 | 4.735 | 1.000 | 5.996 | 1.000 |
| difference | 0.142 | -1.202 | 1.000 | 0.978 | 1.000 | 1.356 | 1.000 |
| difference (p-value) | <i>(0.00)</i> | <i>(0.56)</i> | <i>(0.00)</i> | <i>(0.47)</i> | <i>(0.00)</i> | <i>(0.21)</i> | <i>(0.00)</i> |
| <u>State-driven Conflict (UCDP)</u> | | | | | | | |
| non-partitioned ethnic groups (N=596) | 0.299 | 8.886 | 0.000 | 4.518 | 0.000 | 3.160 | 0.000 |
| partitioned ethnic groups (N=229) | 0.446 | 9.677 | 0.000 | 6.820 | 0.000 | 6.872 | 0.000 |
| difference | 0.147 | 0.791 | 0.000 | 2.302 | 0.000 | 3.713 | 0.000 |
| difference (p-value) | <i>(0.00)</i> | <i>(0.87)</i> | | <i>(0.31)</i> | | <i>(0.01)</i> | |
| <u>One-Sided Violence (UCDP)</u> | | | | | | | |
| non-partitioned ethnic groups (N=596) | 0.334 | 4.745 | 0.000 | 3.366 | 0.000 | 3.447 | 0.000 |
| partitioned ethnic groups (N=229) | 0.48 | 10.44 | 0.00 | 7.64 | 0.00 | 6.44 | 0.00 |
| difference | 0.147 | 5.696 | 0.000 | 4.278 | 0.000 | 2.998 | 0.000 |
| difference (p-value) | <i>(0.00)</i> | <i>(0.14)</i> | | <i>(0.12)</i> | | <i>(0.07)</i> | |
| <u>Non-State-Driven Conflict (UCDP)</u> | | | | | | | |
| non-partitioned ethnic groups (N=596) | 0.201 | 5.341 | 0.000 | 4.368 | 0.000 | 1.396 | 0.000 |
| partitioned ethnic groups (N=229) | 0.205 | 2.018 | 0.000 | 2.038 | 0.000 | 1.147 | 0.000 |
| difference | 0.004 | -3.323 | 0.000 | -2.330 | 0.000 | -0.249 | 0.000 |
| difference (p-value) | <i>(0.93)</i> | <i>(0.37)</i> | | <i>(0.47)</i> | | <i>(0.50)</i> | |

Appendix Table 10: Test of Means and Medians for Main Civil Conflict Measures
Panel B: Ethnic Homelands close to the National Border

| | Indicator | | Excluding | | Excluding | | |
|---|---------------|---------------------|---------------|---------------|---------------|---------------|---------------|
| | Likelihood | Number of Incidents | | Capitals | Outliers | | |
| | mean | mean | median | mean | median | mean | median |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| <u>All Types of Conflict Events (ACLED)</u> | | | | | | | |
| non-partitioned ethnic groups (N=200) | 0.873 | 32.300 | 1.000 | 29.270 | 1.000 | 24.704 | 1.000 |
| partitioned ethnic groups (N=213) | 0.725 | 66.067 | 12.000 | 47.249 | 10.000 | 52.524 | 12.000 |
| difference | 0.148 | 33.667 | 11.000 | 17.978 | 9.000 | 27.820 | 11.000 |
| difference (p-value) | <i>(0.00)</i> | <i>(0.10)</i> | <i>(0.01)</i> | <i>(0.09)</i> | <i>(0.02)</i> | <i>(0.00)</i> | <i>(0.00)</i> |
| <u>Battles (ACLED)</u> | | | | | | | |
| non-partitioned ethnic groups (N=200) | 0.714 | 15.700 | 0.000 | 13.867 | 0.000 | 9.879 | 0.000 |
| partitioned ethnic groups (N=213) | 0.550 | 30.230 | 3.000 | 22.173 | 2.000 | 20.719 | 3.000 |
| difference | 0.164 | 14.530 | 3.000 | 8.305 | 2.000 | 10.840 | 3.000 |
| difference (p-value) | <i>(0.00)</i> | <i>(0.16)</i> | <i>(0.15)</i> | <i>(0.18)</i> | <i>(0.10)</i> | <i>(0.01)</i> | <i>(0.01)</i> |
| <u>Violence against Civilians (ACLED)</u> | | | | | | | |
| non-partitioned ethnic groups (N=200) | 0.615 | 11.390 | 0.000 | 10.786 | 0.000 | 7.525 | 0.000 |
| partitioned ethnic groups (N=213) | 0.765 | 24.033 | 3.000 | 16.919 | 2.000 | 17.393 | 3.000 |
| difference | 0.150 | 12.643 | 3.000 | 6.133 | 2.000 | 9.868 | 3.000 |
| difference (p-value) | <i>(0.00)</i> | <i>(0.13)</i> | <i>(0.01)</i> | <i>(0.08)</i> | <i>(0.00)</i> | <i>(0.00)</i> | <i>(0.00)</i> |
| <u>Riots and Protests (ACLED)</u> | | | | | | | |
| non-partitioned ethnic groups | 0.420 | 3.325 | 0.000 | 4.528 | 0.000 | 3.325 | 0.000 |
| partitioned ethnic groups | 0.667 | 6.394 | 1.000 | 2.924 | 1.000 | 5.755 | 1.000 |
| difference | 0.247 | 3.069 | 1.000 | 1.604 | 1.000 | 2.430 | 1.000 |
| difference (p-value) | <i>(0.00)</i> | <i>(0.01)</i> | <i>(0.01)</i> | <i>(0.23)</i> | <i>(0.01)</i> | <i>(0.02)</i> | <i>(0.01)</i> |
| <u>State-driven Conflict (UCDP)</u> | | | | | | | |
| non-partitioned ethnic groups (N=596) | 0.195 | 3.315 | 0.000 | 2.515 | 0.000 | 2.281 | 0.000 |
| partitioned ethnic groups (N=229) | 0.432 | 9.662 | 0.000 | 6.503 | 0.000 | 6.645 | 0.000 |
| difference | 0.237 | 6.347 | 0.000 | 3.987 | 0.000 | 4.363 | 0.000 |
| difference (p-value) | <i>(0.00)</i> | <i>(0.05)</i> | | <i>(0.06)</i> | | <i>(0.00)</i> | |
| <u>One-Sided Violence (UCDP)</u> | | | | | | | |
| non-partitioned ethnic groups (N=596) | 0.245 | 4.695 | 0.000 | 3.714 | 0.000 | 1.505 | 0.000 |
| partitioned ethnic groups (N=229) | 0.479 | 11.042 | 0.000 | 8.000 | 0.000 | 6.751 | 0.000 |
| difference | 0.234 | 6.347 | 0.000 | 4.286 | 0.000 | 5.246 | 0.000 |
| difference (p-value) | <i>(0.00)</i> | <i>(0.07)</i> | | <i>(0.13)</i> | | <i>(0.00)</i> | |
| <u>Non-State-Driven Conflict (UCDP)</u> | | | | | | | |
| non-partitioned ethnic groups (N=596) | 0.165 | 9.175 | 0.000 | 9.321 | 0.000 | 1.136 | 0.000 |
| partitioned ethnic groups (N=229) | 0.207 | 2.042 | 0.000 | 2.056 | 0.000 | 1.105 | 0.000 |
| difference | 0.042 | -7.133 | 0.000 | -7.266 | 0.000 | -0.030 | 0.000 |
| difference (p-value) | <i>(0.35)</i> | <i>(0.39)</i> | | <i>(0.40)</i> | | <i>(0.95)</i> | |

The table reports summary statistics and test of means and medians for the ACLED and UCDP civil conflict variables at the ethnic homeland level. Panel A reports test of means/medians at the full sample of ethnic homelands. Panel B reports test of means/medians across ethnic homelands close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland to the national border; 102 kilometers). Column (1) reports the likelihood that a type of conflict (all conflict incidents, battles, violence against the civilian population, riots and protests, state-driven conflict, one-sided violence, and non-state-actor driven conflict) affect ethnic homelands. Columns (2)-(3) report the mean and the median value for each type of conflict, respectively. Columns (4)-(5) report the mean and the median value for each type of conflict excluding ethnic homelands where capital cities fall. Columns (6)-(7) report the mean and the median value for each type of conflict, excluding ethnic regions where the respective variable exceeds the 99th percentile (outliers). For each variable the table reports the mean/median value using all ethnic homelands, partitioned ethnicities and non-partitioned ethnicities. The table also reports the mean and median difference and the p-value of mean-median equality between the group of partitioned and non-partitioned ethnicities. The associated p-values for the test of means are based on double-clustered standard errors at the country level and at the ethnolinguistic level. The associated p-values for the test of medians are based on clustered at the country-level standard errors (recovered via median regression). The Data Appendix gives detailed variable definitions and data sources.

Appendix Table 11: Ethnic Partitioning and Civil Conflict. Baseline Country Fixed-Effects Estimates over 1997-2010

| | All Ethnicity-Country Homelands | | | | | | Ethnicity-Country Homelands Close to the National Border | | | | | |
|--|---------------------------------|-----------|-----------|-----------|----------------|----------------|--|-----------|-----------|----------|----------------|----------------|
| | All Observations | | | | Excl. Outliers | Excl. Capitals | All Observations | | | | Excl. Outliers | Excl. Capitals |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Panel A. Negative Binomial ML Estimates | | | | | | | | | | | | |
| SPLIT (Partitioning) | 0.4764*** | 0.3604** | 0.4930*** | 0.5073*** | 0.5383*** | 0.5109*** | 1.0389*** | 0.7799*** | 0.6505*** | 0.6119** | 0.6119** | 0.6063** |
| Double-clustered s.e. | (0.1369) | (0.1990) | (0.1542) | (0.1498) | (0.1381) | (0.1524) | (0.1858) | (0.2997) | (0.2763) | (0.2794) | (0.2794) | (0.2776) |
| SPIL (Adjacent Split) | 0.2969 | 0.2734 | 0.2444 | 0.2808 | 0.3421 | 0.2031 | 0.3394 | 0.2884 | 0.1758 | 0.1646 | 0.1646 | 0.0925 |
| Double-clustered s.e. | (0.3246) | (0.3904) | (0.3182) | (0.3076) | (0.3137) | (0.3217) | (0.5658) | (0.5069) | (0.4212) | (0.3767) | (0.3767) | (0.3922) |
| Log Likelihood | -3899.33 | -3687.30 | -3542.93 | -3524.40 | -3402.42 | -3221.60 | -1488.05 | -1363.22 | -1327.74 | -1320.62 | -1320.62 | -1259.95 |
| Panel B. Linear Probability Model (LPM) Estimates | | | | | | | | | | | | |
| SPLIT (Partitioning) | 0.0852*** | 0.0833*** | 0.0896*** | 0.0918*** | 0.0889*** | 0.0877*** | 0.1269*** | 0.0967* | 0.1025** | 0.0987** | 0.0987** | 0.0967** |
| Double-clustered s.e. | (0.0307) | (0.0289) | (0.0300) | (0.0304) | (0.0304) | (0.0313) | (0.0457) | (0.0499) | (0.0475) | (0.0461) | (0.0461) | (0.0465) |
| SPIL (Adjacent Split) | 0.0556 | 0.0761 | 0.0836 | 0.0953 | 0.0984* | 0.0957 | 0.1653** | 0.1305* | 0.1310* | 0.1294* | 0.1294* | 0.116 |
| Double-clustered s.e. | (0.0619) | (0.0600) | (0.0596) | (0.0583) | (0.0581) | (0.0602) | (0.0721) | (0.0776) | (0.0766) | (0.0727) | (0.0727) | (0.0734) |
| Adjusted R-squared | 0.283 | 0.448 | 0.457 | 0.461 | 0.462 | 0.46 | 0.286 | 0.467 | 0.471 | 0.479 | 0.479 | 0.474 |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | No | No | Yes | Yes | Yes | Yes | No | No | Yes | Yes | Yes | Yes |
| Geographic Controls | No | No | No | Yes | Yes | Yes | No | No | No | Yes | Yes | Yes |
| Country Fixed Effects | No | Yes | Yes | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 1199 | 1165 | 579 | 579 | 579 | 579 | 579 | 568 |

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in Panel A is the total number of civil conflict incidents at each country-ethnic homeland over the period 1997-2010. The dependent variable in Panel B is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced conflict and zero otherwise over the period 1997-2010. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (2)-(6) and (8)-(12) include country fixed effects (constants not reported). The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). The specifications in columns (5) and (11) exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (6) and (12) exclude country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 12: Ethnic Partitioning and Civil Conflict Intensity. Baseline Country-Fixed-Effects Estimates over 1997-2010

| | All Ethnicity-Country Homelands | | | | | Ethnicity-Country Homelands Close to the National Border | | | | |
|-----------------------|---------------------------------|----------------------------|------------------|------------------------|---------------------------|--|----------------------------|------------------|------------------------|---------------------------|
| | Deadly Incidents | Deadly Incidents Indicator | Total Casualties | Duration All Incidents | Duration Deadly Incidents | Deadly Incidents | Deadly Incidents Indicator | Total Casualties | Duration All Incidents | Duration Deadly Incidents |
| | NB-ML | LPM | NB-ML | Poisson - ML | Poisson - ML | NB-ML | LPM | NB-ML | Poisson - ML | Poisson - ML |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| SPLIT (Partitioning) | 0.4484*** | 0.0963*** | 0.7665*** | 0.2240*** | 0.2087** | 0.6639*** | 0.1180** | 1.4786*** | 0.3156** | 0.5145*** |
| Double-clustered s.e. | (0.1668) | (0.0344) | (0.2094) | (0.0717) | (0.0892) | (0.2681) | (0.0478) | (0.4783) | (0.1385) | (0.1744) |
| SPIL (Adjacent Split) | 0.2718 | 0.1101* | -0.0442 | 0.2486* | 0.4181** | 0.3548 | 0.1707** | 0.5772 | 0.2709 | 0.4806 |
| Double-clustered s.e. | (0.3554) | (0.0566) | (0.4216) | (0.1433) | (0.1904) | (0.4350) | (0.0678) | (0.6975) | (0.2506) | (0.3287) |
| Log Likelihood | -2452.01 | — | -3913.681 | -2487.784 | -1971.27 | -892.18 | — | -1441.2 | -932.983 | -710.46 |
| Adjusted R-squared | — | 0.381 | — | — | — | — | 0.417 | — | — | — |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1199 | 1212 | 1212 | 579 | 579 | 575 | 579 | 579 |

The table reports estimates associating civil conflict with ethnic partitioning at the country-ethnicity homeland level. In columns (1) and (6) the dependent variable is the total number of deadly civil conflict incidents at each country-ethnic homeland over the sample period (1997-2010). These models are estimated with the negative binomial ML model. In columns (2) and (7) the dependent variable is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced at least one deadly conflict incident over the period 1997-2013 and zero otherwise. These columns give linear probability model estimates. In columns (3) and (8) the dependent variable is the total number of fatalities at each country-ethnic homeland over 1997-2010. These models are estimated with the negative binomial ML model. For the estimation we exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. In columns (4) and (9) the dependent variable is the number of years that each country-ethnic homeland has experienced conflict over the period 1997-2010. These columns give Poisson ML estimates. In columns (5) and (10) the dependent variable is the number of years that each country-ethnic homeland has experienced deadly conflict (at least one casualty) over the period 1997-2010. These columns give Poisson ML estimates. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include country fixed effects (constants not reported) and a rich set of controls. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 13: Ethnic Partitioning and Main Aspects of Civil Conflict over 1997-2010

| | All Ethnicity-Country Homelands | | | | | | Ethnicity-Country Homelands Close to the National Border | | | | | |
|-----------------------|---------------------------------|------------|-------------------|------------|------------------|------------|--|------------|-------------------|------------|------------------|------------|
| | Battles | | Civilian Violence | | Riots & Protests | | Battles | | Civilian Violence | | Riots & Protests | |
| | <u>NB-ML</u> | <u>LPM</u> | <u>NB-ML</u> | <u>LPM</u> | <u>NB-ML</u> | <u>LPM</u> | <u>NB-ML</u> | <u>LPM</u> | <u>NB-ML</u> | <u>LPM</u> | <u>NB-ML</u> | <u>LPM</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| SPLIT (Partitioning) | 0.5360*** | 0.0839** | 0.4036*** | 0.0654** | 0.0917 | -0.0002 | 0.6111** | 0.0976** | 0.5448*** | 0.0717 | 0.0077 | 0.004 |
| Double-clustered s.e. | (0.1540) | (0.0366) | (0.1453) | (0.0315) | (0.1868) | (0.0290) | (0.3186) | (0.0420) | (0.2344) | (0.0450) | (0.2769) | (0.0382) |
| SPIL (Adjacent Split) | 0.3345 | 0.0497 | 0.0247 | 0.0438 | 0.0896 | 0.0235 | 0.3491 | 0.085 | -0.3973 | 0.0666 | 0.5937 | 0.026 |
| Double-clustered s.e. | (0.3193) | (0.0525) | (0.3709) | (0.0576) | (0.2963) | (0.0399) | (0.4126) | (0.0620) | (0.4326) | (0.0674) | (0.4842) | (0.0595) |
| Log Likelihood | -2511.84 | — | -2420.43 | — | -1620.2 | — | -957.28 | — | -868.121 | — | -469.618 | — |
| Adjusted R-squared | — | 0.453 | — | 0.393 | — | 0.413 | — | 0.461 | — | 0.442 | — | 0.375 |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 1212 | 1212 | 579 | 579 | 579 | 579 | 579 | 579 |

The table reports Negative Binomial Maximum Likelihood (ML) estimates in odd-numbered columns and linear probability model (LPM) estimates in even-numbered columns, associating the main categories of civil conflict with ethnic partitioning at the country-ethnicity homeland level. Columns (1)-(2) and (7)-(8) focus on battles. Columns (3)-(4) and (9)-(10) focus on violence against the civilian population. Columns (5)-(6) and (11)-(12) focus on riots and protests. In odd-numbered columns the dependent variable is the total number of battles (in columns (1) and (7)), violent events against the civilian population (in columns (3) and (9)) and riots and protests events (in columns (5) and (11)) over the period 1997-2010. In even-numbered columns the dependent variable is an indicator (dummy) variable for country-ethnic homelands that have experienced at least one battle (in columns (2) and (8)), at least one violent event against the civilian population (in columns (4) and (10)) and at least one event of riots and protests (in columns (6) and (12)) over the period 1997-2010 (and zero otherwise). SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating civil conflict by actor with ethnic partitioning at the country-ethnicity homeland level over the period 1997-2010. Columns (1) and (7) focus on conflict where government forces participate. Columns (2) and (8) focus on conflict where rebels and militias participate. Columns (3) and (9) focus on riots and protests. Columns (4) and (10) focus on violence against the civilian population. Columns (5) and (11) focus on military interventions of adjacent (nearby) African countries. Columns (6) and (12) focus on foreign interventions by peace-keeping forces (UN, African Union, etc.). In Panel A the dependent variable is the total number of events of each category across country-ethnic homelands over the period 1997-2010. In Panel B the dependent variable is an indicator (dummy) variable for country-ethnic homelands that have experienced at least one event from each type of civil conflict over the period 1997-2010 (and zero otherwise).

SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 15: Ethnic Partitioning and Civil Conflict. Alternative Estimation Techniques

| | All Ethnicity-Country Homelands | | | | | | Ethnicity-Country Homelands Close to the National Border | | | | | |
|--|---------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | All Observations | | | | Excl. Outliers | Excl. Capitals | All Observations | | | | Excl. Outliers | Excl. Capitals |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Panel A. Conditional Negative Binomial ML Estimates (Hausman, Hall, and Griliches (1984)) | | | | | | | | | | | | |
| SPLIT (Partitioning) | 0.1304** (0.0650) | 0.1304** (0.0650) | 0.2473*** (0.0699) | 0.2537*** (0.0711) | 0.2705*** (0.0713) | 0.2470*** (0.0742) | 0.2618** (0.1171) | 0.2618** (0.1171) | 0.4177*** (0.1241) | 0.4139*** (0.1253) | 0.4139*** (0.1253) | 0.3577*** (0.1228) |
| SPIL (Adjacent Split) | 0.3482*** (0.1172) | 0.3482*** (0.1172) | 0.4839*** (0.1153) | 0.5055*** (0.1162) | 0.5440*** (0.1189) | 0.6342*** (0.1188) | 0.3047 (0.1871) | 0.3047 (0.1871) | 0.3959** (0.1806) | 0.3425* (0.1832) | 0.3425* (0.1832) | 0.4491** (0.1914) |
| Log Likelihood | -3959.55 | -3959.55 | -3837.44 | -3826.74 | -3709.21 | -3531.86 | -1344.66 | -1344.66 | -1315.7 | -1309.81 | -1309.81 | -1242.05 |
| Observations | 1212 | 1212 | 1212 | 1212 | 1199 | 1162 | 579 | 579 | 579 | 579 | 579 | 568 |
| Panel B. Fixed-Effects Poisson ML Estimates (excl. Outliers) | | | | | | | | | | | | |
| SPLIT (Partitioning) | 0.2546** (0.1022) | 0.2372*** (0.0794) | 0.3712*** (0.1117) | 0.3892*** (0.1026) | | 0.3435*** (0.1016) | 0.5054** (0.2005) | 0.4868*** (0.1409) | 0.6954*** (0.1754) | 0.7213*** (0.1574) | | 0.5835*** (0.1589) |
| SPIL (Adjacent Split) | 0.1119 (0.2264) | 0.1443 (0.2288) | 0.2800 (0.2413) | 0.3069 (0.2346) | | 0.1975 (0.2447) | 0.4569 (0.3658) | 0.4853 (0.3998) | 0.6001* (0.3193) | 0.4922 (0.3000) | | 0.5426 (0.3986) |
| Log Likelihood | -19100 | -14200 | -12000 | -11700 | | -11000 | -7200.19 | -4746.57 | -4061.54 | -3934.52 | | -3697.48 |
| Observations | 1151 | 1151 | 1151 | 1151 | | 1122 | 570 | 570 | 570 | 570 | | 562 |
| Simple Controls | Yes | Yes | Yes | Yes | | Yes | Yes | Yes | Yes | Yes | | Yes |
| Location Controls | No | No | Yes | Yes | | Yes | No | No | Yes | Yes | | Yes |
| Geographic Controls | No | No | No | Yes | | Yes | No | No | No | Yes | | Yes |
| Country Fixed Effects | No | Yes | Yes | Yes | | Yes | No | Yes | Yes | Yes | | Yes |

Panel A reports Conditional Negative Binomial Maximum Likelihood (ML) estimates, using the method of Hausman, Hall, and Griliches (1984) to account for country-level unobservable features. Panel B reports country fixed-effects Poisson Maximum Likelihood estimates. Both panels associate civil conflict with ethnic partitioning at the country-ethnicity homeland level. The dependent variable in both panels is the total number of civil conflict incidents at each country-ethnic homeland over the period 1997-2013. For the Poisson ML estimates we exclude country-ethnic homelands where the dependent variable exceeds the 95th percentile. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands.

The specifications in columns (2)-(6) and (8)-(12) include country fixed effects (constants not reported). The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). The specifications in columns (5) and (11) in Panel A exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (6) and (12) exclude country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. Panel B reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 16: Ethnic Partitioning and Main Aspects of Civil Conflict
Alternative Estimation Techniques**

| | All Ethnicity-Country Homelands | | | Ethnicity-Country Homelands Close to the National Border | | |
|---|---------------------------------|--------------------------|-----------------------------|--|--------------------------|-----------------------------|
| | <u>Battles</u> | <u>Civilian Violence</u> | <u>Riots & Protests</u> | <u>Battles</u> | <u>Civilian Violence</u> | <u>Riots & Protests</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Conditional Negative Binomial ML Estimates | | | | | | |
| SPLIT (Partitioning) | 0.2589*** (0.0891) | 0.2229*** (0.0831) | 0.2841*** (0.0949) | 0.5204*** (0.1584) | 0.4289*** (0.1461) | 0.3690* (0.1954) |
| SPIL (Adjacent Split) | 0.2976** (0.1462) | 0.5722*** (0.1350) | 0.4158*** (0.1536) | 0.4559** (0.2256) | 0.2999 (0.2216) | 0.0743 (0.2790) |
| Log Likelihood | -2696.18 | -2676.78 | -1983.24 | -898.64 | -863.07 | -523.86 |
| Observations | 1212 | 1212 | 1212 | 570 | 577 | 579 |
| Panel B: Fixed-Effects Poisson ML Estimates (Excl. Outliers) | | | | | | |
| SPLIT (Partitioning) | 0.3247*** (0.1089) | 0.4543*** (0.1299) | 0.2827** (0.1115) | 0.7334*** (0.2577) | 0.8135*** (0.2294) | 0.1207 (0.1658) |
| SPIL (Adjacent Split) | 0.5559*** (0.1776) | 0.3573 (0.2913) | 0.245 (0.1945) | 0.5343* (0.2962) | 0.1605 (0.3755) | 0.5298 (0.4281) |
| Adjusted R-square | -5281.91 | -4415.02 | -2742.90 | -1926.75 | -1511.16 | -842.91 |
| Observations | 1151 | 1151 | 1151 | 565 | 569 | 572 |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes | Yes |

Panel A reports Conditional Negative Binomial Maximum Likelihood (ML) estimates, using the method of Hausman, Hall, and Griliches (1984) to account for country-level unobservable features. Panel B reports fixed-effects Poisson Maximum Likelihood estimates. Both panels associate civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in both panels is the total number of civil conflict incidents at each country-ethnic homeland over the period 1997-2013. For the Poisson ML estimates (in Panel B) we exclude country-ethnic homelands where the dependent variable exceeds the 95th percentile. Columns (1) and (4) focus on battles. Columns (2) and (5) focus on violence against the civilian population. Columns (3) and (6) focus on riots and protests. In both panels the dependent variable is the total number of battles (in columns (1) and (4)), violent events against the civilian population (in columns (2) and (5)) and riots and protests events (in columns (3) and (6)). SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications in Panel B include country fixed effects (constants not reported). All specifications in both panels include rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. Panel B reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 17: Ethnic Partitioning and Conflict Actors
Fixed-Effects Poisson ML Estimates (excl. Outliers)

| | All Ethnicity-Country Homelands | | | | | | Ethnicity-Country Homelands Close to the National Border | | | | | |
|-----------------------|---------------------------------|-------------------|------------------|-------------------|-----------------|----------------|--|-------------------|------------------|-------------------|-----------------|----------------|
| | Government Forces | Rebels & Militias | Riots & Protests | Civilian Violence | Nearby External | Other External | Government Forces | Rebels & Militias | Riots & Protests | Civilian Violence | Nearby External | Other External |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| SPLIT (Partitioning) | 0.4290*** | 0.3531*** | 0.2970** | 0.4688*** | 0.5570** | -0.4439 | 0.8499*** | 0.9103*** | 0.1910 | 0.7984*** | 0.6471 | 0.2613 |
| Double-clustered s.e. | (0.1200) | (0.1114) | (0.1481) | (0.1262) | (0.2666) | (0.2746) | (0.2548) | (0.1624) | (0.2242) | (0.2116) | (0.3952) | (0.5226) |
| SPIL (Adjacent Split) | 0.5140*** | 0.3607** | 0.2584 | 0.3611 | 0.7794** | -0.8987* | 0.4196 | 0.2414 | 0.6637 | 0.0192 | 0.8836** | -1.0169 |
| Double-clustered s.e. | (0.2229) | (0.1827) | (0.2146) | (0.2787) | (0.2747) | (0.4835) | (0.3467) | (0.2529) | (0.5411) | (0.3652) | (0.4303) | (0.7607) |
| Log Likelihood | -6196.64 | -8442.94 | -2810.11 | -4482.95 | -1010.49 | -399.206 | -2037.086 | -2786.1 | -888.711 | -1517.88 | -410.167 | -111.053 |
| Observations | 1149 | 1151 | 1151 | 1151 | 1116 | 621 | 568 | 567 | 572 | 568 | 421 | 173 |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

The table reports country-fixed effects Poisson Maximum Likelihood (ML) estimates associating civil conflict by actor with ethnic partitioning at the country-ethnicity homeland level. Columns (1) and (7) focus on conflict where government forces participate. Columns (2) and (8) focus on conflict where rebels and militias participate. Columns (3) and (9) focus on riots and protests. Columns (4) and (10) focus on violence against the civilian population. Columns (5) and (11) focus on military interventions of adjacent (nearby) African countries. Columns (6) and (12) focus on foreign interventions by peace-keeping forces (UN, African Union, etc.). The dependent variable is the total number of events of each category across country-ethnic homelands over the period 1997-2013, excluding country-ethnic homelands where the dependent variable exceeds the 95th percentile. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 18: Ethnic Partitioning and Civil Conflict
Not Accounting for Spillovers**

| | All Ethnicity-Country Homelands | | | Ethnicity-Country Homelands Close to the National Border | | |
|--|---------------------------------|-----------------------|-----------------------|--|-----------------------|-----------------------|
| | <u>All</u> | <u>Excl. Outliers</u> | <u>Excl. Capitals</u> | <u>All</u> | <u>Excl. Outliers</u> | <u>Excl. Capitals</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Negative Binomial ML Estimates | | | | | | |
| SPLIT (Partitioning) | 0.4543*** | 0.4755*** | 0.4376*** | 0.5044** | 0.5040** | 0.5009** |
| Double-clustered s.e. | (0.1319) | (0.1287) | (0.1327) | (0.2585) | (0.2576) | (0.2566) |
| Log Likelihood | -3752.875 | -3637.696 | -3462.719 | -1379.347 | -1371.668 | -1323.013 |
| Panel B: Linear Probability Model (LPM) Estimates | | | | | | |
| SPLIT (Partitioning) | 0.0830*** | 0.0831*** | 0.0828*** | 0.0804* | 0.0804* | 0.0809* |
| Double-clustered s.e. | (0.0303) | (0.0306) | (0.0306) | (0.0487) | (0.0487) | (0.0489) |
| Adjusted R-square | 0.458 | 0.458 | 0.457 | 0.465 | 0.465 | 0.463 |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1199 | 1165 | 579 | 579 | 568 |

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating civil conflict with ethnic partitioning at the country-ethnicity homeland level. The dependent variable in Panel A is the total number of main civil conflict incidents (excluding riots and protests) at each country-ethnic homeland over the period 1997-2013. The dependent variable in Panel B is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced conflict and zero otherwise (in Panel B) over the period 1997-2013.. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers).The specifications in columns (2) and (5) exclude country-ethnic homelands where conflict events exceed the 99th percentile. The specifications in columns (3) and (6) exclude country-ethnic homelands where capital cities fall.

All specifications include country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 19: Ethnic Partitioning and Civil Conflict
Alternative Measure of Ethnic Partitioning

| | All Ethnicity-Country Homelands | | | Ethnicity-Country Homelands Close to the National Border | | |
|--|---------------------------------|------------------------------|------------------------------|--|------------------------------|------------------------------|
| | <u>All</u> (1) | <u>Excl. Outliers</u> (2) | <u>Excl. Capitals</u> (3) | <u>All</u> (4) | <u>Excl. Outliers</u> (5) | <u>Excl. Capitals</u> (6) |
| Panel A: Negative Binomial ML Estimates | | | | | | |
| SPLIT-5PC (Partitioning) | 0.4977*** | 0.5385*** | 0.5024*** | 0.7601*** | 0.7597*** | 0.7390*** |
| Double-clustered s.e. | (0.1107) | (0.1029) | (0.1184) | (0.2185) | (0.2174) | (0.2160) |
| Log Likelihood | -3752.14 | -3636.22 | -3461.39 | -1377.07 | -1369.39 | -1321.07 |
| Panel B: Linear Probability Model Estimates | | | | | | |
| SPLIT-5PC (Partitioning) | 0.0670** | 0.0701*** | 0.0691** | 0.0649 | 0.0649 | 0.0633 |
| Double-clustered s.e. | (0.0268) | (0.0269) | (0.0275) | (0.0495) | (0.0495) | (0.0502) |
| Adjusted R-square | 0.456 | 0.456 | 0.455 | 0.462 | 0.462 | 0.460 |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1199 | 1165 | 579 | 579 | 568 |

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating civil conflict with ethnic partitioning at the country-ethnicity homeland level. The dependent variable in Panel A is the total number of main civil conflict incidents (excluding riots and protests) at each country-ethnic homeland over the period 1997-2013. The dependent variable in Panel B is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced conflict and zero otherwise (in Panel B) over the period 1997-2013. SPLIT-5PC is an indicator variable that identifies partitioned ethnicities as those with at least 5% of the historical homeland falling into more than one contemporary country. The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). The specifications in columns (2) and (5) exclude country-ethnic homelands where conflict events exceed the 99th percentile. The specifications in columns (3) and (6) exclude country-ethnic homelands where capital cities fall. All specifications include country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 20: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Controlling for Unobservables. Distance to the Border. 3rd-order Polynomial

| | All Ethnic Homelands | | | | Ethnic Homelands close to the National Border | | | |
|--|----------------------|----------------|-----------------|--------------|---|----------------|-----------------|--------------|
| | <u>All Events</u> | <u>Battles</u> | <u>Violence</u> | <u>Riots</u> | <u>All Events</u> | <u>Battles</u> | <u>Violence</u> | <u>Riots</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Panel A: Fixed-Effects Negative Binomial ML Estimates | | | | | | | | |
| SPLIT (Partitioning) | 0.5234*** | 0.4483*** | 0.5425*** | 0.1792 | 0.5712*** | 0.4637 | 0.5158*** | 0.0484 |
| Double-clustered s.e. | (0.1294) | (0.1631) | (0.1501) | (0.1740) | (0.2183) | (0.2949) | (0.1803) | (0.2437) |
| SPIL (Adjacent Split) | 0.5301** | 0.4553 | 0.4276 | 0.5437** | 0.4490 | 0.4201 | -0.0115 | 0.9463* |
| Double-clustered s.e. | (0.2654) | (0.2900) | (0.3478) | (0.2347) | (0.3580) | (0.3702) | (0.3718) | (0.4930) |
| Log Likelihood | -4108.06 | -2917.84 | -2874.90 | -2198.11 | -1510.36 | -1067.17 | -1000.35 | -648.12 |
| Panel B: Fixed-Effects Linear Probability (LPM) Estimates | | | | | | | | |
| SPLIT (Partitioning) | 0.0911*** | 0.1015*** | 0.0606 | 0.0192 | 0.0838* | 0.0863* | 0.0674 | 0.0205 |
| Double-clustered s.e. | (0.0327) | (0.0387) | (0.0370) | (0.0326) | (0.0467) | (0.0495) | (0.0447) | (0.0548) |
| SPIL (Adjacent Split) | 0.1504*** | 0.0678 | 0.1784*** | 0.0836 | 0.2409*** | 0.1692*** | 0.1860*** | 0.084 |
| Double-clustered s.e. | (0.0406) | (0.0449) | (0.0583) | (0.0541) | (0.0570) | (0.0610) | (0.0722) | (0.0760) |
| Adjusted R-square | 0.445 | 0.465 | 0.422 | 0.441 | 0.49 | 0.458 | 0.435 | 0.422 |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 579 | 579 | 579 | 579 |

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating civil conflict with ethnic partitioning at the country-ethnicity homeland level. The dependent variable in Panel A is the total number of main civil conflict incidents at each country-ethnic homeland over the period 1997-2013. The dependent variable in Panel B is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced main conflict and zero otherwise over the period 1997-2013. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. Columns (1) and (5) focus on all types of civil conflict. Columns (2) and (6) focus on battles. Columns (3) and (7) focus on violence against the civilian population. Columns (4) and (8) focus on riots and protests. In Panel A the dependent variable is the total number of all conflict events (in columns (1) and (5)), battles (in columns (2) and (6)), violent events against the civilian population (in columns (3) and (7)) and riots and protests events (in columns (4) and (8)). In Panel B the dependent variable is an indicator (dummy) variable for country-ethnic homelands that have experienced at least one conflict event (in columns (1) and (5)), a battle (in columns (2) and (6)), at least one violent event against the civilian population (in columns (3) and (7)) and at least one event of riots and protests (in columns (4) and (8)) over the period 1997-2013 (and zero otherwise). The specifications in columns (5)-(8) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include a country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. All specifications include a third-order polynomial on distance from the centroid of each country-ethnic homeland to the national border. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 21: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Controlling for Unobservables. Distance to the Border. 4th-order Polynomial

| | All Ethnic Homelands | | | | Ethnic Homelands close to the National Border | | | |
|--|----------------------|----------------|-----------------|--------------|---|----------------|-----------------|--------------|
| | <u>All Events</u> | <u>Battles</u> | <u>Violence</u> | <u>Riots</u> | <u>All Events</u> | <u>Battles</u> | <u>Violence</u> | <u>Riots</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Panel A: Fixed-Effects Negative Binomial ML Estimates | | | | | | | | |
| SPLIT (Partitioning) | 0.5183*** | 0.4332*** | 0.5299*** | 0.1782 | 0.5787*** | 0.4806 | 0.5112*** | 0.0496 |
| Double-clustered s.e. | (0.1339) | (0.1630) | (0.1481) | (0.1733) | (0.2254) | (0.3056) | (0.1831) | (0.2459) |
| SPIL (Adjacent Split) | 0.5233** | 0.4350 | 0.4104 | 0.5416** | 0.4313 | 0.4046 | -0.0023 | 0.9455* |
| Double-clustered s.e. | (0.2651) | (0.2838) | (0.3444) | (0.2510) | (0.3666) | (0.3817) | (0.3767) | (0.4914) |
| Log Likelihood | -4108.008 | -2917.585 | -2874.157 | -2198.107 | -1509.827 | -1066.38 | -1000.234 | -648.12 |
| Panel B: Fixed-Effects Linear Probability (LPM) Estimates | | | | | | | | |
| SPLIT (Partitioning) | 0.0917*** | 0.1006*** | 0.0635* | 0.0193 | 0.0841* | 0.0869* | 0.0676 | 0.0204 |
| Double-clustered s.e. | (0.0325) | (0.0385) | (0.0368) | (0.0332) | (0.0466) | (0.0495) | (0.0443) | (0.0539) |
| SPIL (Adjacent Split) | 0.1511*** | 0.0667 | 0.1820*** | 0.0837 | 0.2418*** | 0.1712*** | 0.1867*** | 0.0839 |
| Double-clustered s.e. | (0.0402) | (0.0446) | (0.0584) | (0.0558) | (0.0577) | (0.0600) | (0.0719) | (0.0761) |
| Adjusted R-square | 0.445 | 0.465 | 0.423 | 0.441 | 0.491 | 0.461 | 0.436 | 0.422 |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 579 | 579 | 579 | 579 |

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating civil conflict with ethnic partitioning at the country-ethnicity homeland level. The dependent variable in Panel A is the total number of main civil conflict incidents at each country-ethnic homeland over the period 1997-2013. The dependent variable in Panel B is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced main conflict and zero otherwise over the period 1997-2013. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. Columns (1) and (5) focus on all types of civil conflict. Columns (2) and (6) focus on battles. Columns (3) and (7) focus on violence against the civilian population. Columns (4) and (8) focus on riots and protests. In Panel A the dependent variable is the total number of all conflict events (in columns (1) and (5)), battles (in columns (2) and (6)), violent events against the civilian population (in columns (3) and (7)) and riots and protests events (in columns (4) and (8)). In Panel B the dependent variable is an indicator (dummy) variable for country-ethnic homelands that have experienced at least one conflict event (in columns (1) and (5)), a battle (in columns (2) and (6)), at least one violent event against the civilian population (in columns (3) and (7)) and at least one event of riots and protests (in columns (4) and (8)) over the period 1997-2013 (and zero otherwise). The specifications in columns (5)-(8) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include a country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. All specifications include a fourth-order polynomial on distance from the centroid of each country-ethnic homeland to the national border. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 22: Ethnic Partitioning and Civil Conflict
Ethnic Family Fixed Effects Specifications**

| | All Ethnicity-Country Homelands | | | Ethnicity-Country Homelands Close to the National Border | | |
|-----------------------------|---------------------------------|------------------|-------------------|--|------------------|-------------------|
| | <u>All Incidents</u> | <u>Indicator</u> | <u>Duration</u> | <u>All Incidents</u> | <u>Indicator</u> | <u>Duration</u> |
| | | <u>Conflict</u> | <u>Conflict</u> | | <u>Conflict</u> | <u>Conflict</u> |
| | <u>NB-ML</u> | <u>LPM</u> | <u>Poisson-ML</u> | <u>NB-ML</u> | <u>LPM</u> | <u>Poisson-ML</u> |
| (1) | (2) | (3) | (4) | (5) | (6) | |
| SPLIT (Partitioning) | 0.3605** | 0.0657** | 0.1804*** | 0.5296** | 0.0879 | 0.3629*** |
| Double-clustered s.e. | (0.1487) | (0.0327) | (0.0681) | (0.2419) | (0.0604) | (0.1422) |
| SPIL (Adjacent Split) | 0.3516 | 0.1264*** | 0.1245 | 0.7025 | 0.2425*** | 0.4645* |
| Double-clustered s.e. | (0.2687) | (0.0484) | (0.1424) | (0.4192) | (0.0806) | (0.2504) |
| Log Likelihood | -3991.88 | — | -2560.59 | -1431.32 | — | -941.82 |
| Adjusted R-square | — | 0.49 | — | — | 0.58 | — |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Ethnic Family Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1199 | 1165 | 579 | 579 | 579 |

The table reports estimates associating civil conflict with ethnic partitioning at the country-ethnicity homeland level conditioning on both country fixed effects and ethnic family fixed effects. In columns (1) and (4) the dependent variable is the total number of main civil conflict incidents (excluding riots and protests) at each country-ethnic homeland over the sample period (1997-2013). These models are estimated with the negative binomial ML model. In columns (2) and (5) the dependent variable is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced at least one main conflict incident over the period 1997-2013 and zero otherwise. These columns give linear probability model estimates. In columns (3) and (6) the dependent variable is the number of years that each country-ethnic homeland has experienced a main conflict over the period 1997-2013. These models are estimated with Poisson ML. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include a country fixed effects (constants not reported), ethnic family fixed effects and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 23. Ethnic Partitioning and Civil Conflict
Sensitivity Analysis. Dropping Iteratively Each African Region**

| | Excluding | | | | | | | | | |
|--|-------------------|----------------------|-------------------|----------------------|-------------------|----------------------|-------------------|----------------------|-------------------|-----------------------|
| | North | | South | | West | | East | | Central | |
| | <u>All</u> (1) | <u>Border</u> (2) | <u>All</u> (3) | <u>Border</u> (4) | <u>All</u> (5) | <u>Border</u> (6) | <u>All</u> (7) | <u>Border</u> (8) | <u>All</u> (9) | <u>Border</u> (10) |
| Panel A: Fixed-Effects Negative Binomial ML Estimates | | | | | | | | | | |
| SPLIT (Partitioning) | 0.4801*** | 0.5310*** | 0.4999*** | 0.5370** | 0.5421*** | 0.8586*** | 0.5328*** | 0.5950** | 0.2538* | 0.2071 |
| Double-clustered s.e. | (0.1281) | (0.2335) | (0.1346) | (0.2422) | (0.1296) | (0.1649) | (0.1420) | (0.2775) | (0.1481) | (0.2274) |
| SPIL (Adjacent Split) | 0.3463 | 0.2237 | 0.3773 | 0.2725 | 0.2222 | -0.3947 | 0.6717** | 0.4515 | 0.3254 | 0.5987 |
| Double-clustered s.e. | (0.3132) | (0.3538) | (0.3154) | (0.3428) | (0.3311) | (0.3373) | (0.3063) | (0.4129) | (0.3272) | (0.4258) |
| Log Likelihood | -3494.47 | -1342.93 | -3347.17 | -1259.38 | -2710.46 | -882.393 | -2657.01 | -1048.1 | -2750.78 | -956.73 |
| Panel B: Fixed-Effects Linear Probability Model (LPM) Estimates | | | | | | | | | | |
| SPLIT (Partitioning) | 0.0817** | 0.0763 | 0.0817** | 0.0786 | 0.0736** | 0.0976* | 0.0803** | 0.0984* | 0.1014*** | 0.1028* |
| Double-clustered s.e. | (0.0340) | (0.0516) | (0.0332) | (0.0499) | (0.0324) | (0.0551) | (0.0357) | (0.0534) | (0.0353) | (0.0542) |
| SPIL (Adjacent Split) | 0.1489*** | 0.2185*** | 0.1418*** | 0.2125*** | 0.1311** | 0.1874** | 0.1582*** | 0.2293*** | 0.1440*** | 0.2477*** |
| Double-clustered s.e. | (0.0504) | (0.0595) | (0.0501) | (0.0606) | (0.0572) | (0.0786) | (0.0522) | (0.0638) | (0.0516) | (0.0642) |
| Adjusted R-square | 0.463 | 0.482 | 0.457 | 0.477 | 0.435 | 0.452 | 0.471 | 0.488 | 0.496 | 0.500 |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1127 | 556 | 1067 | 511 | 838 | 353 | 907 | 460 | 909 | 436 |

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating civil conflict with ethnic partitioning at the country-ethnicity homeland level. The dependent variable in Panel A is the total number of main civil conflict incidents (excluding riots and protests) at each country-ethnic homeland over the period 1997-2013. The dependent variable in Panel B is an dummy variable that takes on the value of one for country-ethnic homelands that have experienced conflict and zero otherwise (in Panel B) over the period 1997-2013.

In columns (1)-(2) we exclude ethnicity-country observations that fall in North Africa. In columns (3)-(4) we exclude observations that fall in South Africa. In columns (5)-(6) exclude observations that fall in West Africa. In columns (7)-(8) we exclude observations that fall in East Africa. In column (9)-(10) we exclude observations that fall in Central Africa. The regional classification follows Nunn (2008). SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. Odd-numbered specifications report estimates in the full sample of country-ethnic homelands. Even-numbered columns focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers).

All specifications include country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 24: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Accounting for Spillovers with Spatial Models**

| Weighting Matrix | Linear in Euclidian Distance | | | Quadratic in Euclidian Distance | | |
|---|------------------------------|-----------------------|-----------------------|---------------------------------|-----------------------|-----------------------|
| Spatial Model Type | Simple | Durbin | Generalized | Simple | Durbin | Generalized |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Linear Probability Model (LPM) Estimates | | | | | | |
| SPLIT (Partitioning) | 0.0883*** (0.0245) | 0.0702*** (0.0247) | 0.0880*** (0.0246) | 0.0863*** (0.0245) | 0.0698*** (0.0256) | 0.0866*** (0.0249) |
| Log Likelihood | -414.8 | -352.653 | -414.792 | -413.051 | -358.925 | -414.933 |
| rho | 0.76 [0.00] | 1.27 [0.02] | 0.746 [0.00] | 0.28 [0.00] | 0.10 [0.22] | 14.62 [0.00] |
| lamda | | | 0.068 [0.90] | | | 0.0733 [0.79] |
| Panel B: Log Linear Model Estimates [dep. var: ln(1+events)] | | | | | | |
| SPLIT (Partitioning) | 0.1628** (0.0731) | 0.1259* (0.0716) | 0.1542** (0.0732) | 0.1512** (0.0725) | 0.1101 (0.0734) | 0.1313* (0.0750) |
| Log Likelihood | -1742.28 | -1638.09 | -1738.49 | -1730.05 | -1640.76 | -1732.20 |
| rho | 0.92 [0.00] | 0.09 [0.84] | 0.90 [0.00] | 0.52 [0.00] | 0.35 [0.00] | 0.17 [0.07] |
| lamda | | | 0.82 [0.00] | | | 0.44 [0.00] |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 1212 | 1212 |

The table reports spatial auto-regressive model maximum-likelihood (ML) estimates, associating civil conflict with ethnic partitioning at the country-ethnicity level. In Panel A the dependent variable is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced conflict and zero otherwise over the period 1997-2013. In Panel B the dependent variable is the log of one plus the total number of main civil conflict incidents (excluding riots and protests) in an ethnic region within a country over the period 1997-2010. Columns (1) and (4) report spatial lag models that control for conflicts in neighbouring homelands. Columns (2) and (5) report Durbin spatial models that include as additional controls the vector of independent variables in neighbouring regions. Columns (3) and (6) report generalized spatial lag models that control both for conflicts in neighbouring regions and for the effect of the independent variables in nearby regions. In columns (1)-(3) we use a linear in Euclidian distance to the centroid of each country-ethnic region weighting matrix. In columns (4)-(6) we use a quadratic in Euclidian distance to the centroid of each country-ethnic region weighting matrix. All specifications include country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses standard errors accounting for heteroskedasticity and spatial correlation. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 25: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Accounting for Spillovers at the Country Level and at the Ethnic Family Level

| | All Ethnic Homelands | | | | | | Ethnic Homelands close to the National Border | | | | | |
|-----------------------|----------------------|------------|--------------|--------------|--------------|--------------|---|------------|--------------|--------------|--------------|-----------|
| | Main | Main | Log | Deadly | Duration | Duration | Main | Main | Log | Deadly | Duration | Duration |
| | Events | Events | (1+Main | Events | Main | Duration | Events | Events | (1+Main | Events | Main | Duration |
| | Events | Indicator | Events) | Events | Events | Deadly | Events | Indicator | Events) | Events | Events | Deadly |
| <u>NB-ML</u> | <u>LPM</u> | <u>OLS</u> | <u>NB-ML</u> | <u>PO-ML</u> | <u>PO-ML</u> | <u>NB-ML</u> | <u>LPM</u> | <u>OLS</u> | <u>NB-ML</u> | <u>PO-ML</u> | <u>PO-ML</u> | |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | |
| SPLIT (Partitioning) | 0.8591*** | 0.1176*** | 0.3667*** | 0.6554*** | 0.2456*** | 0.1909** | 1.1938*** | 0.1276*** | 0.4322*** | 0.9180*** | 0.4126*** | 0.4639*** |
| Double-clustered s.e. | (0.1638) | (0.0256) | (0.1062) | (0.1804) | (0.0902) | (0.1001) | (0.2537) | (0.0433) | (0.1140) | (0.2259) | (0.1196) | (0.1344) |
| Log Conflict Family | 0.1004*** | 0.0195*** | 0.0937*** | 0.0964*** | 0.0677*** | 0.0628*** | 0.0946** | 0.0168 | 0.0660* | 0.0531 | 0.0607* | 0.0436 |
| Double-clustered s.e. | (0.0327) | (0.0066) | (0.0252) | (0.0252) | (0.0186) | (0.0160) | (0.0388) | (0.0108) | (0.0381) | (0.0371) | (0.0330) | (0.0315) |
| Log Conflict Country | 0.1557*** | 0.0352*** | 0.1286*** | 0.1171*** | 0.0857*** | 0.0938*** | 0.1424** | 0.0369*** | 0.1294*** | 0.1066** | 0.1067*** | 0.1047*** |
| Double-clustered s.e. | (0.0458) | (0.0067) | (0.0313) | (0.0380) | (0.0257) | (0.0290) | (0.0583) | (0.0092) | (0.0350) | (0.0491) | (0.0321) | (0.0328) |
| Log Likelihood | -4042.71 | — | — | -3113.31 | -3000.30 | -2542.00 | -1636.35 | — | — | -1209.11 | -1357.25 | -1051.10 |
| Adjusted R-square | — | 0.35 | 0.48 | — | — | — | — | 0.35 | 0.41 | — | — | — |
| Region Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 1212 | 1212 | 606 | 606 | 606 | 606 | 606 | 606 |

The table reports estimates associating civil conflict with ethnic partitioning at the country-ethnic homeland level. In columns (1) and (7) the dependent variable is the total number of main civil conflict incidents (excluding riots and protests) at each country-ethnic homeland over the sample period (1997-2013). These models are estimated with the negative binomial ML model. In columns (2) and (8) the dependent variable is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced at least one deadly conflict incident over the period 1997-2013 and zero otherwise. These columns give linear probability model (LPM) estimates. In columns (3) and (9) the dependent variable is the log of one plus the total number of main civil conflict incidents (excluding riots and protests) in an ethnic region within a country over the period 1997-2013. These models are estimated with OLS. In columns (4) and (10) the dependent variable is the total number of deadly civil conflict incidents at each country-ethnic homeland over the sample period (1997-2013). These models are estimated with the negative binomial ML model.

In columns (5) and (11) the dependent variable is the number of years that each country-ethnic homeland has experienced a main conflict over the period 1997-2013. In columns (6) and (12) the dependent variable is the number of years that each country-ethnic homeland has experienced deadly conflict (at least one casualty) over the period 1997-2013. These models are estimated with Poisson ML.

SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country.

In all specifications we control for the log of one plus the total number of all conflict incidents in each ethnic family and the log of one plus the total number of all conflict incidents in each country minus conflicts in each country-ethnic area. The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include a set of (five) region fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 26: Ethnic Partitioning and Civil Conflict.
Accounting for Pre-colonial Conflict and Political Centralization**

| Historical (Pre-colonial) | All Ethnicity-Country Homelands | | | Ethnicity-Country Homelands Close to the National Border | | |
|--|---------------------------------|---------------------|----------------|--|---------------------|----------------|
| | <u>Conflict</u> | <u>Slave Trades</u> | <u>Kingdom</u> | <u>Conflict</u> | <u>Slave Trades</u> | <u>Kingdom</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Negative Binomial ML Estimates | | | | | | |
| SPLIT (Partitioning) | 0.4794*** | 0.4246*** | 0.5269*** | 0.5561*** | 0.4923** | 0.5838*** |
| Double-clustered s.e. | (0.1278) | (0.1460) | (0.1268) | (0.2200) | (0.2298) | (0.2069) |
| SPIL (Adjacent Split) | 0.4177 | 0.4107 | 0.4708 | 0.3695 | 0.2673 | 0.3724 |
| Double-clustered s.e. | (0.2930) | (0.2929) | (0.2936) | (0.3732) | (0.3510) | (0.3512) |
| Log Likelihood | -3750.59 | -3749.42 | -3744.98 | -1377.27 | -1377.85 | -1375.77 |
| Panel B: Linear Probability Model (LPM) Estimates | | | | | | |
| SPLIT (Partitioning) | 0.0820** | 0.0860** | 0.0825*** | 0.0896* | 0.0914* | 0.0889* |
| Double-clustered s.e. | (0.0308) | (0.0474) | (0.0307) | (0.0463) | (0.0331) | (0.0504) |
| SPIL (Adjacent Split) | 0.1443*** | 0.1453*** | 0.1446*** | 0.2316*** | 0.2192*** | 0.2282*** |
| Double-clustered s.e. | (0.0492) | (0.0600) | (0.0509) | (0.0610) | (0.0487) | (0.0573) |
| Adjusted R-square | 0.463 | 0.463 | 0.463 | 0.481 | 0.478 | 0.480 |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 579 | 579 | 579 |

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating civil conflict with ethnic partitioning at the country-ethnicity homeland level. The dependent variable in Panel A is the total number of main civil conflict incidents at each country-ethnic homeland over the period 1997-2013. The dependent variable in Panel B is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced main conflict and zero otherwise over the period 1997-2013. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. Specifications (1) and (4) control for an indicator for pre-colonial wars and the log distance of each homeland to the centroid of the closest war during the period 1400-1700, using data from Besley and Reynal-Querol (2014). Specifications (2) and (5) control for an indicator that takes on the value of one for ethnicities that were directly affected by the slave trades and the log of one plus the number of slaves at the ethnicity level normalized by the surface area of each homeland, using data from Nunn (2008) and Nunn and Watchekon (2011). Specifications (3) and (6) control for an indicator that takes the value of one when the historical homeland falls within the boundaries of a large pre-colonial kingdom and empire and log distance to the closest pre-colonial empire/kingdom using data from Besley and Reynal-Querol (2014). The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include a country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 27: Ethnic Partitioning and Civil Conflict
Accounting for Regional Development**

| | All Ethnic Homelands | | | | Ethnic Homelands close to the National Border | | | |
|--|----------------------|----------------|-----------------|--------------|---|----------------|-----------------|--------------|
| | <u>All Events</u> | <u>Battles</u> | <u>Violence</u> | <u>Riots</u> | <u>All Events</u> | <u>Battles</u> | <u>Violence</u> | <u>Riots</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Panel A: Fixed-Effects Negative Binomial ML Estimates | | | | | | | | |
| SPLIT (Partitioning) | 0.4556*** | 0.4367*** | 0.4209*** | 0.0871 | 0.5722*** | 0.5124* | 0.4819** | 0.0478 |
| Double-clustered s.e. | (0.1220) | (0.1466) | (0.1246) | (0.1548) | (0.2180) | (0.2811) | (0.1952) | (0.2388) |
| SPIL (Adjacent Split) | 0.4891* | 0.4818 | 0.3883 | 0.4404 | 0.4219 | 0.4101 | -0.0027 | 0.9475* |
| Double-clustered s.e. | (0.2636) | (0.3073) | (0.3530) | (0.2666) | (0.3467) | (0.3730) | (0.3648) | (0.4955) |
| Log GDP p.c. | -0.1298 | -0.3084*** | -0.2789 | 0.3569 | -0.8097*** | -0.8035** | -0.7639*** | 0.1650 |
| Double-clustered s.e. | (0.1089) | (0.1484) | (0.1689) | (0.2826) | (0.2495) | (0.3287) | (0.2869) | (0.3418) |
| Log Likelihood | -4107.81 | -2916.114 | -2874.751 | -2201.445 | -1507.682 | -1066.397 | -998.492 | -648.303 |
| Panel B: Fixed-Effects Linear Probability (LPM) Estimates | | | | | | | | |
| SPLIT (Partitioning) | 0.0818*** | 0.0907** | 0.0511 | 0.0195 | 0.0921* | 0.0891* | 0.0673 | 0.0062 |
| Double-clustered s.e. | (0.0269) | (0.0374) | (0.0322) | (0.0305) | (0.0471) | (0.0468) | (0.0464) | (0.0548) |
| SPIL (Adjacent Split) | 0.1420*** | 0.0596 | 0.1689*** | 0.0776 | 0.2367*** | 0.1635*** | 0.1717** | 0.0763 |
| Double-clustered s.e. | (0.0415) | (0.0444) | (0.0571) | (0.0543) | (0.0559) | (0.0614) | (0.0756) | (0.0784) |
| Log GDP p.c. | -0.0454 | -0.0392 | -0.0837*** | -0.0074 | -0.0948** | -0.0535 | -0.1453** | 0.0211 |
| Double-clustered s.e. | (0.0295) | (0.0331) | (0.0305) | (0.0279) | (0.0442) | (0.0600) | (0.0611) | (0.0701) |
| Adjusted R-square | 0.445 | 0.465 | 0.424 | 0.438 | 0.49 | 0.458 | 0.439 | 0.417 |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1211 | 1211 | 1211 | 1211 | 578 | 578 | 578 | 578 |

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and linear probability model (LPM) estimates in Panel B, associating civil conflict with ethnic partitioning at the country-ethnicity homeland level. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. Columns (1) and (5) focus on all types of civil conflict. Columns (2) and (6) focus on battles. Columns (3) and (7) focus on violence against the civilian population. Columns (4) and (8) focus on riots and protests. In Panel A the dependent variable is the total number of all conflict events (in columns (1) and (5)), battles (in columns (2) and (6)), violent events against the civilian population (in columns (3) and (7)) and riots and protests events (in columns (4) and (8)). In Panel B the dependent variable is an indicator (dummy) variable for country-ethnic homelands that have experienced at least one conflict event (in columns (1) and (5)), a battle (in columns (2) and (6)), at least one violent event against the civilian population (in columns (3) and (7) and at least one event of riots and protests (in columns (4) and (8)) over the period 1997-2013 (and zero otherwise). The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include a country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. All specifications include the log of GDP per capita in 2000 (data come from the G-Econ project). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

The table reports linear probability model (LPM) estimates, associating civil conflict incidence by actor with ethnic partitioning at the country-ethnicity homeland level over the period 1997-2013. Columns (1) and (7) focus on conflict where government forces participate. Columns (2) and (8) focus on conflict where rebels and militias participate. Columns (3) and (9) focus on riots and protests. Columns (4) and (10) focus on violence against the civilian population. Columns (5) and (11) focus on military interventions of adjacent (nearby) African countries. Columns (6) and (12) focus on foreign interventions by peace-keeping forces (UN, African Union, etc.). The dependent variable is an indicator (dummy) variable for country-ethnic homelands that have experienced at least one event from each type of civil conflict over the period 1997-2013 (and zero otherwise). SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include the log of GDP per capita in 2000 (data come from the G-Econ project). All specifications include country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 29: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis. Accounting for Measurement Error in the Civil Conflict Databases**

| | All Ethnic Homelands | | | | Ethnic Homelands close to the National Border | | | |
|-----------------------|----------------------|---------------------|---------------|---------------|---|---------------------|---------------|---------------|
| | <u>Trichotomous</u> | <u>Trichotomous</u> | <u>Binary</u> | <u>Binary</u> | <u>Trichotomous</u> | <u>Trichotomous</u> | <u>Binary</u> | <u>Binary</u> |
| | <u>All</u> | <u>Main</u> | <u>All</u> | <u>Main</u> | <u>All</u> | <u>Main</u> | <u>All</u> | <u>Main</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| SPLIT (Partitioning) | 0.1250** | 0.1148** | 0.0720** | 0.0661** | 0.1394 | 0.1311 | 0.0957** | 0.0894** |
| Double-clustered s.e. | (0.0500) | (0.0509) | (0.0311) | (0.0317) | (0.0869) | (0.0865) | (0.0458) | (0.0437) |
| SPIL (Adjacent Split) | 0.2260** | 0.2058** | 0.1309*** | 0.1262** | 0.2159** | 0.2008** | 0.0937* | 0.0910* |
| Double-clustered s.e. | (0.0913) | (0.0983) | (0.0455) | (0.0497) | (0.1059) | (0.1020) | (0.0550) | (0.0486) |
| Adjusted R-square | 0.564 | 0.56 | 0.476 | 0.477 | 0.567 | 0.566 | 0.485 | 0.487 |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 579 | 579 | 579 | 579 |

The table reports OLS estimates, associating civil conflict with ethnic partitioning at the country-ethnicity homeland level. The dependent variable in columns (1) and (5) is a trichotomous (0, 1, 2) civil conflict index that takes on the value of two when a country-ethnic homeland has experienced conflict according to both the ACLED and the UCDP GED databases; the index takes on the value of one if a country-ethnic homeland has experienced conflict according to either the ACLED or the UCDP GED; and the index takes on the value of zero when the country-ethnic area has not experienced conflict according to both the ACLED and the UCDP GED. The dependent variable in columns (2) and (6) is a trichotomous main civil conflict index that takes on the value of two when a homeland has experienced a main conflict according to both the ACLED and the UCDP GED databases (excluding riots and protests that are only covered by ACLED); the index takes on the value of one if a homeland has experienced main conflict according to either the ACLED or the UCDP GED; and the index takes on the value of zero when the country-ethnic area has not experienced main conflict according to both the ACLED and the UCDP GED. The dependent variable in columns (3) and (7) is a dichotomous (binary) civil conflict index that takes on the value of one when a homeland has experienced conflict according to both the ACLED and the UCDP GED databases and zero otherwise. The dependent variable in columns (4) and (8) is a dichotomous (binary) main civil conflict index that takes on the value of one when a country-ethnic homeland has experienced main conflict according to both the ACLED and the UCDP GED databases and zero otherwise (excluding riots and protests that are only covered by ACLED). SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (5)-(8) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include country fixed effects (constants not reported) and a rich set of controls. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 30: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Accounting for Measurement Error in the Civil Conflict Databases**

| | State (Government Forces) Conflict | | | | One-Sided Violence Against the Civilians | | | |
|-----------------------|------------------------------------|---------------|--|---------------|--|---------------|--|---------------|
| | All Ethnic Homelands | | Homelands close to the National Border | | All Ethnic Homelands | | Homelands close to the National Border | |
| | <u>Trichotomous</u> | <u>Binary</u> | <u>Trichotomous</u> | <u>Binary</u> | <u>Trichotomous</u> | <u>Binary</u> | <u>Trichotomous</u> | <u>Binary</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| SPLIT (Partitioning) | 0.1684*** | 0.1117*** | 0.1947*** | 0.1157** | 0.0947** | 0.0683** | 0.108 | 0.0377 |
| Double-clustered s.e. | (0.0463) | (0.0290) | (0.0685) | (0.0473) | (0.0466) | (0.0293) | (0.0725) | (0.0432) |
| SPIL (Adjacent Split) | 0.1535* | 0.0815 | 0.1148 | 0.0856 | 0.1288 | 0.0416 | 0.0872 | 0.0368 |
| Double-clustered s.e. | (0.0897) | (0.0584) | (0.1032) | (0.0709) | (0.0874) | (0.0617) | (0.1111) | (0.0784) |
| Adjusted R-square | 0.53 | 0.461 | 0.536 | 0.483 | 0.499 | 0.427 | 0.52 | 0.459 |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 579 | 579 | 1212 | 1212 | 579 | 579 |

The table reports OLS estimates, associating civil conflict with ethnic partitioning at the country-ethnicity homeland level. The dependent variable in columns (1) and (3) is a trichotomous state conflict index that takes on the value of two when a country-ethnic homeland has experienced state conflict according to both the ACLED and the UCDP GED databases; the index takes on the value of one if a country-ethnic homeland has experienced state conflict according to either the ACLED or the UCDP GED; and the index takes on the value of zero when the country-ethnic area has not experienced state conflict according to both the ACLED and the UCDP GED. The dependent variable in columns (5) and (7) is a trichotomous one-sided violence against the civilian population that takes on the value of two when a country-ethnic homeland has experienced a violence against civilians according to both the ACLED and the UCDP GED databases; the index takes on the value of one if a country-ethnic homeland has experienced violence against civilians according to either the ACLED or the UCDP GED; and the index takes on the value of zero when the country-ethnic area has not experienced violence against civilians according to both the ACLED and the UCDP GED. The dependent variable in columns (2) and (4) is a dichotomous (binary) state civil conflict index that takes on the value of one when a country-ethnic homeland has experienced state conflict according to both the ACLED and the UCDP GED databases and zero otherwise. The dependent variable in columns (6) and (8) is a dichotomous (binary) one-sided violence against the civilian population index that takes on the value of one when a country-ethnic homeland has experienced violence against civilians according to both the ACLED and the UCDP GED databases and zero otherwise. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (5)-(8) focus on country-ethnicity areas close to the national border (using as a cut-off the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include a country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 31: Heterogeneous Effects by Country-Ethnicity Features

Panel A: Negative Binomial Maximum Likelihood (NB-ML) Estimates

| | <u>Adjacent Split</u> | <u>Population Share</u> | <u>Adjacent Largest</u> | <u>Share Groups Same Family</u> | <u>Share Adjacent Same Family</u> | <u>High - Low Fractal</u> | <u>Across & Within Colony</u> | <u>2-Way Splits vs. More-than-2 Splits</u> |
|--|-----------------------|-------------------------|-------------------------|---------------------------------|-----------------------------------|---------------------------|-----------------------------------|--|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| SPLIT - High Double-clustered s.e. | 0.5394*** (0.1667) | | | | | 0.4216** (0.1817) | | |
| SPLIT - Low Double-clustered s.e. | 0.4361*** (0.1435) | | | | | 0.5635*** (0.1325) | | |
| SPLIT - Small Share Double-clustered s.e. | | 0.4927*** (0.1565) | 0.5012*** (0.1707) | 0.5897*** (0.1632) | 0.2148 (0.1462) | | | |
| SPLIT - Large Share Double-clustered s.e. | | 0.4693*** (0.1382) | 0.4609*** (0.1250) | 0.3479*** (0.1338) | 0.7711*** (0.1643) | | | |
| SPLIT - Between Colonial Powers Double-clustered s.e. | | | | | | | 0.6674*** (0.1427) | |
| SPLIT - Within Colonial Power Double-clustered s.e. | | | | | | | 0.1155 (0.1519) | |
| SPLIT - Two-Way Splits Double-clustered s.e. | | | | | | | | 0.5559*** (0.1502) |
| SPLIT - Multiple-Way Splits Double-clustered s.e. | | | | | | | | 0.3572*** (0.1325) |
| SPIL (Adjacent Split) Double-clustered s.e. | 0.4883* (0.2727) | 0.4184* (0.2330) | 0.4155* (0.2337) | 0.4149* (0.2375) | 0.4234* (0.2294) | 0.4532* (0.2543) | 0.4462* (0.2291) | 0.4567* (0.2390) |
| Log Likelihood | -3750.46 | -3750.61 | -3750.58 | -3749.48 | -3743.86 | -3570.01 | -3744.46 | -3749.61 |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 1212 | 1131 | 1212 | 1212 |

Appendix Table 31: Heterogeneous Effects by Country-Ethnicity Features

Panel B: Linear Probability Model (LPM) Estimates

| | <u>Adjacent Split</u> | <u>Population Share</u> | <u>Adjacent Largest</u> | <u>Share Groups Same Family</u> | <u>Share Adjacent Same Family</u> | <u>High - Low Fractal</u> | <u>Across & Within Colony</u> | <u>2-Way Splits vs. More-than-2 Splits</u> |
|---------------------------------|-----------------------|-------------------------|-------------------------|---------------------------------|-----------------------------------|---------------------------|-----------------------------------|--|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| SPLIT - High | 0.0869** | | | | | 0.0405 | | |
| Double-clustered s.e. | (0.0396) | | | | | (0.0413) | | |
| SPLIT - Low | 0.0801** | | | | | 0.1125*** | | |
| Double-clustered s.e. | (0.0334) | | | | | (0.0320) | | |
| SPLIT - Small Share | | 0.0781** | 0.0377 | 0.0600 | 0.0719* | | | |
| Double-clustered s.e. | | (0.0395) | (0.0462) | (0.0383) | (0.0385) | | | |
| SPLIT - Large Share | | 0.0890*** | 0.1273*** | 0.1117*** | 0.0945** | | | |
| Double-clustered s.e. | | (0.0328) | (0.0274) | (0.0382) | (0.0434) | | | |
| SPLIT - Between Colonial Powers | | | | | | | 0.0885** | |
| Double-clustered s.e. | | | | | | | (0.0378) | |
| SPLIT - Within Colonial Power | | | | | | | 0.0732** | |
| Double-clustered s.e. | | | | | | | (0.0361) | |
| SPLIT - Two-Way Splits | | | | | | | | 0.0623 |
| Double-clustered s.e. | | | | | | | | (0.0403) |
| SPLIT - Multiple-Way Splits | | | | | | | | 0.1163*** |
| Double-clustered s.e. | | | | | | | | (0.0402) |
| SPIL (Adjacent Split) | 0.1494** | 0.1429*** | 0.1389*** | 0.1491*** | 0.1445*** | 0.1427*** | 0.1454*** | 0.1387*** |
| Double-clustered s.e. | (0.0628) | (0.0488) | (0.0493) | (0.0488) | (0.0482) | (0.0491) | (0.0479) | (0.0509) |
| adjusted R-square | 0.46 | 0.47 | 0.47 | 0.46 | 0.46 | 0.46 | 0.46 | 0.47 |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 1212 | 1131 | 1212 | 1212 |

The table reports negative binomial maximum likelihood (NB-ML) estimates (in Panel A) and linear probability model estimates (in Panel B), civil conflict with ethnic partitioning. In Panel A the dependent variable is the total number of all conflict over the period 1997-2013. In Panel B the dependent variable is an indicator (dummy) variable for country-ethnic homelands that have experienced at least one conflict event over the period 1997-2013. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country.

The coefficient on SPLIT is allowed to differ by the share of adjacent groups that are split, in column (1) by the population share of each partition relative to the country's population, in column (2) by the population size of your co-ethnics on the other side of the border relative to the population of the neighboring country, in column (3) by share of adjacent groups that belong to the same ethnic family, in column (4) by the share of groups in the country that belong to the same ethnic family, in column (5) by the share of adjacent groups that belong to the same ethnic family, in column (6) by whether the group is partitioned by a relatively straight border (low fractal) or a relatively squiggly one (high fractal), in column (7) by whether the group is split between or within a colonial power, and in column (8) by whether the group is split between 2 or more countries (3, 4, 5 or 6).

SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. All specifications include a vector of country fixed effects (constants not reported). The specifications in all columns condition on a rich set of controls that includes: log of land area, the log of population in 1960, an indicator for lakes, an indicator for rivers, distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country, an indicator for country-ethnic areas that are by the sea coast, an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 32: Heterogeneous Effects by Country Features

Panel A: Negative Binomial Maximum Likelihood (NB-ML) Estimates

| | <u>Ethnic</u> <u>Fractionalization</u> (1) | <u>Ethnic</u> <u>Fragmentation</u> (2) | <u>Ethnic</u> <u>Polarization</u> (3) | <u>Landlocked -</u> <u>Coastal</u> (4) | <u>Large - Small</u> <u>Land Area</u> (5) |
|--|--|--|---|--|---|
| SPLIT - High Double-clustered s.e. | 0.5403*** (0.1366) | 0.5524*** (0.1368) | 0.1195 (0.2454) | | |
| SPLIT - Low Double-clustered s.e. | 0.2299 (0.2480) | 0.2811 (0.2363) | 0.6252*** (0.1385) | | |
| SPLIT - Landlocked Double-clustered s.e. | | | | 0.8424*** (0.2393) | |
| SPLIT - Coastal Double-clustered s.e. | | | | 0.2250 (0.1708) | |
| SPLIT - Big Countries Double-clustered s.e. | | | | | 0.6007*** (0.1306) |
| SPLIT - Small Countries Double-clustered s.e. | | | | | 0.1003 (0.2113) |
| SPIIL (Adjacent Split) Double-clustered s.e. | 0.4736* (0.2702) | 0.4654* (0.2658) | 0.4753* (0.2683) | 0.5058* (0.2685) | 0.5047* (0.2588) |
| Log Likelihood | -4107.74 | -4103.57 | -4107.63 | -4105.28 | -4106.12 |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 1212 |

Appendix Table 32: Heterogeneous Effects by Country Features

Panel B: Linear Probability Model (LPM) Estimates

| | <u>Ethnic</u> <u>Fractionalization</u> (1) | <u>Ethnic</u> <u>Fragmentation</u> (2) | <u>Ethnic</u> <u>Polarization</u> (3) | <u>Landlocked -</u> <u>Coastal</u> (4) | <u>Large - Small</u> <u>Land Area</u> (5) |
|--|--|--|---|--|---|
| SPLIT - High Double-clustered s.e. | 0.0638** (0.0317) | 0.0713** (0.0327) | 0.0760** (0.0330) | | |
| SPLIT - Low Double-clustered s.e. | 0.1296*** (0.0472) | 0.1014** (0.0446) | 0.0933*** (0.0377) | | |
| SPLIT - Landlocked Double-clustered s.e. | | | | 0.1512*** (0.0345) | |
| SPLIT - Coastal Double-clustered s.e. | | | | 0.0479 (0.0336) | |
| SPLIT - Big Countries Double-clustered s.e. | | | | | 0.0820*** (0.032) |
| SPLIT - Small Countries Double-clustered s.e. | | | | | 0.0817* (0.047) |
| SPIL (Adjacent Split) Double-clustered s.e. | 0.1398*** (0.0416) | 0.1478*** (0.0412) | 0.1499*** (0.0406) | 0.1451*** (0.0408) | 0.1443*** (0.0408) |
| adjusted R-square | 0.447 | 0.445 | 0.446 | 0.445 | 0.445 |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Rich Set of Controls | Yes | Yes | Yes | Yes | Yes |
| Observations | 1212 | 1212 | 1212 | 1212 | 1212 |

The table reports negative binomial maximum likelihood (NB-ML) estimates (in Panel A) and linear probability model estimates (in Panel B), civil conflict with ethnic partitioning. In Panel A the dependent variable is the total number of all conflict events over the period 1997-2013. In Panel B the dependent variable is an indicator (dummy) variable for country-ethnic homelands that have experienced at least one conflict event over the period 1997-2013. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. In columns (1)-(3) the coefficient on SPLIT is allowed to differ for high and low ethnic fragmentation countries, using as cut-offs the median value of the ethnolinguistic fractionalization index of Alesina et al. (2003) in column (1), the median value of the ethnic fragmentation index of Desmet et al. (2012) in column (2) and the median value of the ethnic polarization index of Desmet et al. (2012) in column (3). In column (4) the coefficient on SPLIT differs for landlocked countries and countries with access to the sea. In column (5) the coefficient on SPLIT differs for large and small countries using as a cut-off the median value of land area. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. All specifications include a vector of country fixed effects (constants not reported). The specifications in all columns condition on a rich set of controls that includes: log of land area, the log of population in 1960, an indicator for lakes, an indicator for rivers (simple controls), distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country, an indicator for country-ethnic areas that are by the sea coast, an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 33: Ethnic Partitioning, Ethnic-based Political Discrimination, and Major Ethnic Civil Wars. Sensitivity Analysis. Alternative Index of Ethnic Partitioning. Linear Probability Model Estimates

| | Ethnic Discrimination | | | | Ethnic War | | | | |
|--------------------------|-----------------------|-----------|-----------|-----------|------------|-----------|-----------|-----------|-----------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| SPLIT-5PC (Partitioning) | 0.1244*** | 0.0933*** | 0.0912*** | 0.0864*** | 0.1057** | 0.1451*** | 0.1408*** | 0.1180*** | 0.0654* |
| Double-clustered s.e. | (0.0311) | (0.0266) | (0.0274) | (0.0282) | (0.0472) | (0.0392) | (0.0424) | (0.0388) | (0.0368) |
| SPIL (Adjacent Split) | | | | 0.0203 | | | | 0.2518** | 0.2415** |
| Double-clustered s.e. | | | | (0.0759) | | | | (0.1140) | (0.1087) |
| Political Discrimination | | | | | | | | | 0.6083*** |
| Double-clustered s.e. | | | | | | | | | (0.1420) |
| Adjusted R-square | 0.027 | 0.479 | 0.498 | 0.523 | 0.017 | 0.425 | 0.428 | 0.488 | |
| Observations | 593 | 593 | 593 | 593 | 593 | 593 | 593 | 593 | 593 |
| Countries | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 |
| Country Fixed Effects | No | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes |
| Simple Controls | No | No | Yes | Yes | No | No | Yes | Yes | Yes |
| Location Controls | No | No | Yes | Yes | No | No | Yes | Yes | Yes |
| Geographic Controls | No | No | No | Yes | No | No | No | Yes | Yes |

The table reports linear probability model estimates, associating ethnic-based political discrimination and ethnic wars with ethnic partitioning. The dependent variable in columns (1)-(4) is a dummy variable that takes on the value of one if an ethnicity has experienced discrimination from the central government for at least one year over the period 1960-2010. The dependent variable in columns (5)-(9) is a dummy variable that takes on the value of one if an ethnicity has experienced a major or minor civil war with an explicit ethnic dimension over the period 1960-2010. Data on ethnic wars and political discrimination from the national government come from the Ethnic Power Relations (EPR) database (Wimmer, Cederman and Min (2009)). SPLIT 5PC is an indicator variable that identifies partitioned ethnicities as those with at least 5% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. The specifications in columns (2)-(4) and (6)-(9) include a vector of country fixed effects (constants not reported). The specifications in columns (3)-(4) and (7)-(9) include log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers (simple controls). The specifications in columns (4), (8) and (9) include a set of location and geographic controls. The specification in column (9) conditions on the political discrimination dummy that takes on the value of one if an ethnicity has experienced discrimination from the central government for at least one year over the period 1960-2010. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 34: Ethnic Partitioning, Ethnic-based Political Discrimination, and Major Ethnic Civil Wars.
Ethnic Power Relations (EPR) Database.
Sensitivity Analysis. Excluding Each Time a Different African Region**

| | Excluding | | | | | | | | | |
|-----------------------|-------------------|-----------------|-------------------|-----------------|-------------------|-----------------|-------------------|-----------------|-------------------|-----------------|
| | North | | East | | West | | Central | | South | |
| | <u>Ethnic War</u> | <u>Discrim.</u> | <u>Ethnic War</u> | <u>Discrim.</u> | <u>Ethnic War</u> | <u>Discrim.</u> | <u>Ethnic War</u> | <u>Discrim.</u> | <u>Ethnic War</u> | <u>Discrim.</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| SPLIT (Partitioning) | 0.1210*** | 0.0766** | 0.1123** | 0.0681** | 0.0906* | 0.0561* | 0.0803** | 0.0624* | 0.1326*** | 0.0824*** |
| Double-clustered s.e. | (0.0428) | (0.0309) | (0.0451) | (0.0333) | (0.0529) | (0.0335) | (0.0389) | (0.0352) | (0.0417) | (0.0269) |
| SPIL (Adjacent Split) | 0.3033** | 0.0487 | 0.2043* | -0.0326 | 0.3144** | -0.045 | 0.1329 | 0.0777 | 0.3016** | 0.0402 |
| Double-clustered s.e. | (0.1227) | (0.0874) | (0.1172) | (0.0631) | (0.1535) | (0.1007) | (0.0866) | (0.0809) | (0.1298) | (0.0864) |
| Adjusted R-squared | 0.485 | 0.520 | 0.505 | 0.492 | 0.516 | 0.568 | 0.481 | 0.546 | 0.487 | 0.519 |
| Observations | 551 | 551 | 434 | 434 | 395 | 395 | 470 | 470 | 522 | 522 |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

The table reports linear probability mode estimates, associating ethnic civil wars and ethnic-based political discrimination with ethnic partitioning. The dependent variable in odd-numbered columns is a dummy variable that takes on the value of one if an ethnicity has experienced a major or minor civil war with an explicit ethnic dimension over the period 1960-2010. The dependent variable in even-numbered columns is a dummy variable that takes on the value of one if an ethnicity has experienced discrimination from the central government for at least one year over the period 1960-2010. Data on ethnic wars and ethnic-based political discrimination from the national government come from the Ethnic Power Relations (EPR) database (Wimmer, Cederman and Min (2009)). In columns (1)-(2) we exclude North Africa. In columns (3)-(4) we exclude East Africa. In columns (5)-(6) we exclude West Africa. In columns (7)-(8) we exclude Central Africa. In column (9)-(10) we exclude South Africa. The regional classification follows Nunn (2008). SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. All specifications include a vector of country fixed effects (constants not reported) and a rich set of controls. The simple set of controls includes log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 35: Ethnic Partitioning, Ethnic-based Political Discrimination, and Major Ethnic Civil Wars.
Ethnic Power Relations (EPR) Database.
Sensitivity Analysis. Excluding Each Time a Different African Region**

| | Excluding | | | | | | | | | |
|-----------------------|-------------------|-----------------|-------------------|-----------------|-------------------|-----------------|-------------------|-----------------|-------------------|-----------------|
| | North | | East | | West | | Central | | South | |
| | <u>Ethnic War</u> | <u>Discrim.</u> | <u>Ethnic War</u> | <u>Discrim.</u> | <u>Ethnic War</u> | <u>Discrim.</u> | <u>Ethnic War</u> | <u>Discrim.</u> | <u>Ethnic War</u> | <u>Discrim.</u> |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| SPLIT-5PC | 0.1355*** | 0.0958*** | 0.1147*** | 0.0765** | 0.1182** | 0.0779** | 0.0752** | 0.0771** | 0.1329*** | 0.0925*** |
| Double-clustered s.e. | (0.0420) | (0.0315) | (0.0432) | (0.0362) | (0.0542) | (0.0368) | (0.0364) | (0.0327) | (0.0417) | (0.0239) |
| SPIL (Adjacent Split) | 0.2996** | (0.0472) | 0.2045* | (0.0320) | 0.3060** | (0.0504) | (0.1298) | (0.0760) | 0.2908** | (0.0326) |
| Double-clustered s.e. | (0.1209) | (0.0874) | (0.1176) | (0.0632) | (0.1484) | (0.1017) | (0.0869) | (0.0818) | (0.1282) | (0.0873) |
| Adjusted R-squared | 0.487 | 0.522 | 0.505 | 0.493 | 0.520 | 0.571 | 0.481 | 0.548 | 0.487 | 0.521 |
| Observations | 551 | 551 | 434 | 434 | 395 | 395 | 470 | 470 | 522 | 522 |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Simple Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Geographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

The table reports linear probability mode estimates, associating ethnic civil wars and ethnic-based political discrimination with ethnic partitioning. The dependent variable in odd-numbered columns is a dummy variable that takes on the value of one if an ethnicity has experienced a major or minor civil war with an explicit ethnic dimension over the period 1960-2010. The dependent variable in even-numbered columns is a dummy variable that takes on the value of one if an ethnicity has experienced discrimination from the central government for at least one year over the period 1960-2010. Data on ethnic wars and ethnic-based political discrimination from the national government come from the Ethnic Power Relations (EPR) database (Wimmer, Cederman and Min (2009)). In columns (1)-(2) we exclude North Africa. In columns (3)-(4) we exclude East Africa. In columns (5)-(6) we exclude West Africa. In columns (7)-(8) we exclude Central Africa. In column (9)-(10) we exclude South Africa. The regional classification follows Nunn (2008). SPLIT 5PC is an indicator variable that identifies partitioned ethnicities as those with at least 5% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands. All specifications include a vector of country fixed effects (constants not reported) and a rich set of controls. The simple set of controls includes log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 36. DHS Descriptive Patterns and Summary Statistics
Ethnic Partitioning at the Identity Level and at the Location Level**

Panel A: Descriptives

| | | Ethnic Homeland (Location) | | |
|-----------------|-----------------|----------------------------|-------------|-------|
| | | Non-Partitioned | Partitioned | Total |
| Ethnic Identity | Non-Partitioned | 36694 | 13256 | 49950 |
| | Partitioned | 12590 | 25631 | 38221 |
| Total | | 49284 | 38887 | 88171 |

Panel B: Summary Statistics

| variable | Obs. | mean | st. dev. | median | min | max |
|------------------------------------|-------|-------|----------|--------|-----|-----|
| Composite Wealth Index | 88171 | 3.168 | 1.433 | 3 | 1 | 5 |
| Education | 88171 | 1.723 | 1.554 | 1 | 0 | 5 |
| Ethnic Partitioning Index | 88171 | 0.433 | 0.496 | 0 | 0 | 1 |
| Location Ethnic Partitioning Index | 88171 | 0.441 | 0.497 | 0 | 0 | 1 |
| Non-Indigenous Indicator | 88171 | 0.612 | 0.487 | 1 | 0 | 1 |

Panel A reports descriptive patterns in the Demographic and Health Surveys (DHS) sample. Panel B reports summary statistics for the main variables employed in the empirical analysis using data from the Demographic and Health Surveys (DHS). The data cover 20 countries. The countries and interview years are Benin in 2001, Burkina Faso in 2010, Central African Republic in 1994, Ethiopia in 2011, Ghana in 2008, Guinea in 2005, Kenya in 2008, Mali in 2006, Mozambique in 2011, Malawi in 2010, Namibia in 2000, Niger in 1998, Senegal in 2010, Sierra Leone in 2008, Togo in 1998, Uganda in 2011, the Democratic Republic of Congo in 2007, and Zambia in 2007. The Data Appendix gives detailed variable definitions and data sources.

**Appendix Table 37: The Long-Run Effects of Ethnic Partitioning.
Enumeration-area (Village) Fixed Effects Estimates**

| | All Observations (Individuals) | | | | Observations close to the Border | | | |
|-----------------------|--------------------------------|------------|-----------|-----------|----------------------------------|-----------|-----------|----------|
| | Wealth Index | | Education | | Wealth Index | | Education | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Ethnic Partitioning | -0.0665*** | -0.0644*** | -0.0856** | -0.0799** | -0.0541** | -0.0506** | -0.0680* | -0.0582 |
| Double-clustered s.e. | (0.0196) | (0.0196) | (0.0364) | (0.0347) | (0.0237) | (0.0243) | (0.0407) | (0.0396) |
| Non-Indigenous | 0.0761*** | 0.0729*** | 0.039 | 0.0328 | 0.1027*** | 0.1022*** | 0.0212 | 0.0164 |
| Double-clustered s.e. | (0.0208) | (0.0205) | (0.0341) | (0.0321) | (0.0243) | (0.0238) | (0.0460) | (0.0430) |
| Adjusted R-square | 0.694 | 0.696 | 0.463 | 0.482 | 0.656 | 0.658 | 0.416 | 0.435 |
| Observations | 88171 | 88171 | 88171 | 88171 | 44090 | 44090 | 44090 | 44090 |
| Individual Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Location Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Village Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

The table reports OLS estimates associating DHS composite wealth index (in columns (1)-(2) and (5)-(6)) and an education index (in columns (3)-(4) and (7)-(8)) with ethnic partitioning at the individual level. The composite wealth index is calculated by the DHS team in each country via a principal component method using easy-to-collect data on a household's ownership of selected assets (e.g., televisions and bicycles), materials used for housing construction and public good access (e.g., type of water access, electrification, and sanitation). The ordered education index (range from 0 to 5) assigns the respondent's level of formal schooling into six categories. A score of 0 indicates "no education"; a score of 1 indicates "incomplete primary"; a score of 2 indicates "complete primary"; a score of 3 indicates "incomplete secondary"; a score of 4 indicates "complete secondary"; and a score of 5 indicates "higher education". The ethnic partitioning index (SPLIT-ID) takes on the value of one for individuals that self-identify with a partitioned ethnicity. The non-indigenous indicator takes on the value of one for individuals residing outside their ethnicity's ancestral homeland and takes on the value of zero for individuals residing in their ethnicity's ancestral homeland ("non-movers" and "movers"). All specifications include a vector of DHS enumeration area (village/town/city) fixed effects. The set of individual controls includes a vector of 10 age-bracket fixed effects, a vector of 6 marital-status fixed effects, and a vector of 7 religion fixed effects. The set of location controls includes the distance of each individual to the capital city, the distance to the sea, the distance to the national border and an indicator that takes on the value of one if the individual resides in the capital city. The specifications in columns (5)-(8) focus on individuals residing close to the national border (using as a cut-off the median distance; 80 kilometers). The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the ethnicity and the ethnic homeland dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 38: The Long-Run Effects of Ethnic Partitioning on Individual Well-Being and Education. DHS Data Channels; Location and Identity. Looking at "Movers"

| | Composite Wealth Index | | | | | | Education | | | | | |
|-------------------------|------------------------|-----------|-------------------------------------|----------|---------------------------------|----------|------------|-----------|-------------------------------------|----------|---------------------------------|-----------|
| | All Movers | | Movers in Non-Partitioned Homelands | | Movers in Partitioned Homelands | | All Movers | | Movers in Non-Partitioned Homelands | | Movers in Partitioned Homelands | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Partitioning - Identity | -0.3070** | -0.2469** | -0.4410** | -0.2791* | -0.1787* | -0.1637* | -0.2897* | -0.1887** | -0.4192** | -0.1950* | -0.1764** | -0.1496** |
| Double-clustered s.e. | (0.1440) | (0.1055) | (0.1921) | (0.1512) | (0.0974) | (0.0841) | (0.1499) | (0.0928) | (0.2115) | (0.1112) | (0.0841) | (0.0758) |
| Partitioning - Location | -0.2804** | -0.0609 | | | | | -0.1582* | 0.0021 | | | | |
| Double-clustered s.e. | (0.1354) | (0.0973) | | | | | (0.0833) | (0.0675) | | | | |
| Adjusted R-squared | 0.042 | 0.195 | 0.067 | 0.231 | 0.072 | 0.176 | 0.135 | 0.251 | 0.137 | 0.290 | 0.152 | 0.221 |
| Observations | 53992 | 53992 | 30606 | 30606 | 23386 | 23386 | 53992 | 53992 | 30606 | 30606 | 23386 | 23386 |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Individual Controls | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Location Controls | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |

The table reports OLS estimates, associating DHS composite wealth index (in columns (1)-(6)) and an education index (in columns (7)-(12)) with ethnic partitioning at the individual level focusing on individuals that reside outside their ethnicity's ancestral homeland ("movers"). The composite wealth index is calculated by the DHS team in each country via a principal component method using easy-to-collect data on a household's ownership of selected assets (e.g., televisions and bicycles), materials used for housing construction and public good access (e.g., type of water access, electrification, and sanitation). The ordered education index (range from 0 to 5) assigns the respondent's level of formal schooling into six categories. A score of 0 indicates "no education"; a score of 1 indicates "incomplete primary"; a score of 2 indicates "complete primary"; a score of 3 indicates "incomplete secondary"; a score of 4 indicates "complete secondary"; and a score of 5 indicates "higher education".

The ethnic partitioning identity index (SPLIT-ID) takes on the value of one for individuals that identify with a partitioned ethnicity. The location based ethnic partitioning index (SPLIT-LOC) takes on the value of one for individuals that reside in ethnic homelands that have been partitioned by the national border and zero otherwise. The specifications in columns (3)-(4) and (9)-(10) restrict estimation to individuals (movers) residing in non-partitioned ethnic homelands. The specifications in columns (5)-(6) and (11)-(12) restrict estimation to individuals (movers) residing in partitioned ethnic homelands.

All specifications include a vector of country fixed effects (constants not reported). The set of individual controls includes a vector of 10 age-bracket fixed effects, a vector of 6 marital-status fixed effects, and a vector of 7 religion fixed effects. The set of location controls includes the distance of each individual to the capital city, the distance to the sea, the distance to the national border and an indicator that takes on the value of one if the individual resides in the capital city. The specifications in columns (3)-(4) and (9)-(10) focus on individuals residing close to the national border (using as a cut-off the median distance; 80 kilometers). The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the ethnicity and the ethnic homeland dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 39: The Long-Run Effects of Ethnic Partitioning.
Examining Persistence**

| | DHS Composite Wealth Index | | | | Education | | | |
|---|----------------------------|-----------|-----------|-----------|-----------|----------|-----------|----------|
| | Young | Old | Young | Old | Young | Old | Young | Old |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Baseline Ethnic Partitioning - Identity | -0.2175** | -0.1809** | -0.2378** | -0.1803** | -0.1719* | -0.1279* | -0.1996** | -0.1351* |
| Double-clustered s.e. | (0.0988) | (0.0874) | (0.0998) | (0.0903) | (0.0893) | (0.0751) | (0.0936) | (0.0780) |
| Ethnic Partitioning - Location | | | 0.0784 | -0.0025 | | | 0.1072 | 0.032 |
| Double-clustered s.e. | | | (0.0827) | (0.0805) | | | (0.0701) | (0.0582) |
| Non-Indigenous Indicator | 0.1754** | 0.1815*** | 0.1789** | 0.1814*** | 0.1122 | 0.0875 | 0.117 | 0.0881 |
| Double-clustered s.e. | (0.0732) | (0.0657) | (0.0719) | (0.0657) | (0.0846) | (0.0601) | (0.0837) | (0.0598) |
| Adjusted R-squared | 0.163 | 0.173 | 0.164 | 0.173 | 0.262 | 0.236 | 0.262 | 0.236 |
| Observations | 43283 | 44888 | 43283 | 44888 | 43283 | 44888 | 43283 | 44888 |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Individual Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

The table reports OLS estimates, associating the DHS composite wealth index (in columns (1)-(4)) and an education index (in columns (5)-(8)) with ethnic partitioning at the individual and at the location level. Odd-numbered columns report estimates restricting estimation to individuals born after 1977 (young). Even-numbered columns report estimates restricting estimation to individuals born before (or on) 1977 (old). The composite wealth index is calculated by the DHS team in each country via a principal component method using easy-to-collect data on a household's ownership of selected assets (e.g., televisions and bicycles), materials used for housing construction and public good access (e.g., type of water access, electrification, and sanitation). The ordered education index (range from 0 to 5) assigns the respondent's level of formal schooling into six categories. A score of 0 indicates "no education"; a score of 1 indicates "incomplete primary"; a score of 2 indicates "complete primary"; a score of 3 indicates "incomplete secondary"; a score of 4 indicates "complete secondary"; and a score of 5 indicates "higher education". The ethnic partitioning identity index (SPLIT-ID) takes on the value of one for individuals that identify with a partitioned ethnicity. The location based ethnic partitioning index (SPLIT-LOC) takes on the value of one for individuals that reside in ethnic homelands that have been partitioned by the national border and zero otherwise. The non-indigenous indicator takes on the value of one for individuals residing outside their ethnicity's ancestral homeland and takes on the value of zero for individuals residing in their ethnicity's ancestral homeland ("movers"). All specifications include a vector of country fixed effects (constants not reported), a set of individual controls and a set of location controls. The set of individual controls includes a vector of 10 age-bracket fixed effects, a vector of 6 marital-status fixed effects, and a vector of 7 religion fixed effects. The set of location controls includes the distance of each individual to the capital city, the distance to the sea, the distance to the national border and an indicator that takes on the value of one if the individual resides in the capital city. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the ethnicity and the ethnic homeland dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table A: Partitioned Ethnicities

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

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

| | | | | | | | |
|-----------|------|-----|---|-----------|------|-----|---|
| 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 3.1 gives details on our approach in identifying partitioned ethnicities. In the empirical analysis we include all partitioned groups (larger than 100sq km). There are some minor discrepancies in the sample employed in the empirical analysis that emerge because we drop country-ethnic areas with zero population in the first post independence census. Moreover in the empirical analysis we do not consider the Comorians, as neither ACLED nor UCDP nor EPR nor DHS cover the Comoros.